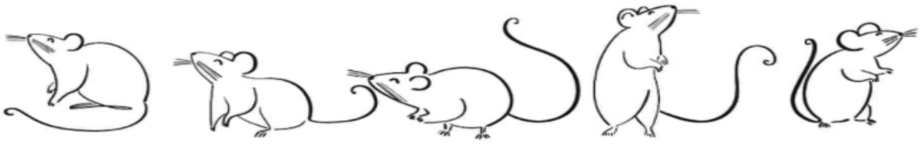
Pavlovian Conditioning

How excitation and inhibition **D**etermine **I**deomotion

Adela Florentina Iliescu



A thesis presented for the degree of

Doctor of Philosophy

Cardiff University

School of Psychology

November, 2019

Summary

Associative theories assume a simple ordinal mapping between the strength of an association between a conditioned stimulus (CS) and an unconditioned stimulus (US) and conditioned behaviour in an experimental preparation. Recent studies that have taken multiple measures of conditioned behaviour challenge this assumption. The purpose of this thesis is a better understanding of the nature of these individual differences in Pavlovian conditioning, combining empirical evidence with theoretical development. It has been observed that simple auto-shaping procedures result in marked individual differences: some rats show learning by interacting with the sign (sign-trackers, STs), others by interacting with the food-well or the goal (goal-trackers, GTs). In Chapter 2, I examined the sensitivity of these two behaviours (sign-tracking and goal-tracking) to changing contingencies. In both STs and GTs, US-oriented behaviour was more sensitive to contingency changes than CS-oriented behaviour. Most attempts to explain this dissociation have appealed to a dual-mechanisms approach. In Chapter 3, I present a new theoretical model, HeiDI, which integrates learning and performance from a single-process perspective. In Chapter 4, I examine two of these predictions. The first prediction relates to how the US-value affects the distribution of conditioned behaviour. According to HeiDI, a higher US-value will result in higher levels of goal-tracking in contrast with lower US value. Experiment 3 suggested that that a higher US-value results in more CS-oriented behaviour, however this was not replicated in Experiment 4. The second prediction concerns an analysis of the feature positive effect, where the discrimination emerges more readily for a feature positive design. HeiDI predicts that a feature positive effect should be more evident in CS-oriented behaviour. Experiments 4 and 5 addressed this prediction, however the animals did not show learning in the feature negative design. The implications of the new findings and the new model are discussed in Chapter 5.

Acknowledgements

I would like to first express my gratitude to my supervisors, Professor Rob Honey and Professor Dominic Dwyer, for the guidance, patience, encouragement and support during these wonderful years. I cannot express how lucky I feel to have you as my mentors.

Sergio, thank you for supporting me through the good and the bad. You have given me the power to overcome all the obstacles. You have been my inspiration and I would not have achieved this without you.

I would also like to thank all my colleagues from the BNL, Patricia, Sophie, and outside the 'banana world', Silvia, Oana, for being incredibly supportive and understanding. Thank you for the 'shoulder'.

Mami, mulţumesc pentru sprijinul acordat în toţi aceşti ani. Mulţumesc pentru exemplul acordat, pentru cã mi-ai fost mereu alãturi, pentru sfaturi şi inspiraţie.

Declaration

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

Signed: Adela F. Iliescu

Date 18/09/2019

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed: Adela F. Iliescu

Date 18/09/2019

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed: Adela F. Iliescu

Date 18/09/2019

STATEMENT 3

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

Signed: Adela F. Iliescu

Date 18/09/2019

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

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

Signed: Adela F. Iliescu

Date 18/09/2019

Large parts of Chapter 2, 3 and 5 appear in:

Iliescu, A. F., Hall, J., Wilkinson, L. S., Dwyer, D. M., & Honey, R. C. (2018). The nature of phenotypic variation in Pavlovian conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition, 44*(4), 358–369.

Honey, R.C., Dwyer, D.M., & Iliescu, A.F. (2019). HeiDI: A model for Pavlovian learning and performance with reciprocal associations. *Psychological Review* (in revision).

I contributed to the conceptual development of HeiDI, and to its formal presentation for publication. I also built the HeiDI app, which generates the simulated results, and derived predictions from it. The accuracy of these predictions was tested in Chapter 4.

Contents

List of Figures

Chapter 1

General Introduction

"Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select—doctor, lawyer, artist, merchant-chief and, yes, even beggarman and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors" (Watson, 1924, p. 104).

Watson believed that experience or training was the principal factor that contributes to what a person becomes. It is difficult to believe that anyone could not have accepted at least some version of Watson's claim, that different training allows people to become *doctors, lawyers, artists, merchant-chiefs*. According to this view, all individual differences in behaviour reflect different learning experiences. However, the focus of this thesis is on what is unacknowledged by Watson's statement, namely that material individual differences still emerge, in spite of shared experience and training (e.g., Flagel, Akil, & Robinson, 2009; Patitucci, Nelson, Dwyer, & Honey, 2016).

A key question that this thesis will address, is whether these individual differences reflect differences in the content of learning (what is learnt in a given situation) as opposed to possible differences in the mechanisms of learning between subjects. In order to address this key question, one needs to understand what is learnt (the content of learning) in a given situation, what is the mechanism underlying that learning, and how what has been learnt is translated into behaviour (Rescorla, 1988). Before addressing these issues, I will first briefly review the concept of learning from an associative perspective, before turning to individual differences in a ubiquitous form of learning (Pavlovian conditioning) and outline contemporary accounts of such differences.

The formal study of animal learning has a long history and it is possible to identify two main approaches that have led to what we now know. The first approach was concerned with the study of adaptive behaviour from an evolutionary point of view, contrasting adaptive specialisation of abilities to the environment of different species (Mackintosh, 1974, p. 2). This comparative approach focussed on naturally occurring acquired behaviours (e.g., the development of birdsong and filial imprinting), and can be contrasted with the study of the processes underlying modifications of behaviour in procedures chosen to involve more arbitrary stimuli, responses and the relationships between them. The latter approach is exemplified by the experimental investigation of changes in behaviour and the conditions when these changes occur in laboratory tasks, which have attempted to reveal general principles of learning. This form of investigation involved attempts to understand the processes involved in learning, or the general principles in operation, through an analysis of when such changes in behaviour occur. In this case, changes in behaviour can be seen as both the thing to be understood and the means of doing so, which has attendant problems. For example, if one assumes that the change in behaviour brought about by a Pavlovian conditioning procedure reflects an association between the events that have been presented (e.g., a chiming bell and food), then one needs to specify the performance rule(s) by which that association is revealed in behaviour. In fact, as I will show, theories of associative learning have typically offered only the most impoverished of performance rules.

In the field of animal learning, Mackintosh (1974) noted two historical figures, Pavlov and Thorndike, who profoundly shaped decades of research, based on their initial procedures and theoretical insights. From their original experimental work, two experimental paradigms have determined the course of most subsequent research on learning: classical conditioning (Pavlov) and instrumental conditioning (Thorndike). Note that the distinction between classical and instrumental conditioning refers to the experimental arrangements to study learning (operationally the rules that determined the delivery of a reinforcer or reward), but does not necessarily imply different learning processes. Experiments on classical conditioning are those in which the contingencies are arranged between stimuli presented independently of the actions of an animal (e.g., a stimulus preceding an outcome) and those on instrumental conditioning, the contingencies are arranged between the actions of an animal and other events (e.g., a response has the scheduled consequence of an outcome).

Thorndike's law of effect stated that any behaviours that are followed by positive consequences are likely to be repeated and the ones followed by negative consequences are likely to be suppressed (Thorndike, 1927). In instrumental conditioning, learning is generally assessed by measuring the behaviour or set of behaviours that are followed by reinforcement, and an increase (or decrease) in responding is taken as evidence of learning. In classical conditioning the subject is presented with a relationship between a stimulus and an outcome (e.g., a tone followed by food) and changes in behaviour occurs without any obvious contingencies between the behaviour itself and the outcome that follows it. In a typical classical or Pavlovian conditioning procedure, a neutral conditioned stimulus (CS) is paired with a motivationally significant unconditioned stimulus (US e.g., food), which produces unconditioned responses (UR; salivation). As a result of pairing the CS and US the CS comes to evoke a conditioned response (CR). This fundamental learning process was first formally documented by Pavlov (1927), who noticed that dogs would salivate in the presence of cues that predicted being fed (e.g., the experimenter coming into the room).

To explain why a CS comes to evoke a CR, Pavlov (1930) suggested that the animal treats the CS as it would be the US (stimulus substitution theory). This analysis predicts that the CS will come to evoke a response greatly resembling that evoked by the US, as it appeared to do in Pavlov's original studies and many of those that followed. For example, in pigeon autoshaping, when the CS is an illuminated key and the US is the delivery of grain to a food-well, the CR to the keylight resembles the UR to the delivery of food (i.e., pecking). Moreover, an experiment by Jenkins and Moore (1973) manipulated the nature of the US (grain or water) in a pigeon autoshaping procedure and showed that when the US was grain the presentation of the keylight elicited "food pecks" whereas when it was water the keylight elicited "water pecks": the CRs resembled the distinct consummatory responses to food and water. However, the principle of stimulus substitution is challenged by conditioning procedures which result in a CR that is different from the UR. For example, when rats are given pairings of a CS with electric shock (the US), while the CR is often a cessation of all activity (i.e., freezing), the US itself *initially* provokes a burst in activity (e.g., Blanchard, Fukunaga, & Blanchard, 1976). While the idea of stimulus substitution has been challenged, contemporary theories of Pavlovian conditioning, the focus of this thesis, have done little by way of addressing how (associative) learning becomes evident in acquired behaviour.

1.1 Theoretical analyses of Pavlovian learning and behaviour

More recent theoretical treatments of Pavlovian conditioning have followed Pavlov's lead in assuming that pairings of a CS with a US result in the formation of an association between the central representations of the events (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Wagner, 1981). This proposition – that the critical association is stimulus-stimulus (S-S) rather than between a stimulus and response (S-R) – carries with it a need to specify how learning is translated into performance.

The Rescorla-Wagner model remains one of the most influential in the field, and its application to the issues that are central to this thesis will be outlined next, including its approach to observed performance. The Rescorla-Wagner Model (Rescorla & Wagner, 1972) assumes that conditioned behaviour reflects the formation of an association between the representations activated by the CS and US. The presentation of the CS comes to associatively activate the representation of the US and thereby behaviour. The change in the associative strength (ΔV_{CS-US}) of a CS on a given trial is determined by the difference between the maximum associative strength supportable by a US (λ) and the pooled associative strength of all stimuli presented on that trial ($\sum V_{Total-US}$). The current associative strength of that stimulus V_{CS-US} is the sum of V_{CS-US} on the previous trial and the change in associative strength on the current trial (ΔV_{CS-US}).

The rate of learning is modulated by two learning parameters α_{CS} and β_{US} and the model assumes that α_{CS} and β_{US} are related to the physical properties (i.e., salience or intensity) of the CS and US. I will discuss the Rescorla-Model in more detail in Chapter 3. However, the resulting equation for the change in learning on a given trial proposed by this model takes the form: $\Delta \mathrm{V}_{CS-US} = \alpha_{CS}\ \beta_{US}\ (\lambda - \sum \mathrm{V}_{Total-US})$.

A consequence of adopting the view that Pavlovian conditioning involves S-S associations (as opposed to S-R associations) is that models like the one developed by Rescorla and Wagner (1972) needed to provide some analysis of how such associations are translated into observable behaviour. This need is shared by other formal associative learning models (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Wagner, 1981). Of course, associative theorists are not blind to this issue, and so have typically offered the simplifying assumption that responding elicited by the associative activation of a representation of a US is monotonically related to the strength of the underlying association. While this assumption has served for some time, I will next discuss the results of recent studies, which pose a challenge to any learning theory that assumes a monotonic relationship between learning and performance.

1.2 Individual differences in conditioned behaviour

Individual differences were a central focus of psychologists in the early 20th century, but the study of individual differences in animal learning has played little role in the development of contemporary theory (Matzel & Sauce, 2017). Perhaps this fact reflected the view that the controlled and shared environments and restricted genetic background of many laboratory animals should limit the extent of individual differences in learnt behaviour between them. However, recent evidence suggests that Pavlovian conditioning procedures can produce marked individual differences in conditioned behaviour. For example, Patitucci, Nelson, Dwyer, and Honey (2016, Experiment 2) presented rats with two levers. One lever (e.g., left lever), the reinforced lever, was inserted in the experimental chamber for 10s and was immediately followed by access to sucrose as reward. The control lever (e.g., right lever) on the other hand, was inserted in the chamber, however it was not reinforced. The rats received twenty trials of each lever for 12 days. Some rats showed that they had learnt the relationship between the reinforced lever and food by predominantly interacting with the lever (or the sign) and others by investigating the food-well (or the goal), where the reinforcer was about to be delivered. Like others, I refer to the behaviour of engaging with the lever (typically automatically recorded as lever presses) as sign-tracking, and to the animals that show a tendency towards the sign or the lever as sign-trackers (STs). Similarly, I refer to the behaviour of engaging with the food-well (typically automatically recorded as food-well entries) as goal-tracking, and to the animals that show a tendency to engage with the goal as goal-trackers (GTs) (Brown & Jenkins, 1968; Hearst & Jenkins, 1974).

In the study conducted by Patitucci et al. (2016, Experiment 2), in which the insertion of one lever was followed by sucrose and the insertion of another (control lever) was not, a median split was used to separate rats into two groups (STs and GTs) on the basis of whether their activity during the final block of training was predominantly directed towards the lever or food-well. This analysis allowed the development of the sign-tracking and goal-tracking behaviours to be traced across training. However, analysis at the level of individual rats reveals that the bias towards sign-tracking or goal-tracking was relatively continuous in nature. When lever activity or sign-tracking was used to asses learning, the STs showed better learning than the GTs; but when food-well activity or goal-tracking is used then the reverse is true. Focussing on one measure (e.g., sign-tracking) leads to the conclusion that learning had proceeded more readily in one set of rats than the other, while focussing on the second measure (e.g., goal-tracking) leads to the opposite conclusion. As it will be discussed in length in future sections, learning theories that assume a monotonic relationship between learning and acquired behaviour (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Wagner, 1981) are unable to explain this dissociation in the expression of learning: there is no simply monotonic mapping between the strength of the CS-US association and the different levels of lever press and food-well entries seen across sign- and goal-tracking animals.

These experiments highlight the importance of how learning is assessed, how it translates in performance, and the fact that one measurement alone might not be enough to draw general conclusions about what the animal has learnt in any given set of circumstances. In the following section, I discuss studies on Pavlovian conditioning where multiple measurements of conditioned behaviour have been contrasted, which suggests that sign-tracking and goal-tracking behaviours are affected differently by different types of manipulations.

1.3 Conditions that affect sign and goal tracking

This thesis is about how learning is translated into behaviour. There has been a great deal of research concerning sign- and goal-tracking, although not necessarily in the context of individual differences. I will focus next on the experimental manipulations that affect sign- and goal tracking. Consideration of the effects of these manipulations are key to the development of any general theory that integrates learning and performance; such theory will be presented in Chapter 3.

Studies by Wasserman, Franklin, and Hearst (1974) show that positive contingency between the CS and US is needed for the two critical behaviours to emerge. Wasserman et al. (1974) exposed pigeons to either a positive contingency (a lighted key CS which was always followed by grain delivery, US), negative contingency (the CS and US were explicitly unpaired), backwards pairing (the US was presented before the CS) or CS only (the CS was never followed by any grain delivery) procedure. The only birds that showed sign-tracking, by pecking at the lighted key, were the ones exposed to a positive contingency between the CS and US. On the basis of key-pecking measurement alone the behaviour of birds in the other groups could not be differentiated from each other. However, on the basis of approach-withdrawal measurements, the groups of birds could be clearly distinguished. Birds in the CS only and backward pairings group did not display any approach-withdrawal behaviour, while birds in the negative contingency group spent their time on the other side of the chamber, away from the CS. Even if Wasserman et al. (1974) did not measure goal-tracking directly, their measurements of approach and withdrawal seem to suggest that the same is true for goal-tracking, as pigeons spent their time away from the site of US delivery, making it hard for them to show any goal-tracking behaviour either. In the CS positive contingency group, grain always followed the CS, however it has been seen that partial reinforcement of as little as of 25 percent of the CS trials, produces sign-tracking behaviour (Wasserman, 1974).

Partial reinforcement increased the number of trials to the first peck, however it produced higher asymptotic sign-tracking levels (Davey & Cleland, 1982; Gottlieb, 2004, 2006). In contrast, goal-tracking was suppressed with partial reinforcement schedules when compared with a continuously reinforced group (Davey & Cleland, 1982).

Prior experience with the CS or the US affects differently sign- and goal-tracking behaviour. Interestingly, studies in pigeons showed that preexposure to CS alone trials before pairing with a US interfered with the development of sign-tracking, but not with goal-tracking responses directed at the site of food delivery (Boughner & Papini, 2003). Preexposure to the US before pairing it with a CS delayed the development of sign-tracking (Engberg, Hansen, Welker, & Thomas, 1972), but had no effect on goal-tracking behaviour (Costa & Boakes, 2009). In Section 1.4, theoretical background, I will discuss the proposed mechanisms underlying sign- and goal-tracking, which provide some insight into why some manipulations affect one behaviour more than the other.

Morrison, Bamkole, and Nicola (2015) changed the value of the US after the initial training (lever CS – sucrose US pairings) by subsequently pairing the US with illness in the absence of the CS (i.e., conditioned taste aversion). This procedure resulted in enhanced, rather than diminished, sign-tracking behaviour, but in contrast, a reduction in goal-tracking behaviour. This dissociation has commonly been used as evidence for differential mechanisms underlying these behaviours, one relying on a stimulus-stimulus association (goal-tracking) and the other on a stimulus-response association (sign-tracking). In the Theoretical Background, section 1.4, I will return to this issue at more length.

Temporal parameters between the CS onset and offset and US delivery have important effects on the strength of CRs. In general, the CR is stronger the more widely spaced the trials are and shorter the gap between CS offset and US delivery (Hearst & Jenkins, 1974). With longer gaps between trials (intertrial interval, ITI) sign-tracking behaviour is enhanced (e.g., Terrace, Gibbon, Farrell, & Baldock, 1975) and with shorter ITIs, goal-tracking is enhanced (e.g., Cinotti, Marchand, Roesch, Girard, & Khamassi, 2019; Lee et al., 2018). It has been suggested that variations in ITI could produce this difference through affecting the "predictiveness" or "informativeness" of the CS (Hearst & Jenkins, 1974). Short ITIs could lead to the CS not being very informative of when the reinforcer will be delivered. In contrast, with large ITIs the occurrence of the CS is informative for the US delivery. CS duration seems not to affect the general topography of the response, durations of 3 and 8 seconds were equally effective (Brown & Jenkins, 1968).

The physical properties of the CS, as well as the nature of the US affect the topography of the CRs evoked by the CS. Furthermore, different species display different pattern of CRs (Shettleworth, 1972). A localised CS generally evokes more sign-tracking that a diffuse one, and the response depends on the type of CS, and how it supports sign-tracking behaviour (a lever can be touched, bitten, licked, however a light can only be pecked or oriented towards, Leslie & Millenson, 1996). The location of the CS in relation with the US delivery site is also important. Generally, close spatial proximity between the CS and the US delivery site sustains more sign-tracking behaviour (Leslie & Millenson, 1996).

CRs evoked by CSs depend also on the nature of the US. Jenkins and Moore (1973) compared food and water reinforcers. Most pecks at water predictive signals were characterized by a very slight opening of the beak, sometimes licking and swallowing movements, which were irregularly spaced, sustained and relatively weak. In contrast, pecks at the food-predictive signal were with the beak wide-open, evenly spaced, strong and brief. Rackham Hearst and Jenkins (1974) observed that if the CS predicts the opportunity to mate in male pigeons, sign-tracking behaviour was characterised by approach, nodding, cooing, circling and bowing, typical courtship behaviour.

There is still much to learn about the conditions that promote or reduce sign- and goal-tracking behaviour. Most of the experiments described above assume that any measurement of any CR would lead to the same conclusion about what has been learnt (or the current value of associative strength) in any given experimental procedure. The different patterns observed when multiple measurements of different CRs have been taken challenge this assumption (e.g., Patitucci et al., 2016). In the following section I focus on different approaches that have attempted to provide an explanation of the different ways in which learning is translated into conditioned behaviour.

1.4 Theoretical background

General process models of learning (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Wagner, 1981) focus on learning and how associative strength changes under various experimental conditions and not on how learning is expressed. According to such models, the predictive relationship between the CS (e.g., lever) and US (e.g., food) should result in the formation of an association between the representations of the two allowing future presentations of the lever to activate a memory of the food. This analysis allows the exact form of conditioned response to differ from the specific nature of the response to food itself (Wagner & Brandon, 1989), but does not provide a coherent account for individual differences in conditioned behaviour. This is perhaps unsurprising given the fact that these general-process accounts were not intended to provide a characterization of individual differences. However, we have seen that different behavioural assessments of learning lead to different conclusions about learning (e.g., Patitucci et al., 2016). Most theoretical approaches have ignored the issue of how learning is translated into performance. However, there are some general principles that offer an explanation for some of the challenging results.

One of the proposed mechanisms to explain individual differences in the form Pavlovian conditioned behaviour is centred on whether it is based on a direct association between the CS and the US (stimulus-stimulus or S-S learning) or the CS is associated with the response elicited by the US (stimulus response or S-R learning). I discuss the implications of this view in the following section.

1.4.1 S-S vs S-R learning

Formal theories of associative learning assume that organisms form associations between the representations of conditioned and unconditioned stimuli. These stimulus–stimulus (S-S) associations have been contrasted with the formation of stimulus–response (S-R) associations between the processes activated by the stimulus and the motor program for generating a response (Hull, 1943).

Skinner (1948) suggested that superstitious conditioning of responses directed towards the CS might explain the development and maintenance of sign-tracking. One could argue that it is possible for pigeons being exposed to a lighted key, to perform a set of behaviours (orienting, approaching, contacting) which accidentally are associated with the reinforcer, in an operant conditioning fashion (S-R).

Consistent with view, Patitucci et al. (2016) suggested that goal-tracking could be based on an S-S learning mechanism and sign-tracking on an S-R mechanisms that operate in parallel, each of them controlling part of the CR repertoire. Goal-tracking activity seems to reflect the current status of the relationship between the lever and the reinforcer (i.e., an S-S association) as devaluating the US affects rats' tendency to enter the food-well more than lever-pressing behaviour (Morrison et al., 2015). Moreover, Patitucci et al. (2016) observed that the bias toward approaching the food-well relative to lever pressing was positively correlated with the palatability of the reinforcer, and that sating rodents with the reinforcer affects conditioned food-well activity but not lever-oriented activity. Evidence motivating the view of sign-tracking behaviour reflects a S-R association comes from studies where contingency changes affect less sign-tracking than goal-tracking (e.g., extinction, devaluation). Rats keep pressing the lever regardless of the current status of the relationship between the lever and the reinforcer, viewed as habitual behaviour (e.g., Ahrens, Singer, Fitzpatrick, Morrow, & Robinson, 2016; Morrison et al., 2015). However, the study conducted by Ahrens et al. (2016) needs to be treated with some caution as they did not measure both lever-oriented and food-well oriented behaviours in rats classified as STs and GTs. Thus it is unclear whether the dissociation in effects between ST and GT animals is truly a difference between individuals themselves or is instead a difference in the specific behaviours measured in each set of animals. I will come back to this issue in more detail in the Reinforcement Learning section and in Chapter 2, as this is the aim of the experiments presented in this chapter.

The view of sign-tracking as a stimulus-response association is challenged however by an interesting experiment conducted by Browne (1976), which suggests that sign-tracking emerges even when responses towards the CS are not reinforced. In this experiment, pigeons could only observe the relationship between the CS (key light) and US (grain) but could not physically interact with the CS or access the grains as they were separated from the CS and US by a transparent Plexiglas screen. Pigeons were exposed to either positive, negative or no correlation between the presentations of the CS and US. When subsequently given access to both the key light and grain hopper, the animals in the positive correlation group interacted with the CS far more than the no correlation group. While the birds in the negative correlation group barely interacted with the CS. Visual inspection of the birds' behaviour during the observation phase did not suggest evidence of sign-tracking responses. Responses towards the CS were never reinforced during the learning (observation) phase and yet they were rapidly expressed when given the opportunity. Browne's interpretation of his pattern of results was that it seemed implausible to suppose the animals have learnt to respond to the keylight through a process of stimulus-response learning.

Evidence from omission procedures from Williams and Williams (1969), where any peck at the lighted key would terminate the trial and prevent scheduled grains, also suggests that response – reinforcer relations do not explain or at least are not sufficient to fully explain sign-tracking behaviour towards the CS. Pigeons still pecked at the lighted key, albeit to a lesser degree, even if this behaviour was never reinforced.

These experiments suggest that sign-tracking emerges as the animals learn the relationship between the cues and the reinforcer, and secondly, sign-tracking is directed towards the most predictive cues. Sign-tracking emerges even if contact responses are never paired with a reinforcer, suggesting that stimulus-response learning is not solely responsible for this behaviour, or at least not for the acquisition stage. However, the fact that outcome devaluation or extinction (e.g., Ahrens et al., 2016; Morrison et al., 2015) affects sign-tracking less than goal-tracking might suggest a possible role for stimulus-response association in the maintenance of the behaviour. I will present an alternative theoretical analysis for these and other results in Chapter 3.

This theoretical framework (S-S vs S-R) addresses what it is learnt during Pavlovian conditioning. However, it is not able to explain why some animals engage with the sign and others with goal. The evidence suggests that S-S learning is necessary for acquisition of sign- and goal-tracking, but the role of S-R learning seems to be less clear and could potentially contribute at maintaining sign-tracking behaviour. However, it is unclear if these processes operate in parallel or, if there is a shift, when it would occur. If we assume that goal-tracking is based on S-S learning, and sign-tracking on S-R learning (at least during the later stages), it would provide a general account for why sign-tracking is less sensitive to change than goal-tracking, but still leaves unexplained the origin of these individual differences: why some animals engage in S-S and some in S-R learning. The experiments that I present in Chapter 2 aim to develop a better understanding of the different sensitivity to change of these behaviours, in STs and GTs.

1.4.2 Behaviour systems theory

Timberlake (1994) proposed that an animal comes to any learning situation with an evolutionary history that has shaped pre-existing hierarchically organized network of behaviour. A given motivational system (like the feeding system) is formed of different sub-systems (e.g., predation or social foraging), modes (General Search, Focal Search, and Handling/Consuming), and more specific stimulus-response (S-R) modules. A feeding system would be an exhaustive set of behaviours available to the animal that generally have the function of finding and consuming food. During Pavlovian conditioning, the CR varies according of which mode gets conditioned to the CS. If the CS comes to associate with a general search mode, then behaviours that are distant to the goal object should be conditioned. However, if the CS becomes associated with a more specific focal search mode, then behaviours that are more proximal to the goal object should dominate. In one experiment, Timberlake, Wahl, and King (1982) paired a ball bearing CS with food with a 6s (long) CS-US interval and they noticed that rats learned to approach, contact, and carry the ball bearing. However, when they used a 2s CS-US interval (short), the rats avoided contact with the ball bearing and directly approached the food magazine when the ball bearing CS appeared. In other words, it looked as though a more general search mode was conditioned with the long CS-US interval and a focal search mode was conditioned with a short CS-US interval.

There is the possibility that during a given conditioning procedure different animals associate the CS with different modes (e.g., a ST associates the CS with a general search mode and a GT with a more focal search mode) which then lead to the observed differences in engaging with either the CS or the goal location. Why this would be the case remains unclear. However, one possible explanation could be the related to difference in animals' perception of the timing between the CS and US. But while the behaviour systems approach has a clear value in terms of an analysis of performance, it leaves learning largely aside. It does not provide testable predictions in relation to what the animal has learnt and how it is translated into the activation of a specific system. Furthermore, it also fails to predict what mode an animal would be in given an experimental procedure or what leads to the activation of that specific system.

1.4.3 Reinforcement learning

Computational analyses of learning have paid substantial attention to the distinction between model-free and model-based forms of learning. Model-based strategies are held to generate goal-directed choices employing a model or cognitive-style representation, which is an internal map of events and stimuli from the external world (Dickinson & Balleine, 2002a). The internal model supports prospective assessment of the consequences of taking particular actions. By contrast, model-free strategies have no model of outside events, but instead merely learn by caching information about the utilities of outcomes encountered on past interactions with the environment. Learning in Model-Free systems relies on a computed reinforcement signal, the reward prediction error (Lesaint, Sigaud, Flagel, Robinson, & Khamassi, 2014). This leads to direct rules for how to behave, or propensities for performing particular actions (Dayan & Berridge, 2014). Lesaint et al. (2014) modelled individual differences in Pavlovian conditioning using this dual-process approach. The assumption is that rats use the two learning systems, employing distinct mechanisms to learn the same task, with both systems present in any given animal and the balance between them is given by a parameter (ω). For a ST, this balance favours the model-free system, and for a GT it favours a model-based system. This approach implies that observed differences between sign-tracking and goal-tracking behaviour are due to a natural tendency of the animal to engage with either the CS or the location of the US or the goal. STs are thought to attribute the CS with incentive salience, making the CS more desirable. A potential issue arises from the fact that what determines ω is unclear. There was no suggestion ω could be predicted beforehand, resulting in a something of circular argument: a ST is a ST because it displays sign-tracking behaviour (albeit, there are suggestions of other independently-assessable differences between sign- and goal-tracking animals, such as their impulsivity or sensitivity to conditioned reinforcement, e.g., Lovic, Saunders, Yager, & Robinson, 2011).

Dopamine is thought to be central for reward-related processes, but the exact nature of its role remains controversial (Schultz, 2006). Phasic neurotransmission in the mesolimbic dopamine system is initially triggered by the receipt of reward (US), but shifts to a cue that predicts a reward (CS) after associative learning (when the reward prediction error is reduced; Flagel et al., 2011; Schultz, Dayan, & Montague, 1997). Lesaint et al. (2014) examined phasic dopamine release in the nucleus accumbens, and observed dopamine was released during the acquisition of a sign-tracking but not goal-tracking response.

A potential issue with this interpretation arises from the fact that dopamine release is measured in STs while they are lever-pressing but not when they are entering the food-well, and vice versa for the GTs. From this study is unclear whether the dopamine release is related to the propensity of the animal to engage with the lever or the food-well, or of the behaviour itself (lever-pressing or food-well entries).

In any case, the link between dopamine release and sign-tracking, together with the apparently 'non-adaptive' nature of this behaviour (e.g., resistance to extinction or engaging with a CS even if this prevents the delivery of a relevant food reward for a hungry animal, Ahrens et al., 2016; Williams & Williams, 1969) has led to the proposal that STs represent a potential model for addiction or impulsivity (Bissonette et al., 2015; Lovic, Saunders, Yager, & Robinson, 2011; Ostlund & Balleine, 2008). This view assumes that cues associated with rewards not only evoke CRs but they can also become attractive and desirable in their own right, acting as incentive stimuli. Along the same lines, Lesaint et al. (2014) assume that differences in sign-tracking and goal-tracking behaviour are because of a natural tendency of the animal to attribute some CSs with incentive salience and motivate behaviour.

Lovic et al. (2011) compared STs on two tests of so-called impulsive action (a two-choice serial reaction time task and a differential reinforcement of low rates of responding task) and one test of impulsive choice (a delay discounting choice procedure). For the first test, they initially trained rats with a lever, that was also illuminated, followed by the delivery of one pellet. Subsequently, they divided the rats in STs and GTs using an index which indicated the propensity to approach either the lever or the food-well (see Meyer et al., 2012). For the subsequent phases of the experiment, they only used animals that would predominantly sign- or goal-track. During the two-choice serial reaction test they presented rats with two illuminated nose-poke ports, of which only one was followed by a reinforcer. Responses into the illuminated nose-poke port during the trial terminated the trial and resulted in the delivery of a food pellet (for the correct port). If the animal made no responses, the nose-poke port stayed illuminated for 5s and was followed by the reward. They found that STs made more premature responses than GTs. A problem with this experiment arises from the shift from Pavlovian conditioning to operant conditioning. In the first stage, the lever was reinforced regardless of the animal's response, but in the second, a nose-poke in the port (a response) is now reinforced. GTs already had a predominant behaviour of engaging with the food-well, while the CS for the STs has changed from a lever to nose-poke port. It is unclear how these changes might have affected the general tendency to interact with the food-well and the results should be to be treated with caution.

For the second test of impulsive action, another group of rats went through the same training procedure, after which they were placed on an FR 1 (fixed ratio of 1 response) operant task, until they reached 100 responses. Subsequently, they were tested on DRL-10s (differential reinforcement of low rates of responding task) schedule for 5 days and DRL-20s schedule for 15 days. On the DRL-10s and DRL-20s schedules, rats were reinforced only if at least 10s or 20s, respectively, elapsed between responses. They found that STs were less efficient in the task than GTs. Again, previous sign-tracking behaviour in terms of lever interaction might have affected the results, given the tendency in the training phase to engage with the lever. For their delay discounting choice procedure, they found no differences between sign- and goal-traking groups. On the basis of this pattern of results, the authors conclude that STs are more 'impulsive' than GTs, there are however, alternative interpretations.

Ahrens, Singer, Fitzpatrick, Morrow, and Robinson (2016) trained rats with a lever

CS followed by a reinforcer. Then they introduced periods of non-reinforcement (extinction) of the CS and noticed that STs extinguished lever presses more slowly than GTs extinguished food-well entries. They concluded that STs are less flexible in their behaviour. This conclusion was based on lever presses data for STs and food-well entries data for the GTs. As noted before, the fact that ST and GT groups were only assessed in terms of their predominant behaviour means that the difference in flexibility could be attributed either to the behaviour (lever press or food-well entry) or the animals themselves (STs or GTs). Thus to fully support the claim that this reflect differences in the animals, it would be necessary for both behaviours (lever presses and food-well entries) to be assessed in animals classified as STs and GTs. It is unclear whether slower extinction is a property of the animal or the behaviour itself, which will be discussed at length in Chapter 2.

In another experiment which investigated the link between sign-tracking and addiction, Pena-Oliver et al. (2015) examined whether selectively bred alcohol-preferring and alcohol-non-preferring rats show differential levels of impulsivity and conditioned behavioural responses to food incentives. Interestingly, they found that alcohol-preferring rats showed more goal-oriented behaviour than controls and no evidence of impulsive behaviour was found.

Part of the argument linking sign-tracking and addiction is based on the fact that Flagel et al. (2011) observed that rats classified as either STs or GTs differ in dopamine release during learning. In STs they observed an increase in CS-evoked dopamine and a decrease in US-evoked dopamine release. This pattern was not observed for GTs when assessing their food-well entries behaviour. However, they only examined STs while they lever pressed and GTs while they were interacting with the food-well. It would be interesting to assess whether an increase is observed in sign-tracking animals when they goal-track, as this would point to a definitive origin of these differences in the animal, not the behaviour per se. But, as will be shown in Chapter 2, this issue can also be addressed at a behavioural level.

The link between dopamine, addiction and sign-tracking is generally in terms of similarities of certain types of behaviours or traits assumed to be characteristic of addicted individuals (e.g., impulsivity, inflexibility). It is generally based on the idea that dopamine is involved in reward-processes and the fact that individuals addicted to harmful substances have increased concentration of dopamine release in limbic regions (Volkow, Fowler, Wang, & Swanson, 2004). However, the link with sign-tracking behaviour is somewhat unclear and based on experiments that do not assess both behaviours (sign- and goal-tracking) in animals classified as STs or GTs.

Whether there is a link or not between addiction and sign-tracking behaviour is not the central issue here, and I presented this approach because it is relevant for understanding individual differences in conditioned responses. The assumption for this line of argument is that genetic individual differences are responsible for the rat's tendency to sign or goal-track. A study by Patitucci et al. (2016) challenges this view. They presented two levers, one reinforced with food and another one sucrose. They found that the rat's bias towards sign- or goal-tracking behaviour did not correlate across these two levers. In other words, it was possible for a rat to be a ST on one lever but not the other. Differences in the balance between sign- and goal-tracking behaviour across situations, appears to be inconsistent with there being a general tendency to one type of behaviour or other, and hence the idea of 'an addictive personality' as revealed in a general propensity to sign-tracking across situations appears difficult to sustain.

There are similarities between the analyses based on the division between model-based and model-free processes on one hand, and S-S and S-R processes on the other. Model-based is based on the current value of the CS-US relationship and it is a goal-directed action (Beierholm, Anen, Quartz, & Bossaerts, 2011). Model free, in contrast, is held to be a fast system based on general responses or habits performed in any given situation. In the same way as the S-S and S-R framework, we could assume behaviour directed toward the food-well and lever are generated by independent systems (model-based or model-free) that operate in parallel. This analysis predicts that a given form of response will exhibit the same characteristics in a STs and GTs (e.g., different sensitivity to change being reflected in the behaviour). However, instead of both systems operating in parallel, the model proposed by Lesaint et al. (2014) assumes that the behaviour of a given rodent is governed predominantly by the operation of a single system (either model-based or model-free), which controls both types of behaviours. This account predicts that sign- and goal-tracking behaviours will exhibit different properties in STs and GTs. In Chapter 2 I investigate these predictions by examining both behaviours in STs and GTs by changing the contingency between the CS and US or the value of the US.

In summary, the few attempts to explain individual differences in conditioned responses, or more specifically sign- and goal tracking behaviour, have appealed to a dual mechanism which underpin each of the two behaviours. There is evidence consistent with a dual-system account where actions are controlled by either a fast habitual system or a slower more goal-directed system, based on current associations between stimuli (Beierholm et al., 2011). However, neither approach provides clear predictions of how each type of behaviour (sign- or goal-tracking) will be affected by a certain experimental manipulation. Moreover, given the continuous nature of the balance between the tendency to sign- or goal track (e.g., Patitucci et al., 2016), a dual-mechanism account would need to address the interaction between the two processes in a given rat or procedure. This issue has not been dealt with by any of the dual-processes frameworks. That said, general learning theories, which rely on a single-process (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Wagner, 1981), are unable to predict differences in responding, as performance is viewed merely as a translation of associative strength based on learning the contingency between the CS and US.

1.5 Thesis Outline

This thesis is concerned with better understanding of the nature of individual differences in Pavlovian conditioning and behaviour. I present new empirical evidence in

Chapter 2 and Chapter 4. In Chapter 2, I examine the differential sensitivity to contingency changes in both STs and GTs, measuring both sign- and goal-tracking behaviour in both groups. This evidence contributed to the development of a theoretical model, HeiDI, presented in Chapter 3. The model distances itself from a dual process approach in explaining performance and from a single-process perspective differences and, at the same time, leads to novel predictions. In Chapter 4, I examine two predictions. One prediction relates to how changes in the value of the US affect the distribution of sign- and goal tracking. The second prediction concerns an analysis of the feature positive effect, where the discrimination emerges more readily for a feature positive design, where the compound is reinforced (AB+) and the element is non-reinforced (B-), in contrast with a feature negative design, where the compound is non-reinforced (AB-) and the element is reinforced (B+).

Chapter 2

Investigating the nature of phenotypic variation

2.1 Introduction

Pavlovian conditioning is perhaps the most well-known psychological phenomenon, and its theoretical importance was evident from the point of its initial description. The use of this paradigm is widespread, particularly across the fields of behavioural and cognitive neuroscience (for a recent review, see Murphy & Honey, 2016) and behavioural genetics (e.g., Duvarci, Nader, & LeDoux, 2008; Lonsdorf et al., 2009). In the field of behavioural neuroscience, one of its principal uses has been in providing a test bed for formal theories of associative learning, which assume that organisms form associations between the representations of conditioned and unconditioned stimuli (e.g., Mackintosh, 1975; Pearce, 1994; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Wagner, 1981). These stimulus–stimulus (S-S) associations have been contrasted with the formation of stimulus–response (S-R) associations between the processes activated by the stimulus and the motor program for generating a response (Hull, 1943; Spence, 1936, 1937). The idea that two (associative) systems might underpin conditioned behaviour has clear counterparts in cognitive neuroscience as discussed in Chapter 1 (Reinforcement learning section) (e.g., Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Dayan & Berridge, 2014). Like Pavlov, the formal theories of associative learning identified above appeal to the idea that the memory or representation of one stimulus can come to excite (or to inhibit) the representation of another stimulus through an excitatory (or inhibitory) association formed between them. Unlike Pavlov, however, such theories have eschewed consideration of individual differences in Pavlovian conditioning, apart from insofar as they represent one source of variance in behavioural measures of learning (see Matzel et al., 2003). The view that there are marked individual differences in simple conditioning has been amply confirmed in more recent experiments with rodents, where the differences are perhaps more striking and well characterized. These differences too have potential translational significance (see Flagel et al., 2009; Lovic et al., 2011).

As discussed in Chapter 1 (Individual differences section), critical behavioural observations come from simple autoshaping procedures. In a procedure where rats receive brief presentations of a lever that are paired with the delivery of a reinforcer into a food-well (e.g., Patitucci et al., 2016), marked individual differences are observed in behaviour: some rats predominantly interact with the lever while others approach the food-well during the lever presentations. The basis for these different phenotypes is the central issue that is addressed here.

One analysis of individual differences in sign- and goal-tracking behaviour can be derived from the assumption that the types of associative structures described above (i.e., S-S or S-R) might be differently represented across individuals (see Lesaint et al., 2014; Patitucci et al., 2016). The general idea that acquired behaviours might be the product of different systems with distinct characteristics has a clear precedent in the context of studies of instrumental conditioning (e.g., Dickinson & Balleine, 2002b), and there are two sources of evidence that are consistent with it from studies of phenotypic differences in Pavlovian conditioned responding. First, food-well activity in rats classified as goal-trackers (GTs) declines more rapidly during an extinction procedure than does lever pressing in rats classified as sign-trackers (STs; Ahrens et al., 2016). These observations suggest that food-well activity reflects the current status of the relationship between the lever and the reinforcer (i.e., an S-S association), whereas lever-oriented behaviour was based on a S-R habit that was more resistant to changes in contingencies. Second, the bias toward approaching the food-well relative to lever pressing is positively correlated with the palatability of the reinforcer (Patitucci et al., 2016), and sating rodents with the reinforcer affects conditioned food-well activity but not lever-oriented activity (Morrison et al., 2015; Patitucci et al., 2016).

The results described in the preceding paragraph are consistent with the general idea that there are two learning systems that operate differently across rats; but there are at least two forms that this analysis could take. For example, while food-well behaviour might be the dominant response generated by the S-S system and lever-oriented behaviour the dominant response generated by the S-R system, both systems might have the capacity to generate both responses (see Lesaint et al., 2014). If a single system governed all behaviour in a given rat then both food-well and lever-press responses should exhibit the characteristic property of that system: When governed by an S-S system, activity directed toward both the lever and the food-well will change rapidly in the face of a change in contingencies; whereas when governed by an S-R system both will change relatively slowly. In principle, the accuracy of this prediction could have been assessed by Ahrens et al. (2016), but, they only presented activity directed toward the food-well for GTs and toward the lever for STs. It is not, therefore, possible to assess whether the two forms of response were affected differently in rats classified as GTs or STs. The single-system analysis just outlined is, however, challenged by the following observation: A given rodent can be classified as a GT (or ST) with respect to their behaviour on a lever that predicts one reinforcer (e.g., food pellets), but not classified in the same way on another lever that predicts a different reinforcer (e.g., sucrose; Patitucci et al., 2016, Experiment 1). If a single-system (S-S or S-R) governed behaviour in a given animal then the patterns of behaviour should be consistent across different manipulanda (i.e., the left and right levers). A simple alternative to the analysis described in the previous paragraph assumes that behaviours directed toward the food-well and lever are generated by independent systems (S-S and S-R, respectively) that operate in parallel. This analysis predicts that a given form of response will exhibit the same characteristics independently of whether the animal in which it is observed is classified as a ST or a GT; with food-well activity being derived from the operation of a S-S system and lever-oriented behaviour being derived from a S-R system that operate to different degrees in all rodents. The dominant response might be toward the food-well in one rodent and lever in another, but in both rats food-well activity should more rapidly track changes in reinforcement contingencies than should lever-oriented activity. As already noted, this prediction was not assessed by Ahrens et al. (2016), but Patitucci et al. (2016, Experiment2) reported that satiation had a marked effect on food-well activity when the effect of this manipulation was considered across rats that had been classified as GTs or STs. This observation is consistent with the idea that activity directed toward the food-well and lever have the same properties irrespective of whether they were exhibited in STs or GTs.

To summarize, to the best of my knowledge nobody has directly investigated the following simple question: Does a given type of behaviour (e.g., lever oriented) have the same or different characteristics when assessed in STs and GTs? Here, I addressed this question in two experiments. In both experiments, rats received training procedures that should allow the two phenotypes to develop, and then the contingencies were changed (e.g., the reinforced lever became non-reinforced and vice versa, or reinforcer is changed from a more desirable one to a less desirable one and vice versa). The changes in behaviours directed toward the lever and food-well were then assessed as a function of whether the rodents had been classified as STs or GTs at the end of the first stage of training. Evidence favouring the claim that the S-S system generates food-well activity and the S-R system generates lever-oriented behaviour would take the form of a compelling dissociation: More rapid changes in food-well activity than in lever-oriented behaviour in both ST & GT groups, despite the quite different levels of performance anticipated in rats classified as predominantly goal-tracking or sign-tracking.

2.2 Experiment 1

2.2.1 Method

Animals and apparatus

Sixteen female Sprague Dawley rats were used (supplied by Charles River, UK). They had been subjects in a behavioural task involving drinking different concentrations of sucrose, but were naïve with respect to the apparatus and procedures used in Experiment 1[1]. Their mean ad libitum weight before the start of the experiment was 321g (range: 280-366g) and they were maintained at between 85 and 95 % of these weights by giving them restricted access to food at the end of each day. The animals were housed in groups ranging from two to four in standard cages with environmental enrichment (small wood, paper and cardboard tubes) and maintained on 12-hr/12-hr light/dark cycle (lights on at 7 a.m.). The research was conducted in accordance with Home Office regulations under the Animal (Scientific Procedures) Act 1986 (PPL 303243, PI Dominic M. Dwyer). The apparatus consisted of eight identical conditioning boxes measuring 30×24×21 cm (H×W×D; Med Associates, Georgia, VT). Each box was placed in a sound-attenuating shell that incorporated a ventilation fan, which maintained the background noise at 68 dB(A). The boxes had aluminum side walls and clear acrylic front back and top. The floor was constructed from 19 steel rods (4.8 mm diameter, 16 mm apart) and was situated above a stainless steel tray. Food pellets (45 mg: supplied by MLab: Richmond, IN) were delivered to a floor-level recessed food-well (aperture: 5.3×5.3 cm) in the centre of the left wall. The food-well was equipped with infrared detectors that allowed the presence of the rat in the well to be automatically recorded. A single response was registered when the detector was interrupted (e.g., when a rat's snout entered the food-well). Two retractable levers (4.5×1.8×0.2 cm) were located 3 cm to the left and right

---

[1] Experiment 1 was replicated as part of a larger study in which the rats were from the same source as Experiment 1, but were male rather than female. The results from this replication matched those of Experiment 1. Experiment 2 also used male rats to confirm the generality of the results from Experiment 1.

of the food-well and at a height of 4.6 cm and 1.5 cm from the edge of the wall. A lever press was recorded each occasion that the lever was depressed by 4mm from its usual horizontal resting position. MED-PC software was used to insert levers, deliver food pellets, and to record food-well entries and lever presses.

Procedure

The rats had two 24-min pre-training sessions where food pellets were delivered on a variable-time (VT) 60-s schedule (range: 40–80 s). Rats then received a single session of training on each of the next 12 days of training, which occurred at the same time of day for a given rat (see design in Table 1). These sessions consisted of 20 trials on which the left lever was inserted for 10s and then retracted and 20 trials on which the right lever was inserted for 10s and was then withdrawn. For half of the rats, the reinforced lever (L_1) was the left lever and the non-reinforced lever (L_2) was the right lever; and for the other half L_1 was the right lever and L_2 the left lever. The order in which the left and right levers were presented was random with the constraint that there could not be more than three same type trials in succession. Rats were assigned to the groups randomly. The trials were delivered on a variable-time (VT) 60-s schedule (range: 40–80 s). All rats then received reversal training for 12 days in which L_1 (e.g., left lever) was non-reinforced and L_2 (e.g., right lever) was reinforced. L_1 denotes the lever that was reinforced during training and non-reinforced during reversal and L_2 the lever that was non-reinforced during training and reinforced during reversal. The procedure used for the reversal stage was in other respects identical to the training stage.

Data Analysis

For manipulating the data, data cleaning, statistical analysis and data visualization, I used open source software R (RStudio, 2015) and JASP (JASP Team, 2018). For data cleaning and manipulation, I used “tidyr” (Wickham & Henry, 2018) and “dplyr”

Table 1: Design of the Experiment 1

Classification	Training	Reversal	Notation
ST	L_1 +	L_1 -	L_1 (+ ‖ -)
	L_2 -	L_2 +	L_2 (- ‖ +)
GT	L_1 +	L_1 -	L_1 (+ ‖ -)
	L_2 -	L_2 +	L_2 (- ‖ +)

Note: ST refers to a sign-tracker rat and GT refers to a goal-tracker rat. L_1 and L_2 refer to two levers (left and right, counterbalanced). During training, L_1 was reinforced ("+"; food pellet) and L_2 was non-reinforced ("-"). During the reversal stage L_1 was not reinforced and L_2 was reinforced (food pellet). Rats were classified as ST or GT on the basis of their bias towards lever pressing or entering the food-well during the final block of training on the reinforced lever.

R packages (Wickham, Henry, & Muller, 2019). For standard hypothesis testing, I used "ez" R package (Lawrence, 2016) and for Bayesian statistics I used JASP (JASP Team, 2018). For data visualization and representation, I used "ggplot2" (Wickham et al., 2018) R package, and "gridExtra" R packages (Auguie & Anotonov, 2017). The analysis are automatically reported in text in APA format style using "apa" R package (Gromer, 2019). For analysis of variance (ANOVA) when the sphericity assumption was violated, the analysis was reported with Greenhouse-Geisser correction. For post-hoc unplanned comparisons Bonferroni correction was applied (the threshold for declaring a p-value significant was equal to 0.05 divided by the number of comparisons).

Standard hypothesis testing does not directly assess whether the absence of a significant effect is sufficient evidence to conclude that there is no effect. In contrast, Bayesian statistics provides a ratio of the probability for the observed data under different models, such as a model based on the null hypothesis relative to a model based on some specified alternative. The resulting Bayes factors can then be interpreted according to the convention suggested by Rouder, Speckman, Sun, Morey, and Iverson (2009), where a Bayes factor between 1 and 3 provides anecdotal support, a factor between 3 and 10 suggests some supporting evidence, while a factor beyond 10 indicates strong evidence. I have, therefore, supplemented standard null-hypothesis statistical testing with the presentation of equivalent Bayes factors, when null results are of theoretical significance. Bayesian analysis was conducted with Bayes factors for main effects and interactions for factorial ANOVA in the way described by Rouder, Morey, Speckman, and Province (2012) and Rouder, Morey, Verhagen, Swagman, and Wagenmakers (2017).

Successive sessions during the training and reversal stages were combined into 12×2-day blocks (6 × training: T1-T6; and 6 × reversal: R1-R6). At the end of the training phase, the rats were split into two groups, sign-trackers (ST) and goal-trackers (GT), based on their tendency to engage with the lever and the food-well. A bias score was calculated using the number of lever presses and food-well entries for the reinforced lever, L_1: (Goal-tracking - Sign-tracking) / (Goal-tracking + Sign-tracking). Individual scores for each rat was calculated based on the raw lever presses and food-well entries averaged across last two days of training (T6). At this stage both behaviours are remarkably stable and have reached asymptote. A median split was used to divide rats into those with higher scores (group GT for goal-trackers) and those with lower scores (group ST for sign-trackers). A score close to 1 indicates a tendency to engage solely with the food-well and a score close to -1 a tendency to engage solely with the lever. With a fairly even distribution the median split criterion is around 0 (scores above 0 are classified as GT and scores below 0 are classified as ST). The distribution of behaviour is shown in Figure 1. Bias scores above -.16 were classified as GTs and below -.09 as STs.

Subsequent analyses were conducted separately for lever presses and food-well entries, with the main focus being on the transition between the final block of training (T6) and the first block of reversal (R1). I also conducted complementary analyses in which the number of lever presses and food-well entries were treated in a continuous fashion. Mixed ANOVAs were conducted separately for lever pressing and food-well entries with within-subjects factors block (training 6 levels: T1-T6; transition blocks 2 levels: T6-R1; reversal 6 levels: R1-R6;) and lever (2 levels: reinforced "+" vs non-reinforced "-") and between-subjects factor classification (2 levels: ST vs GT).

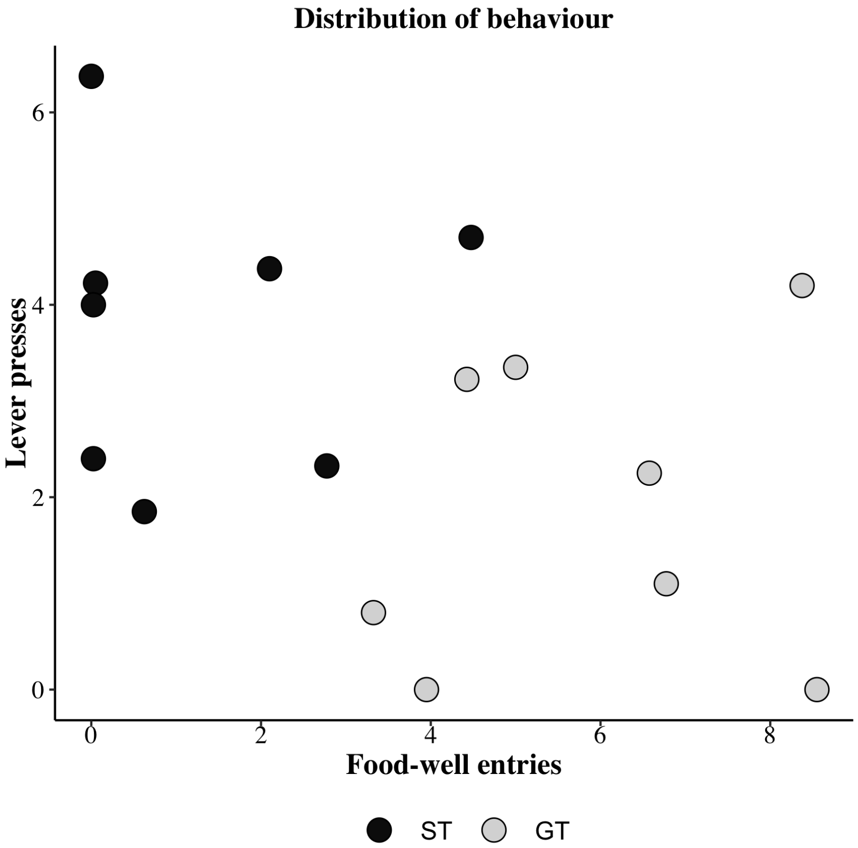


Figure 1: Distribution of sign-tracking (lever presses) and goal-tracking (food-well entries) behaviour for Experiment 1 per (10-s) trial during last block of training (T6) for L_1 (+ || -). The black symbols correspond to sign-trackers (ST) and the clear symbols to goal-trackers (GT).

2.2.2 Results

The main results from Experiment 1 are shown in Figure 2 and in Figure 3. The analysis will begin with results from the training stage (left-hand panels of Figure 2), before moving to the critical transition between training and reversal (identified by the grey section), and finally the reversal stage as a whole (right-hand panels).

Training

Inspection of the results from the first stage of training (left-hand side of the upper and lower panels of Figure 2) suggests that as training progressed rats in both groups (ST and GT) showed more lever presses and food-well entries during the reinforced

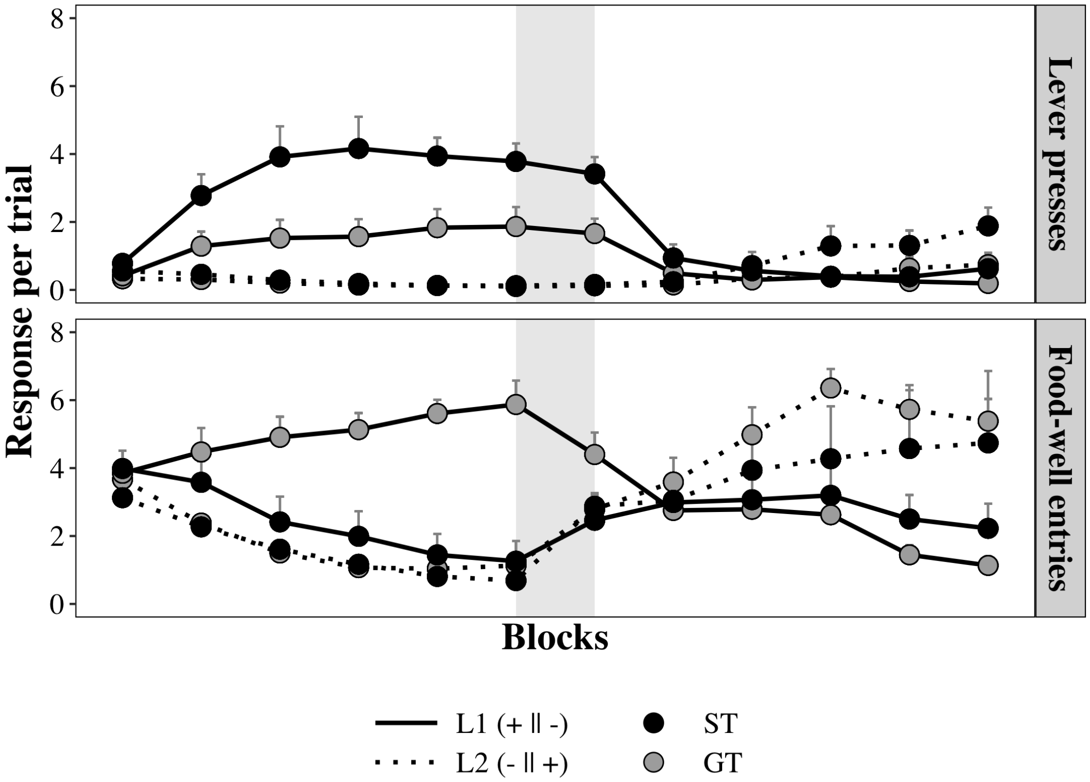


Figure 2: Results from Experiment 1: the effects of a reversal procedure on sign-tracking and goal-tracking. Mean (+SEM) lever presses (upper panel) and food-well entries (lower panel) per (10-s) trial across the two stages: training (T1–T6) and reversal (R1–R6). During training, rats received presentations of one lever paired with food pellets (L_1 (+ || -)) and nonreinforced presentations of a second lever (L_2 (- || +)); rats were classified as sign-trackers (STs) and goal-trackers (GTs) on the basis of their behaviour during the final block of training (T6). They then received a reversal: L_1 non-reinforced and L_2 reinforced. The grey section indicates transition between initial training and the reversal of the contingencies.

L_1 than the nonreinforced L_2. The fact that during the initial training sessions there was a higher level of food-well entries than lever presses probably reflects the impact of the pretraining sessions in which food pellets were delivered into the food-well. In any event, the discrimination involving lever presses was more evident in Group ST than Group GT, while the discrimination involving food-well entries was more evident in Group GT than Group ST, with these between-groups differences being most apparent on reinforced L1 trials. The description of the training results is supported by separate analyses of lever presses and food-well entries.

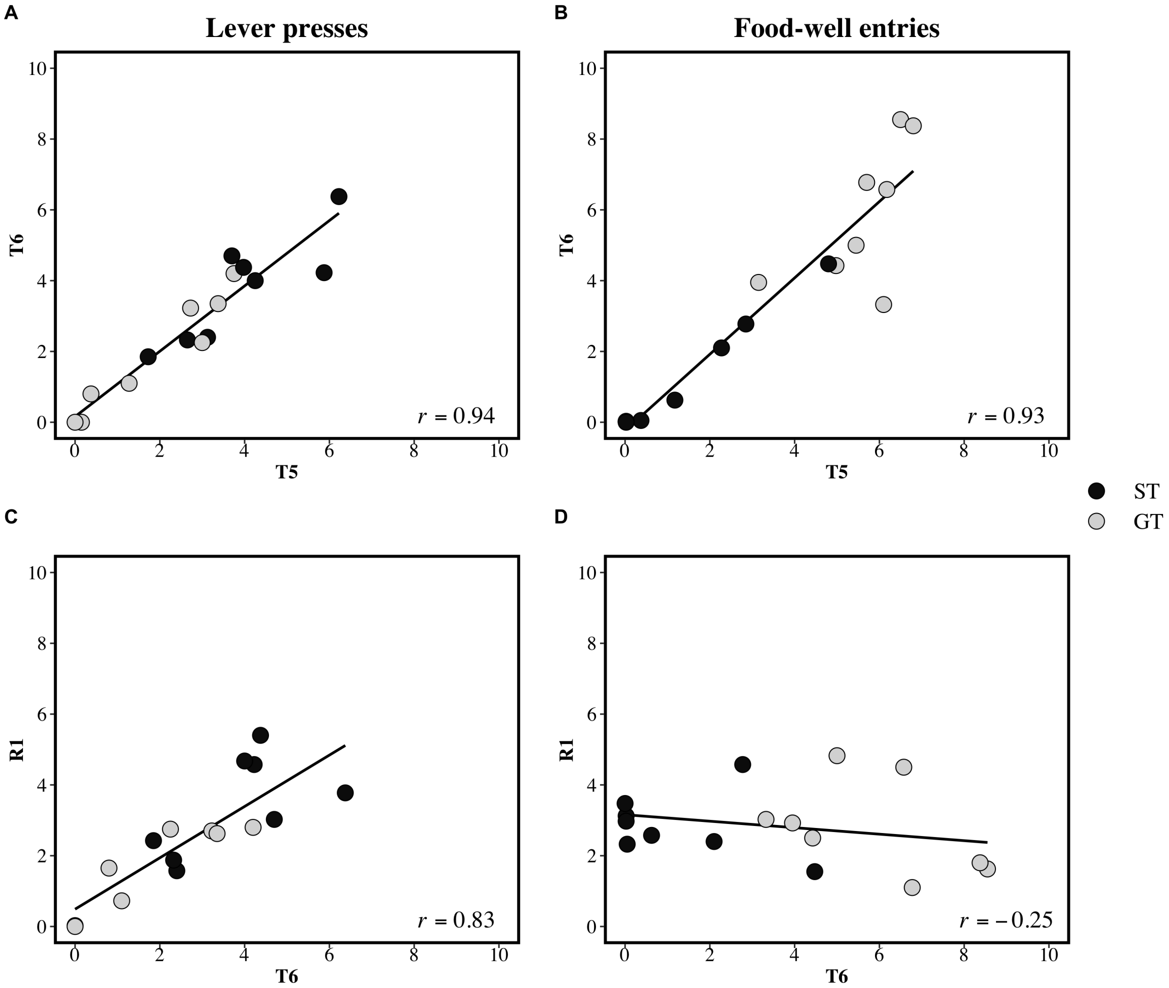


Figure 3: Correlation results from Experiment 1: the effects of a reversal procedure on sign-tracking and goal-tracking. The upper panels show the relationship between the mean number of responses per (10-s) trial between blocks T5 and T6 for lever presses (A) and for food-well entries (B) on reinforced trials. The lower panels show the relationship T6 and R1 for lever presses (C) and food-well entries (D) on reinforced trials. The black symbols correspond to sign-trackers (ST) and the clear symbols to goal-trackers (GT).

Lever presses

An ANOVA conducted for lever presses revealed main effects of classification, $F(1,\ 14) = 6.36$, $p = .024$, $\eta_p^2 = .31$, block, $F(2.96,\ 41.41) = 9.74$, $p < .001$, $\eta_p^2 = .41$ and lever, $F(1,\ 14) = 40.06$, $p < .001$, $\eta_p^2 = .74$, as well as interactions between classification and block, $F(2.96,\ 41.41) = 2.15$, $p = .109$, $\eta_p^2 = .13$, classification and lever $F(1,\ 14) = 6.88$, $p = .020$, $\eta_p^2 = .33$, and block and lever $F(2.60,\ 36.35) = 18.75$, $p < .001$, $\eta_p^2 = .57$. There was a triple interaction between classification, block and lever, $F(2.60,\ 36.35) = 3.42$, $p = .033$, $\eta_p^2 = .20$.

Food-well entries

A parallel analysis was conducted for food-well-entries and revealed main effects of classification $F(1,\ 14) = 12.19$, $p = .004$, $\eta_p^2 = .47$, and lever, $F(1,\ 14) = 37.82$, $p < .001$, $\eta_p^2 = .73$, but not a block effect, $F(2.68,\ 37.49) = 11.98$, $p < .001$, $\eta_p^2 = .46$. There were interactions between classification and lever, $F(1,\ 14) = 12.90$, $p = .003$, $\eta_p^2 = .48$, classification and block, $F(2.68,\ 37.49) = 7.30$, $p < .001$, $\eta_p^2 = .34$ and between block and lever $F(2.83,\ 39.69) = 7.40$, $p < .001$, $\eta_p^2 = .35$. There was a triple interaction between classification, block and lever, $F(2.83,\ 39.69) = 10.33$, $p < .001$, $\eta_p^2 = .42$.

Transition blocks

The results from the transition between the final block of training and first block of reversal (gray section of Figure 2) are of central interest. Inspection of this transition highlights the fact that lever presses remained stable in spite of the reversed reinforcement contingencies (upper panel), whereas food-well entries changed rapidly (lower panel). Moreover, these differences between the effects of the reversal on lever presses and food-well entries were evident in both Groups ST and GT: The levels of lever pressing remained largely unchanged in both groups; and while there was a marked decrease in food-well entries to the previously reinforced lever in Group GT there were marked increases in food-well entries to the previously nonreinforced lever in both Group GT and ST. Also, in Group ST there was a more marked increase in responding between the T6 and R1 for L_2 than L_1.

Lever presses

An ANOVA conducted for lever presses on the critical reversal blocks revealed main effects of classification, $F(1,\ 14) = 7.71$, $p = .015$, $\eta_p^2 = .36$ and lever, $F(1,\ 14) = 55.27$, $p < .001$, $\eta_p^2 = .80$, but no effect of block, $F(1,\ 14) = 0.82$, $p = .380$, $\eta_p^2 = .06$. There was an interaction between classification and lever, $F(1,\ 14) = 6.90$, $p = .020$, $\eta_p^2 = .33$, but no interaction between classification and block $F(1,\ 14) = 0.07$, $p = .800$, $\eta_p^2 < .01$, and critically no interaction between block and lever, $F(1,\ 14) = 1.68$, $p = .216$, $\eta_p^2 = .11$.

There was no triple interaction between classification, block and lever $F(1,\, 14) = 0.13$, $p = .728$, $\eta_p^2 < .01$. The Bayes factor for the best model without the interaction between block and lever relative to the model with the interaction was 9.52, which indicates evidence against the presence of the interaction. The Bayes factor for the best model without the interaction between block, lever and group relative to the model with the interaction was 62.50, which represents strong evidence against the presence of the interaction. This analysis indicates that there were no immediate effects of the reversal on lever presses in either of groups ST and GT.

Food-well entries

A parallel analysis of food-well entries for the critical transition blocks revealed an effect of classification, $F(1,\, 14) = 18.90$, $p < .001$, $\eta_p^2 = .57$, block $F(1,\, 14) = 12.04$, $p = .004$, $\eta_p^2 = .46$ and lever, $F(1, 14) = 11.88$, $p = .004$, $\eta_p^2 = .46$. There were interactions between classification and block, $F(1,\, 14) = 9.68$, $p = .008$, $\eta_p^2 = .41$, classification and lever, $F(1,\, 14) = 10.75$, $p = .005$, $\eta_p^2 = .43$ and more importantly an interaction between block and lever, $F(1,\, 14) = 43.68$, $p < .001$, $\eta_p^2 = .76$. There was a triple interaction between classification, block and lever, $F(1,\, 14) = 11.87$, $p = .004$, $\eta_p^2 = .46$.

The analysis just presented involved dividing rats into two groups (ST and GT) using their biases during the final block of training. However, the same conclusions are supported by an analysis in which their lever presses and food-well entries are treated as a continuum. The upper panels of Figure 3 depict the relationship between lever presses (panel A) on reinforced L_1 trials for the final blocks of training (i.e., T5 and T6; left-hand panel), and between food-well entries on L_1 trials for the same blocks (panel B). The lower panels depict the relationships between lever presses on the final block of training and the first block of reversal (i.e., T6 and R1, panel C) on L_1 trials, and between food-well entries for the same two blocks (panel D) on L_1 trials. The group membership of each rat is identified. For both types of response, there was a significant correlation between T5 and T6, lever presses, ($r(14) = .94$, $p < .001$; food-well entries, $r(14) = .93$, $p < .001$). However, while there was a correlation between T6 and R1 for lever presses ($r(14) = .83$, $p < .001$, there was not for food-well entries ($r(14) =$ -.25, $p = .355$). Food-well entries changed between T6 and R1, but lever presses did not.

Reversal

Across the blocks of reversal training (right-hand side of the upper and lower panels of Figure 2), the numbers of lever presses increased during L_2 and decreased during L_1 in Group ST, and this change was numerically smaller in Group GT. In contrast, the number of food-well entries increased during L_2 and decreased in L_1 in Group GT, and this change was less apparent in Group ST.

Lever presses

An ANOVA for lever presses for the blocks following the reversal revealed main effects of block, $F(2.95,\ 41.24) = 17.07$, $p < .001$, $\eta_p^2 = .55$, but no effects of classification, $F(1,\ 14) = 3.60$, $p = .079$, $\eta_p^2 = .20$ or lever, $F(1,\ 14) = 1.03$, $p = .327$, $\eta_p^2 = .07$. There was a block by lever interaction, $F(1.68,\ 23.58) = 21.96$, $p < .001$, $\eta_p^2 = .61$, but no interaction between classification and block, $F(2.95,\ 41.24) = 2.56$, $p = .069$, $\eta_p^2 = .15$, or classification and lever, $F(1,\ 14) = 0.01$, $p = .919$, $\eta_p^2 < .01$. There was no significant interaction between classification, block and lever, $F(1.68,\ 23.58) = 3.48$, $p = .054$, $\eta_p^2 = .20$.

Food-well entries

A parallel analysis for food-well entries revealed a main effect of lever, $F(1,\ 14) = 22.76$, $p < .001$, $\eta_p^2 = .62$, but no effect of classification, $F(1,\ 14) = 0.13$, $p = .721$, $\eta_p^2 < .01$, or block, $F(1.70,\ 23.75) = 1.42$, $p = .259$, $\eta_p^2 = .09$. There was a block by lever interaction $F(1.85,\ 25.90) = 13.04$, $p < .001$, $\eta_p^2 = .48$, but no interaction between classification and block, $F(1.70,\ 23.75) = 0.46$, $p = .604$, $\eta_p^2 = .03$, or classification and lever, $F(1,\ 14) = 2.36$, $p = .147$, $\eta_p^2 = .14$. There was a triple interaction between classification, block and lever, $F(1.85,\ 25.90) = 3.67$, $p = .042$, $\eta_p^2 = .21$.

2.2.3 Discussion Experiment 1

Discrimination training where the presentation of one lever (L_1) was paired with food pellets and another (L_2) was not, resulted in marked individual differences in conditioned responding; with some rats interacting with L1 (but not L2) and others approaching the site of food delivery during L_1 (but not L_2). When the contingencies were reversed, with L_1 now non-reinforced and L_2 reinforced, the different levels of lever pressing to L_1 (and L_2) in Groups ST and GT remained remarkably stable during the first block of reversal. In contrast, the levels of food-well entries changed more rapidly in both Groups ST and GT (Figure 2). This differential sensitivity of the two response forms to changing contingencies was also evident when they were considered as continuous variables (Figure 3). These results demonstrate that the dissociation between lever presses (in rats designated as STs) and food-well entries (in rats designated as GTs) does not reflect a difference in the sensitivity of the two groups to changed reinforcement contingencies per se (cf. Ahrens et al., 2016). Instead, these results show that the lever press and food-well entry responses are differently sensitive to such changes irrespective of the phenotype of the rat. These observations suggest that the distinct behaviours reflect the parallel operation of S-S and S-R systems within an individual, rather than the operation of a single system (either S-S or S-R) that gives rise to both behaviours (cf. Lesaint et al., 2014). Experiment 2 attempted to extend these observations by examining whether changes in the nature of the reinforcer (between alternatives that produce different levels of responding) produce more rapid changes in food-well activity than in lever pressing in STs and GTs.

2.3 Experiment 2

In Experiment 1, I examined how the extinction of an association between a lever and an appetitive US affected behaviour on lever presses and food-well entries in both STs and GTs. In Experiment 2, rats received separate presentations of two levers (L_1

and L_2) that were both paired with the same reinforcer during training (either food pellets or sucrose). Pilot research had established that food pellets maintain higher levels of both lever pressing and food-well entries than does sucrose (see also Patitucci et al., 2016) which should be evident in the first stage of training in Experiment 2. The rats that were reinforced with pellets and sucrose were further divided into two groups (Group ST ad GT) on the basis of their biases at the end of training. During the second stage, the reinforcers associated with the two levers were switched: the rats given pellets during training received sucrose during the switch and those given sucrose during training received food pellets during the switch (see design in Table 2). The issue of central interest was the extent to which the two target behaviours that had developed during L_1 and L_2 (lever presses and food-well entries) would change to reflect the fact that the levers were now paired with reinforcers that maintained different levels of performance (i.e., food pellets and sucrose). If the behaviour of rats in Group ST is generated by a S-R system, then both lever presses and food-well entries should be less sensitive to the change in reinforcer type than those in Group GT, whose behaviour is generated by a S-S system. However, if lever pressing is based on a S-R system whereas food-well entries reflect a S-S system, then lever pressing should be less sensitive to the change in contingencies between the training and switch stages than should food-well entries, irrespective of whether those behaviours are expressed in Group ST or Group GT.

2.3.1 Method

Animals and apparatus

Thirty-two naïve male Lister Hooded rats (supplied by Envigo, Bicester, U.K.) were housed in the same way as described in Experiment 1. Their mean ad libitum weight was 295 g (range 284–320 g). Rats had free access to water, and they were maintained between 85 and 95 % of their ad lib weights by giving them restricted access to food at the end of each day. The experimental chambers were those used in Experiment

1. The design of the experiment is similar to Experiment 1, with the addition that in Experiment 2, the sucrose dipper delivered 0.05 ml of sucrose solution (8% weight/weight with water) for half the rats (see Table 2). The sucrose dipper was situated at the right-hand of the food-well in close proximity of the pellet dispenser. The default position of the dipper is levelled up and accessible to rats and at point of reward delivery the dipper lowers and gets refilled. That is, sucrose remains available indefinitely, like pellets, but only the amount delivered by one lowering of the dipper. Responding was recorded in the same way as described for food pellets in Experiment 1.

Procedure

The rats had two 24-min pretraining sessions before the training and switch stages. During these sessions, the rats received the reinforcer (food pellets or sucrose) that was to be delivered in the immediately succeeding stage. The reinforcers were delivered on a variable 60-s schedule (range 40 – 80 s). Rats received 12 days of training that were arranged in the same way as Experiment 1 with the exception that the presentation of both levers (L_1 and L_2 - left and right levers, counterbalanced) were followed by a reinforcer (food pellets for half of the rats and sucrose for the remainder). The switch stage also consisted of 12 days. This stage was identical to the training stage with the exception that the rats that had received food pellets during the training stage received sucrose during the switch stage, and those that had received sucrose during training received food pellets during the switch. Rats were randomly assigned to the experimental groups.

Data Analysis

There were strong positive correlations between lever press behaviour during the presentations of L_1 and L_2 that were both paired with the same outcome (either food pellets or sucrose) and between food-well behaviour on the two levers. These observations have some theoretical significance when contrasted with the results of Patitucci et al. (2016),

Table 2: Design of Experiment 2

Classification	Training	Switch	Notation
ST	L_1 & L_2 Pel or L_1 & L_2 Suc	L_1 & L_2 Suc or L_1 & L_2 Pel	ST Pel ‖ Suc ST Suc ‖ Pel
GT	L_1 & L_2 Pel or L_1 & L_2 Suc	L_1 & L_2 Suc or L_1 & L_2 Pel	GT Pel ‖ Suc GT Suc ‖ Pel

Note: ST refers to a sign-tracker rat and GT refers to a goal-tracker rat. L_1 and L_2 refer to two levers (left and right, counterbalanced). During training, both levers were paired with one reinforcer (food pellets or sucrose), and during the switch, both levers were then paired with the other reinforcer (sucrose or food pellets, respectively). Rats were classified as ST or GT on the basis of their bias towards lever pressing or entering the food-well during the final block of training.

who reported no correlation between the sign- and goal-tracking biases on two levers that signalled different outcomes. I shall consider the implications of this evidence in the Chapter 2 Discussion. However, to simplify the results section, the principal analysis of the results of Experiment 2 will be conducted with the frequency of responses combined across the left and right levers. As in Experiment 1, the training and switch sessions were combined into 2-day blocks for the purpose of analysis. The rats were split into two groups, STs and GTs, using the bias score described in Experiment 1. The split was conducted separately for the subgroups of rats that received food pellets and sucrose during the training stage. This resulted in four groups (8 in each group): ST Pel ‖ Suc (pellets during training, sucrose during the switch stage), GT Pel ‖ Suc (GTs; pellets during training, sucrose during the switch), ST Suc ‖ Pel (STs; sucrose during training, pellets during the switch) and GT Suc ‖ Pel (GTs; sucrose during training, pellets during the switch). For the Pel ‖ Suc group bias scores above .33 were classified as GTs and scores below .31 as STs. For the Suc ‖ Pel group scores above .58 were classified as GTs and below .54 as STs. The distribution of sign-tracking and goal-tracking behaviour is shown in Figure 4.

Mixed ANOVAs were conducted separately for lever pressing and food-well entries with within-subjects factor block (for the training and the switch stage 6 levels: T1-T6 and S1-S6; transition stage 2 levels: T6-S1), and between-subjects factors reinforcer (2 levels: pel vs suc) and classification (2 levels: ST vs GT).

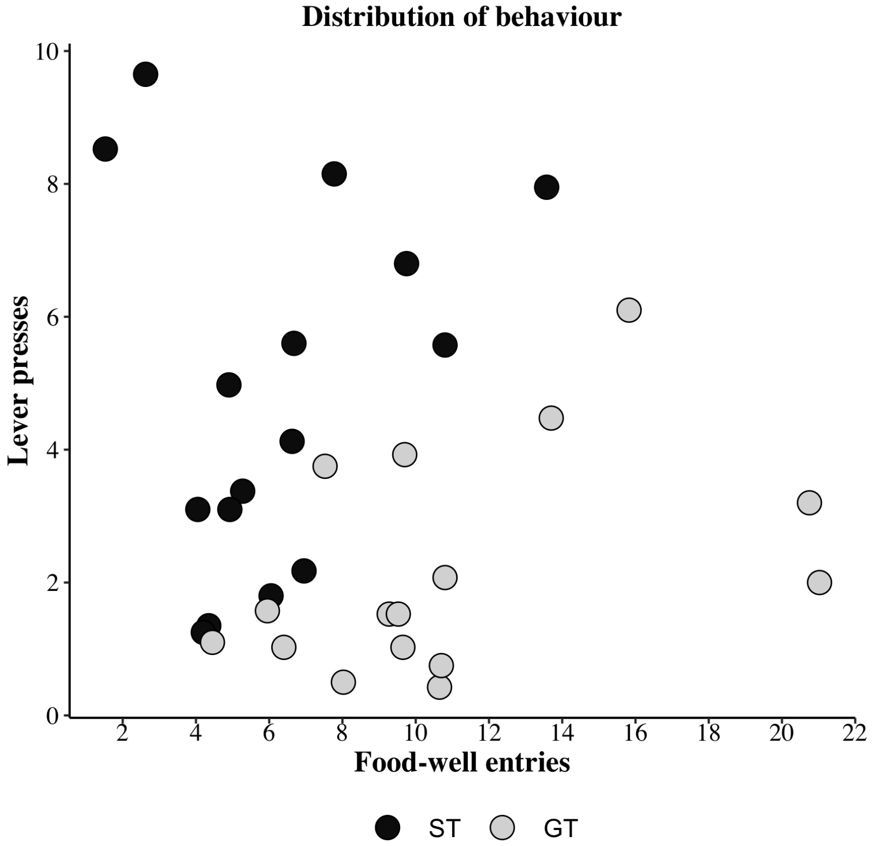


Figure 4: Distribution of sign-tracking (lever presses) and goal-tracking (food-well entries) behaviour for Experiment 2 per (10-s) trial during last block of training (T6) averaged across L_1 and L_2. The black symbols correspond to sign-trackers (ST) and the clear symbols to goal-trackers (GT).

2.3.2 Results

The main results from Experiment 2 are shown in Figure 5 and in Figure 6. As in Experiment 1, the analysis of the results of Experiment 2 will begin with results from the training stage, before moving to a comparison of the final block of training with the first block of reversal (identified by the grey section), and finally the switch stage as a whole.

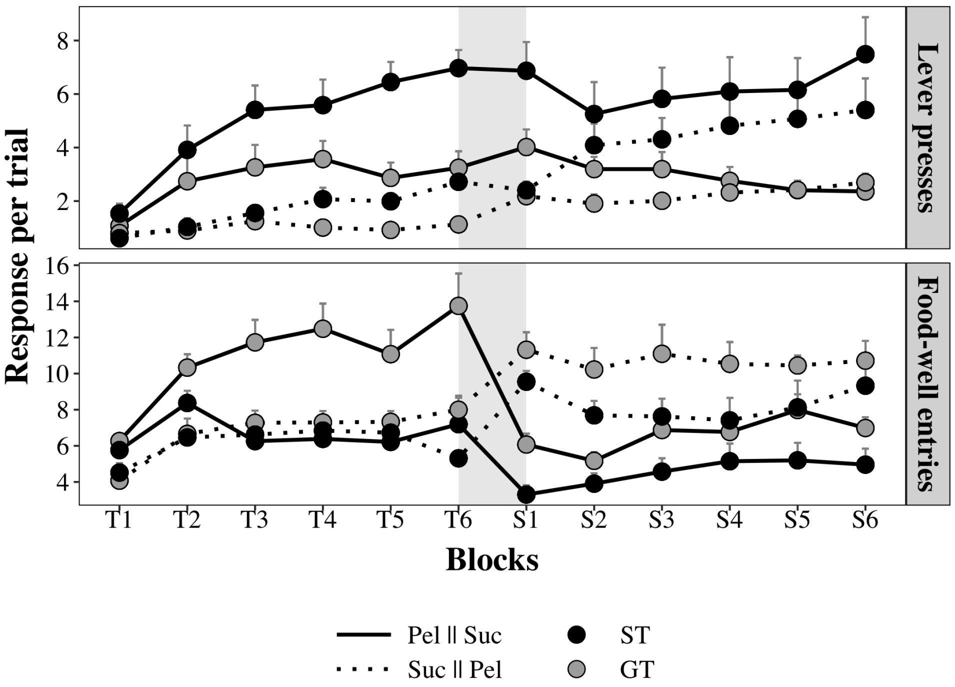


Figure 5: Results from Experiment 2: the effects of a switch in US value on sign-tracking and goal-tracking. Mean (SEM) lever presses (upper panel) and food-well entries (lower panel) per (10-s) trial across the two stages: training (T1–T6) and switch (S1–S6). During training, rats received presentations of two levers paired with either pellets or sucrose. Rats were classified as sign-trackers (ST) and goal-trackers (GT) on the basis of their behaviour during the final block of training (T6). The reinforcers that followed the levers were swapped during the second switch stage. The grey section indicates transition between initial training and the swap from sucrose to food pellet rewards (or from pellets to sucrose). The black symbols correspond to sign-trackers (ST) and the clear symbols to goal-trackers (GT). Solid lines indicate groups receiving pellets in phase 1 and sucrose after the switch to phase 2 (Pel || Suc) and dashed lines those switched from sucrose to pellets (Suc || Pel). For example, the means for the ST group that received food pellets during training and sucrose during the switch are given by black symbols (ST) and continous line (Pel || Suc).

Training

Inspection of the left-hand side of the upper and lower panels in Figure 5 suggests that the ST groups are more likely to engage in lever pressing than are the GT groups, and that the GT groups are more likely to enter the food-well than the ST groups. These group differences, especially in the case of food-well activity, were most marked when food pellets were the reinforcer. This description of the training results presented

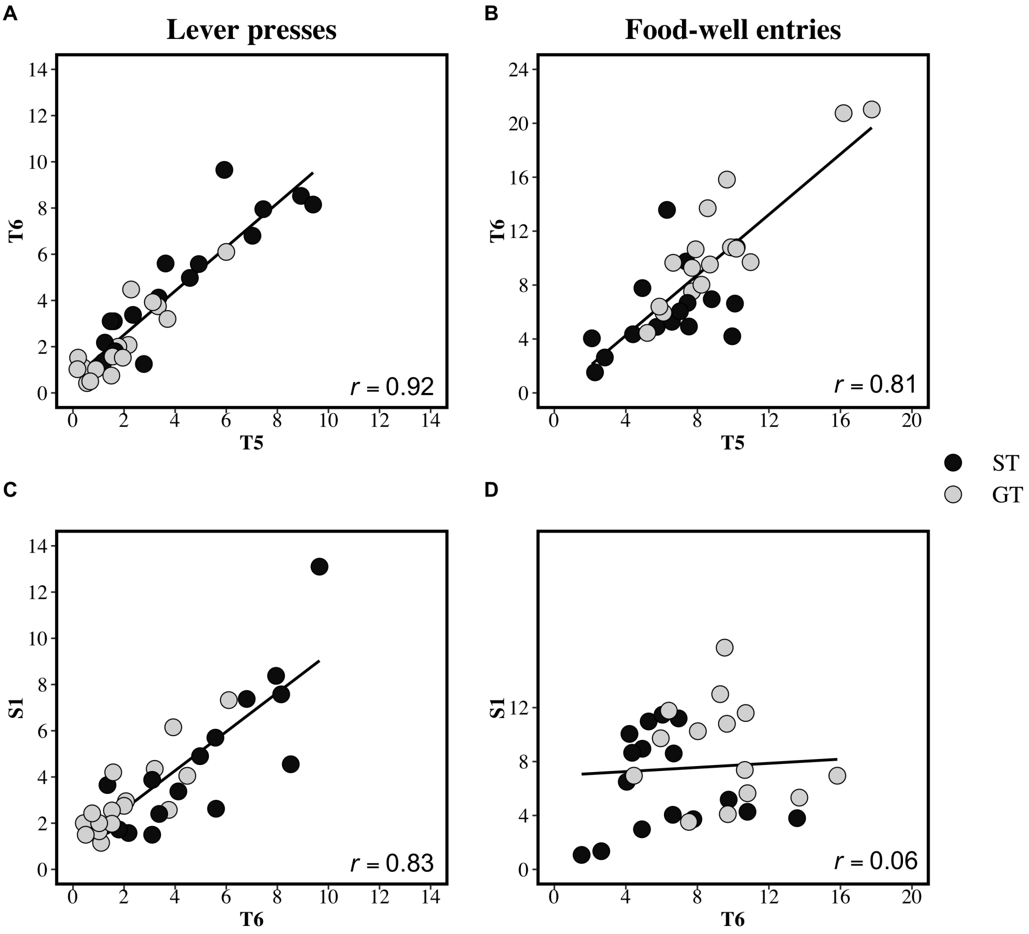


Figure 6: Correlation results from Experiment 2: the effects of a switch in US value on sign-tracking and goal-tracking. The upper panels show the relationship between the mean number of responses per (10-s) trial between blocks T5 and T6 for lever presses (A) and for food-well entries (B) on reinforced trials. The lower panels show the relationship T6 and R1 for lever presses (C) and food-well entries (D) on reinforced trials. The response (lever pressing and food-well entries) has been pooled across L_1 and L_2 trials. The black symbols correspond to sign-trackers (ST) and the clear symbols to goal-trackers (GT).

in Figure 5 is supported by separate analyses of lever presses and food-well entries.

Lever presses

ANOVA conducted on lever presses, pooled across the two levers, confirmed there were main effects of reinforcer, $F(1,\ 28) = 31.02$, $p < .001$, $\eta_p^2 = .53$, classification (ST vs GT), $F(1,\ 28) = 9.66$, $p = .004$, $\eta_p^2 = .26$, and block (T1-T6), $F(3.15,\ 88.33) = 24.96$, $p < .001$, $\eta_p^2 = .47$. There were interactions between reinforcer and block, $F(3.15,\ 88.33) = 7.76$, $p < .001$, $\eta_p^2 = .22$, and between classification and block, $F(3.15,\ 88.33) = 7.11$, $p < .001$, $\eta_p^2 = .20$, but not between reinforcer and classification, $F(1,\ 28) = 2.76$, $p = .108$, $\eta_p^2 = .09$. The three-way interaction between reinforcer, classification and block was not significant, $F(3.15,\ 88.33) = 1.06$, $p = .373$, $\eta_p^2 = .04$.

Food-well entries

A parallel analysis of food-well entries revealed main effects of reinforcer, $F(1,\ 28) = 9.86$, $p = .004$, $\eta_p^2 = .26$, classification, $F(1,\ 28) = 10.47$, $p = .003$, $\eta_p^2 = .27$ and block, $F(2.81,\ 78.57) = 19.48$, $p < .001$, $\eta_p^2 = .41$. There were also interactions between reinforcer and classification, $F(1,\ 28) = 5.42$, $p = .027$, $\eta_p^2 = .16$, and between classification and block, $F(2.81,\ 78.57) = 8.65$, $p < .001$, $\eta_p^2 = .24$. There was no interaction between classification and block, $F(2.81,\ 78.57) = 2.14$, $p = .105$, $\eta_p^2 = .07$, and no triple interaction between reinforcer, classification and block, $F(2.81,\ 78.57) = 2.64$, $p = .059$, $\eta_p^2 = .09$.

Transition blocks

Inspection of the grey panel in Figure 5 shows that there were rapid changes in food-well entries (lower panel) but not in lever pressing (upper panel). To be more specific: The high level of food-well activity - previously maintained by pellets-declined from T6 to S1, and the low level of food-well activity - previously maintained by sucrose-increased from T6 to S1. In contrast, lever pressing was largely unchanged across T6 and S1. This description was supported by separate analysis of lever presses and food-well entries.

Lever presses

ANOVA conducted for lever presses revealed an effect of reinforcer, $F(1,\ 28) = 33.57$, $p < .001$, $\eta_p^2 = .55$ and classification, $F(1,\ 28) = 14.72$, $p < .001$, $\eta_p^2 = .34$, but no effect of block, $F(1,\ 28) = 1.86$, $p = .183$, $\eta_p^2 = .06$. There was an interaction between reinforcer and classification, $F(1,\ 28) = 4.74$, $p = .038$, $\eta_p^2 = .14$, and an interaction between classification and block, $F(1,\ 28) = 4.70$, $p = .039$, $\eta_p^2 = .14$. Critically, there was no interaction between reinforcer and block, $F(1,\ 28) = 0.00$, $p = .952$, $\eta_p^2 < .01$, and no three-way interaction between reinforcer, classification and block, $F(1,\ 28) = 0.23$, $p = .637$, $\eta_p^2 < .01$. The Bayes factor for the best model without the block by lever interaction relative to the best model with the interaction is 5.88, indicating evidence against the presence of the interaction. The Bayes factor for the best model without the block by lever by group interaction relative to the model with the interaction is 100, indicating strong evidence against the presence of the interaction. The switch in reinforcer had little impact on lever press behaviour in either the ST or GT groups.

Food-well entries

A parallel analysis of food-well entries revealed a main effect of classification, $F(1,\ 28) = 15.59$, $p < .001$, $\eta_p^2 = .36$ and block, $F(1,\ 28) = 4.18$, $p = .050$, $\eta_p^2 = .13$, but no effect of reinforcer, $F(1,\ 28) = 1.22$, $p = .278$, $\eta_p^2 = .04$. Critically, there was an interaction between reinforcer and block, $F(1,\ 28) = 95.42$, $p < .001$, $\eta_p^2 = .77$, as well as between clasification and block, $F(1,\ 28) = 5.76$, $p = .023$, $\eta_p^2 = .17$, but not between reinforcer and classification, $F(1,\ 28) = 1.94$, $p = .174$, $\eta_p^2 = .06$, and no triple interaction between reinforcer, classification and block, $F(1,\ 28) = 2.14$, $p = .155$, $\eta_p^2 = .07$. The Bayes factor for the best model without the interaction between block, lever and group relative to the model with the interaction is 25, indicating strong evidence against the interaction. That is the switch in reinforcers had an immediate impact on behaviour directed to the food-well, and critically this was equivalent in both the ST and GT groups.

The same conclusions are supported by an analysis in which lever presses and food-well entries were treated as a continuum. The upper panels of Figure 6 depict the relationship between lever presses on the final blocks of training (i.e., T5 and T6; left-hand panel) and between food-well entries on the same blocks (right-hand panel) pooled across left and right lever trials. The lower panels depict the relationships between lever presses on the final block of training and the first block of switch (i.e., T6 and S1 left-hand panel), and between food-well entries on the same two blocks (right-hand panel). For both types of response there was a significant correlation between T5 and T6 (lever,

$r(30) = .92$, $p < .001$; food-well entries, $r(30) = .81$, $p < .001$). Between T6 and S1 there is a significant correlation for lever presses ($r(30) = .83$, $p < .001$), but not for food-well entries ($r(30) = .06$, $p = .734$).

Switch

The pattern of results evident on the first block of the switch (i.e., S1) was, for the most part, evident across the later blocks of the switch stage. More specifically, the marked changes in food-well entries were sustained across the switch stage and were accompanied by little change in lever pressing: while the low level of lever pressing increased when sucrose was replaced with food pellets during the switch, the high level of lever pressing was maintained when food pellets were replaced with sucrose.

Lever presses

ANOVA conducted on lever presses, pooled across the two levers, confirmed there was a main effect of classification, $F(1, 28) = 13.69$, $p < .001$, $\eta_p^2 = .33$, however there were no effects of reinforcer, $F(1, 28) = 3.35$, $p = .078$, $\eta_p^2 = .11$ or block, $F(3.51, 98.17) = 2.20$, $p = .083$, $\eta_p^2 = .07$. There were interactions between reinforcer and block, $F(3.51, 98.17) = 5.63$, $p < .001$, $\eta_p^2 = .17$, and between classification and block, $F(3.51, 98.17) = 4.52$, $p = .003$, $\eta_p^2 = .14$, but not between reinforcer and classification, $F(1, 28) = 0.68$, $p = .417$, $\eta_p^2 = .02$. The three-way interaction between reinforcer, classification and block was not significant, $F(3.51, 98.17) = 1.98$, $p = .112$, $\eta_p^2 = .07$.

Food-well entries

A parallel analysis of food-well entries revealed main effects of reinforcer, $F(1, 28) = 20.46$, $p < .001$, $\eta_p^2 = .42$, classification, $F(1, 28) = 6.90$, $p = .014$, $\eta_p^2 = .20$, and block $F(2.87, 80.25) = 3.59$, $p = .019$, $\eta_p^2 = .11$. There was an interaction between reinforcer and block, $F(2.87, 80.25) = 5.35$, $p = .002$, $\eta_p^2 = .16$, but no interactions between reinforcer and classification, $F(1, 28) = 0.03$, $p = .862$, $\eta_p^2 < .01$, or between classification and block, $F(2.87, 80.25) = 0.82$, $p = .481$, $\eta_p^2 = .03$, and no triple interaction between reinforcer, classification and block, $F(2.87, 80.25) = 1.41$, $p = .247$, $\eta_p^2 = .05$.

2.3.3 Discussion Experiment 2

The results of Experiment 2 confirm the principal conclusions derived from the results of Experiment 1. First, lever-press behaviour was less sensitive to changes in reinforcement contingencies than was food-well behaviour. Second, this difference in sensitivity was equally apparent in rats that were classified as STs and GTs. In Experiment 1, these conclusions were supported by the effects of a reversal between the relationships between two levers (L_1 and L_2) and the presence and absence of food pellets, whereas in Experiment 2 they were supported by the substitution of reinforcers that maintained more (pellets) or less (sucrose) behaviour. The facts that Experiment 2 used male rats while Experiment 1 used female rats (and its results have been replicated in male rats) and the two experiments used different strains (Sprague-Dawley and Lister Hooded, respectively), suggests that the difference in sensitivity of lever and food-well directed behaviour to changes in reinforcement contingencies is preserved across rat strains and male/female animals.

2.4 Chapter 2 Discussion

During appetitive Pavlovian conditioning, rodents will reliably display behaviour directed both toward the stimulus (sign-tracking) and toward the site of food pellet delivery (goal-tracking). Although individual differences in conditioned responding have typically received scant consideration in theories of associative learning, it is clear that the distribution of these behaviours differs across individuals (e.g., Fitzpatrick et al., 2013). For example, when a lever is temporarily inserted into a conditioning chamber and paired with food pellets some rats develop a consistent tendency to interact with the lever whereas others develop a tendency to approach the food-well. These behaviours are differently sensitive to the current value of the reinforcer and indeed its presence.

Patitucci et al. (2016) demonstrated that the bias toward engaging in food-well activity rather than lever-press activity was positively correlated with the palatability of the reinforcer; and sating rats on the reinforcer reduced food-well but not lever-oriented activity; and Ahrens et al. (2016) showed that lever pressing, in rats that predominantly engaged in sign-tracking, was less sensitive to extinction than food-well activity, in rats that predominantly engaged in goal-tracking.

My results confirm that lever-press behaviour is indeed less sensitive to changes in reinforcement contingencies than is food-well behaviour. In Experiment 1, this was evident in the effects of a reversal in the relationships between two levers and the presence and absence of food pellets, whereas in Experiment 2 it was evident in the effects of the substitution of reinforcers that maintained more (pellets) or less (sucrose) behaviour. Moreover, in both experiments, these conclusions received additional support from treating lever-press and food-well activity in a continuous way: lever-press activity was correlated between the final block of training (T6) and the first block of the changed contingencies (R1 in Experiment 1 and S1 in Experiment 2), but food-well activity was not. Here, I contrasted two possible accounts of the behavioural phenotypes, based on the assumption that sign- and goal-tracking are modulated by two systems (e.g., S-S/S-R, or model-based/model-free). First, that the behaviour of a given rodent is governed by the operation of a single system and that the control of both types of behaviour simply reflects the nature of the governing system (i.e., in a goal-tracking rat the predominant system that drives behaviour is S-S or model-based, and in a sign-tracking rat is S-R or model-free). This account predicts that food-well and lever-oriented behaviours will exhibit different properties in STs and GTs. Second, that behaviour directed toward the food-well and lever are generated by independent systems (S-S/model-based and S-R/model-free, respectively) that operate in parallel. This analysis predicts that a given form of response will exhibit the same characteristics in STs and GTs. In Experiments 1 and 2, lever-press and food-well behaviour in both goal-tracking and sign-tracking rats showed the same pattern of sensitivity to changes in reinforcer contingencies. This pattern of results provides support for the second of these accounts: differential sensitivity to contingency changes is a property of the behaviour, not the animal.

While the results that I have presented so far have clear implications regarding the control of behaviours in the two behavioural phenotypes they do not contribute to our understanding of the origin of the two phenotypes. Patitucci et al. (2016) argued that food-well activity was more likely to dominate in (goal-tracking) rats that, for whatever reason, valued the reinforcer more, based on the fact that goal-tracking correlated with the palatability (preference) of the reinforcer. But, they also observed that the classification of a rat as either a GT (or ST) on a lever that was paired with one reinforcer (e.g., food pellets) was unrelated to the classification of the same rat on a second lever that was paired with a different reinforcer (e.g., sucrose). They argued that if a given rat valued one reinforcer (e.g., food pellets) more than the other (e.g., sucrose) then this would result in more goal-tracking on one lever than another. While this analysis is certainly consistent with other features of their results, a more prosaic account can be developed for the lack of correlations between the behaviours directed to two levers: It might have reflected superstitious reinforcement of different behaviours (e.g., lever or food-well oriented) that happened to occur during the two types of trials. However, in the current Experiment 2, the two levers were both paired with the same reinforcer (food pellets or sucrose), and while this necessarily means that there is no difference in the value of the reinforcer that is paired with the levers, it remains possible that rats will be engaging in different behaviours during the two levers that would be subject to superstitious reinforcement. The results of Experiment 2 provide support for the explanation preferred by Patitucci et al. (2016): When the levers were paired with the same reinforcer there were significant correlations between food-well activity on the left and right levers on Block 6 ($r(30) = .94$, $p < .001$), and between lever-press activity on the two levers during the same block ($r(30) = .74$, $p < .001$). The fact that the extent to which phenotypic variation in sign- and goal-tracking behaviours is consistent across levers depends on whether they are paired with the same or different outcomes suggests that outcome value contributes to response selection.

To summarize, the results presented in this chapter indicate that individual differences in the topography of conditioned behaviour reflect the behaviour that is being measured and its properties. The observed differences reflect the operation of distinct associative processes that differ in their sensitivity to reward value and changes in contingencies, this being a property of the behaviour and not the animal, per se.

As I already mentioned in the introduction, a dual-process account is able to explain some results where individual differences are observed (e.g., differential sensitivity to change) but does not provide a satisfactory explanation of the origin of these differences. The evidence suggests that S-S learning is necessary for acquisition of sign- and goal-tracking, but the role of S-R learning is less clear. Sign-tracking displays the properties of an habitual system given the resistance to change, however we have seen that an association between the CS and a response is not necessary for the behaviour to emerge (e.g., Browne, 1976; Williams & Williams, 1969).

These results motivated the development of a new formal model of learning and performance, HeiDI (described in Chapter 3), in which the associative structures that are acquired during Pavlovian conditioning are integrated with an analysis of how the knowledge embodied in these structures determines the nature of the responses elicited by a CS. Neither of the dual-process accounts is able to explain the origin of the observed individual differences or to predict when or how the balance favours one system or the other. HeiDI, echoes the conclusions presented in Chapter 2 and at the same time distances itself from the dual-process account in explaining conditioned behaviour. The model is explained in detail the following Chapter.

Chapter 3

HeiDI: A model for Pavlovian learning and performance with reciprocal associations

Heidi, one of the world's most popular children's stories, was originally written by Johanna Spyri as two companion pieces: *Heidi: Her years of wandering and learning*, and *Heidi: How she used what she learned.* They describe how Heidi's predisposition to wander and learn was later evident in her behaviour. The central concern of the model developed here is the nature of the associative structures that are acquired during Pavlovian conditioning and how these structures result in their behavioural sequelae. Pavlovian conditioning is probably the best-known phenomenon in the history of the scientific study of psychology. The basic procedure and observations can be recounted by people with little or no other knowledge of the field: dogs given pairings of a ringing bell with food come to salivate when the bell rings. HeiDI is a significant revision of the model of Pavlovian conditioning developed by Rescorla and Wagner (Rescorla & Wagner, 1972; Wagner & Rescorla, 1972), and reflects Pavlov's vision that the study of conditioning provides associative psychology with a scientific basis (Pavlov, 1941, 171). Their model has had a profound and enduring influence on the field of animal learning (e.g., Mackintosh, 1975; McLaren, Kaye, & Mackintosh, 1989; Pearce & Mackintosh, 2010; Wagner, 1981), but also on psychology more broadly (e.g., Gluck & Bower, 1988; Kruschke, 1992; Rumelhart, Hinton, & Williams, 1986), and on neuroscience (e.g, Lee et al., 2018; Schultz, Dayan, & Montague, 1997); with 8667 citations at the time of writing this thesis. However, the Rescorla-Wagner model offers only the most rudimentary analysis of the associative structures that are acquired during conditioning and how these map onto changes in behaviour. Moreover, the model provides no explanation for recent evidence, where different behavioural indices of learning can be taken to support different conclusions about the strength of an association (e.g., Flagel, Akil, & Robinson, 2009; Flagel et al., 2011; Patitucci, Nelson, Dwyer, & Honey, 2016). This fundamental problem, together with others that I shall come to (e.g., Dickinson, Hall, & Mackintosh, 1976; Lubow, 1989; Miller, Barnet, & Grahame, 1995; Rescorla, 2000, 2001), provided the impetus for the development of HeiDI. The name of the model, HeiDI, reflects the literary reference and links to my surname and those of my supervisors to one of the principal issues that the model seeks to address: How excitation and inhibition determine ideo-motion.

3.1 The Rescorla-Wagner Model

The Rescorla-Wagner model proposes that Pavlovian conditioned behaviour reflects the formation of an association between the nodes activated by the conditioned stimulus (CS) and unconditioned stimulus (US). The presentation of the CS comes to associatively activate the representation or idea of the US and thereby behaviour, which can be thus considered ideo-motive: A seemingly reflexive movement effected in response to an idea, in this case the evoked memory of the US. The model has been fundamental to the development of theoretical treatments of associative learning for almost 50 years and has influenced neurobiological analyses of learning and memory. I briefly review the model here because it provides the principal source of inspiration for the new model that is developed in the remainder of Chapter 3.

According to the Rescorla-Wagner model, the change in the associative strength (V_{CS-US}) of a CS on a given trial is determined by the difference between the maximum associative strength supportable by a US (λ) and the pooled associative strength of all stimuli presented on that trial ($\sum\mathrm{V}_{Total-US}$). The global or pooled error term (λ - $\sum\mathrm{V}_{Total-US}$) allows the model to accommodate a wide range of phenomena: blocking (e.g., Kamin, 1969), conditioned inhibition, (e.g., Rescorla, 1969), contingency effects (e.g., Rescorla, 1968), overshadowing (e.g., Mackintosh, 1976), relative validity (e.g., Wagner, Logan, Haberlandt, & Price, 1968), superconditioning (e.g., Rescorla, 1971). These phenomena were beyond the scope of models with separate error terms for each component of a pattern of stimulation (e.g., Bush & Mosteller, 1951; Hull, 1943). It also provides an elegant integration of excitatory conditioning, where the memory of a CS provokes the memory of the US, and inhibitory learning, where a CS can reduce the likelihood of the US memory from becoming active when it otherwise would.

$$\Delta V_{CS-US} = \alpha_{CS-US}(\lambda - \sum V_{Total-US}) \quad (0)$$

Briefly, the pooled error term means that ΔV_{CS-US} is affected not only by the current associative strength of that stimulus (i.e., V_{CS-US}), but also by the presence of other stimuli that have associative strength (i.e., by $\sum V_{Total-US}$). According to the Rescorla-Wagner model, the change in associative strength driven by the discrepancy within the pooled error term (λ - $\sum V_{Total-US}$) is modulated by the product of two learning rate parameters, α_{CS} and β_{US}. Rescorla and Wagner (1972) note that "the value of α_{CS} roughly represents stimulus salience" and that "the assignment of different β_{US} values to different USs indicates that the rate of learning may depend on the particular US employed". The two learning rate parameters were confined to the unit interval: $0 \leq \alpha_{CS}, \beta_{US} \leq 1$, and enabled the model to capture the fact that the salience of the CS (α_{CS}) and nature of the US (β_{US}) affect the rate of excitatory learning[1]. Of particular note, however, is the fact that this model of Pavlovian conditioning did not address – in any systematic fashion – the influence of associative strength (i.e., V) on conditioned responding.

In developing their model and its application to experimental findings, Rescorla and Wagner (1972, p. 77) noted that it was "sufficient simply to assume that the mapping of Vs into magnitude or probability of conditioned responding preserves their ordering", and that any such mapping would inevitably depend on the details of each experimen-

---

[1] To enable inhibitory conditioning to occur on trials when the US is absent, Rescorla and Wagner (1972) assumed that β takes a positive value when the US is absent but the CS is present; with this value assumed to be lower than on trials when both the CS and US are present. This complexity is avoided in HeiDI.

tal situation and on "performance" factors. In a companion paper, when comparing conditioning involving a single CS with conditioning involving a compound of two CSs, they also noted "that the greater the number of cues which is made available, the more likely it is that the subject will be provided (and perhaps idiosyncratically so) with a single salient cue to which conditioning can rapidly occur" (Wagner & Rescorla, 1972, pp. 303-304). This statement acknowledges (parenthetically) the fact that individual differences might affect conditioning (see also, Pavlov, 1941, pp. 378-378) but there has been little appetite to address such differences (empirically or theoretically) and to move beyond simple (group level) assumptions about the translation of learning into performance (see also, Mackintosh, 1975; Miller & Matzel, 1988; Pearce, 1994; Pearce & Hall, 1980). However, there is now evidence demonstrating that the reliance on such assumptions can no longer be sustained; and nor can the idea that Pavlovian conditioning results in unconditioned responses snipped from the US being grafted onto the CS through a process of stimulus substitution (see Dwyer, Burgess, & Honey, 2012; Pavlov, 1927; Wagner & Brandon, 1989).

3.2 Individual differences

It has been discussed in Chapter 1, that even a simple auto-shaping procedure produces marked individual differences in behaviour: some rats predominantly interact with the CS (e.g., lever), others investigate the location where the reinforcer is about to be delivered, and the remainder show patterns of behaviour in between these two extremes (e.g., Patitucci et al., 2016). When lever activity is used as the assay of discrimination learning, the sign-tracking group show better learning than the goal-tracking group; but when food-well activity is used then the reverse is the case. That is, it is not possible to provide a mapping of Vs on to conditioned behaviour that provides a coherent interpretation: Focussing on one measure (e.g., sign-tracking) leads to the conclusion that associative learning had proceeded more readily in one set of rats than the other, while focussing on the second measure (e.g., goal-tracking) leads to the opposite conclusion.

As it stands, the Rescorla-Wagner model is unable to explain why, for any given rat, one response was stronger than the other, and why in some rats goal-tracking was stronger than sign-tracking whereas in other rats this relationship was reversed. In fact, these results pose a problem for any theory of learning that assumes a monotonic relationship between a single construct that represents learning and acquired behaviour (e.g., Gallistel & Gibbon, 2000; Stout & Miller, 2007).

3.3 HeiDI: Rationale, architecture and overarching assumptions

The purpose of HeiDI is to offer an account in which the associative structures that are acquired during Pavlovian conditioning are integrated with an analysis of how the knowledge embodied in these structures determines the nature of the responses elicited by a CS, and their relative strengths. In doing so, the model seeks to address challenges to the Rescorla-Wagner model, and other models of Pavlovian learning (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Wagner, 1981).

Figure 7 provides a schematic for the associative structures, to which I will align the analysis of the learning and performance equations that follow. The left-hand panel shows the structure of the model before conditioning has taken place and the right-hand panel shows the structure of the model after conditioning. Before conditioning, the CS is principally linked to a set of unconditioned responses (r1-r3; e.g., orienting, lever approach, rearing), whereas the US is weakly linked to a set of unconditioned responses (r4-r6; e.g., food-well approach, chewing, swallowing). Unconditioned links from the CS to r4-r6 and the US to r1-r3 are assumed to be very weak; and the weights of the lines between the CS and r1-r6 and between US and r1-r6 denote the relative strengths of these untrained or unconditioned links. In this way, a general distinction is adopted between CS-oriented responses (r1-r3) and US-oriented responses (r4-6; see Holland, 1977, 1984). Importantly, as a consequence of conditioning it is assumed that reciprocal CS-US and

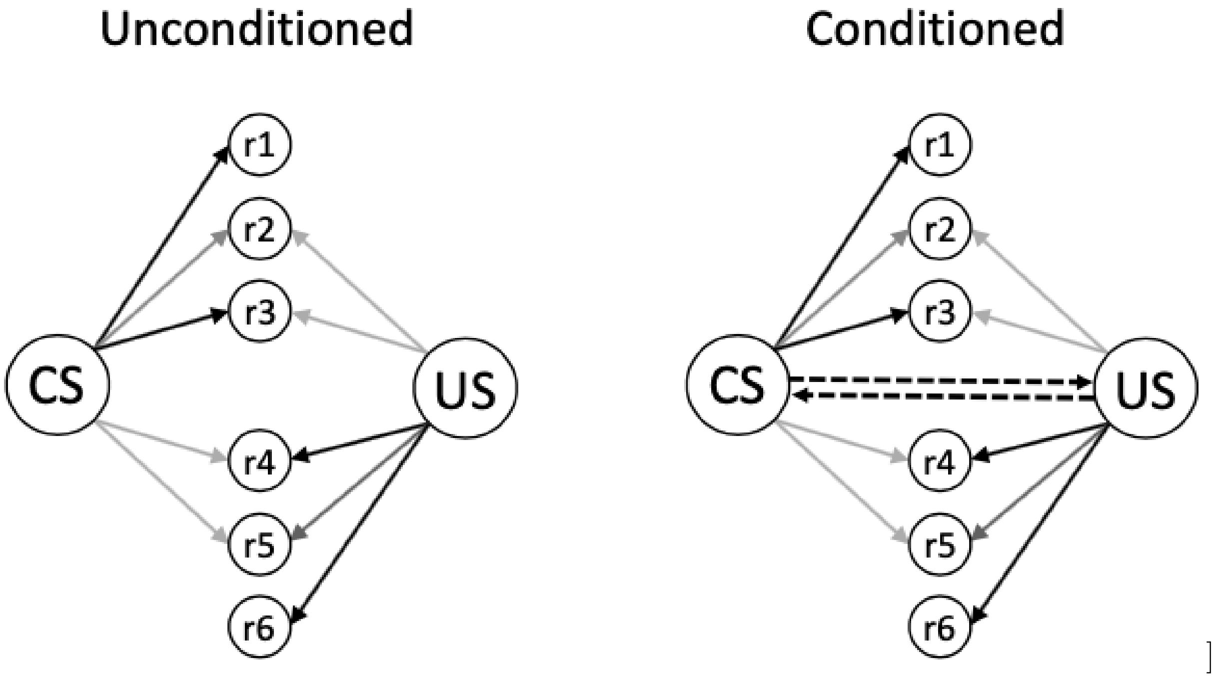


Figure 7: A schematic for associative structures that underpin the translation of excitatory learning into performance. The left-hand side depicts the model before conditioning (i.e., the unconditioned structure), with the darkness of the links indicating their strength, and the right-hand side depicts the model after conditioning (i.e., the conditioned structure). Conditioning results in changes in the strength of the reciprocal CS-US and US-CS associations between nodes activated by the CS and US (denoted by the hashed lines).

US-CS associations are acquired, which are depicted as the presence of hashed lines in the conditioned structure. The general rationale for this assumption, which does not feature in other formal models of Pavlovian conditioning (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972), is outlined next. A more specific justification is reserved until the learning rules for these reciprocal associations are presented. I will present in later sections how the inclusion of US-CS associations, as well as CS-US associations, provides the basis for HeiDI to explain a wide range of phenomena: In particular, those that have proven difficult to reconcile with the Rescorla-Wagner model, unequal change in the associative strengths of the components of a compound (e.g., Rescorla, 2000), downshift unblocking, (e.g., Dickinson, Hall, & Mackintosh, 1976) or that have been taken to provide support for models that have emphasized "predictiveness" (e.g., Mackintosh, 1975; Pearce & Hall, 1980).

The formation of reciprocal associations between the CS and US nodes creates a functional cell assembly and enables “resonance” between them: When the CS is presented activation propagates to the US node, which is propagated back to the CS (e.g., Grossberg, 1980).

Arcediano, Escobar, and Miller (2005), in a series of experiments with both human and non-human subjects, provided evidence of temporal integration between the presentation of two different training experiences (Experiment 1: outcome → S1 and S2 → S1; Experiment 2 (A-C): S2 → S1 and US → S1) based on the element common to both experiences (S1). This temporal integration suggests the formation not only of specific temporal forward associations (i.e., S2 → S1) but also of specific temporal backward associations (i.e., US $\leftarrow S1$).

There is further evidence that such reciprocal associations are acquired during forward conditioning in a variety of preparations (e.g., Asch & Ebenholtz, 1962; Cohen-Hatton, Haddon, George, & Honey, 2013; Honey & Bolhuis, 1997; Honey & Ward-Robinson, 2002); and a complementary literature on the conditions under which US-CS pairings result in conditioned responding to the CS (e.g., Barnet & Miller, 1996; Cole & Miller, 1999). At a theoretical level, in typical Pavlovian conditioning procedures – where the CS precedes but does not co-exist with the US – the memory trace of the CS must be sufficient to support the development of excitatory associations (cf. Barnet & Miller, 1996; Gallistel, 1990; Miller, Barnet, & Grahame, 1995; Wagner, 1981). Importantly, while the development of the CS-US association increases the likelihood that the presentation of the CS will activate the US and thereby provoke r4-r6, without the backward associations there would be little change in the likelihood that the CS would provoke r1-r3. The CS-US association allows the presentation of the CS to activate the US node and US-CS association allows activation of the US node to increase activation in the CS node, which increases the tendency for r1-r3 to become active as a consequence of conditioning.

When a CS is presented, there are two sources of information that are immediately available to an animal upon which performance could be based: The perceived salience of the CS, (which is related to α_{CS}) and the perceived salience of the US that is activated by the CS (which is related to V_{CS-US}). A fully effective CS is held to activate the US representation to the value of the perceived salience of the presented US (which relates to β_{US}). HeiDI assumes that both of these sources contribute to the nature of performance (cf. Hull, 1949). In particular, the model proposes that the relative values of α_{CS} and V_{CS-US} determine how learning is translated into performance through two values, R_{CS} and R_{US}. In advance of describing in detail how R_{CS} and R_{US} are calculated, for the time being simply assume that if $\alpha_{CS} > \mathrm{V}_{CS-US}$, then $\mathrm{R}_{CS} > \mathrm{R}_{US}$, whereas if $\mathrm{V}_{CS-US} > \alpha_{CS}$ then $\mathrm{R}_{US} > \mathrm{R}_{CS}$. Returning to Figure 7, R_{CS} affects behaviour via connections from the CS to r1-r6 in Figure 7, and R_{US} affects behaviour via connections from the US to r1-r6. The model assumes that the precise nature of the (alternative) responses generated in a given conditioning preparation will be a function of the interaction between the nature of the CS and US (Holland, 1977, 1984). In the next sections, I first present the learning rules used by HeiDI to determine the development of the reciprocal CS-US and US-CS associations (Equations 1 and 2); and provide a simple rule for combining these values upon presentation of the CS (Equation 3). It is worth briefly noting that Equations 1 and 2 embody the idea that the perceived salience of the CS and US, and their associatively generated counterparts, that determines learning. This suggestion is consistent with the idea that individual differences in the perceived salience of the CS and US play a central role in determining individual differences in the expression of learning. I then provide a detailed analysis of how this combined value is distributed in performance (Equations 4-6). The corresponding simulations of learning and performance are then presented in each experimental chapter. Finally, I illustrate how HeiDI provides a natural account for phenomena that challenge the Rescorla-Wagner model, and how it provides alternative analyses for results that have provided the basis for models of Pavlovian learning that include learnt changes in attention or associability (e.g., Mackintosh, 1975; Pearce & Hall, 1980).

3.4 Learning rules

The use of a pooled error term was the central contribution of the Rescorla-Wagner model, allowing it to provide a ready account of the conditions under which excitatory and inhibitory learning occur.

$$\Delta V_{CS-US} = \alpha_{CS}(c * \beta_{US} - \sum V_{Total-US}) \quad (1)$$

$$\Delta V_{US-CS} = \beta_{US}(c * \alpha_{CS} - \sum V_{Total-CS}) \quad (2)$$

HeiDI adopts versions of the pooled error term in Equation 1 and Equation 2, for the formation of CS-US and US-CS associations, respectively (a list of the HeiDI equations is available at the end of the chapter). There is recent evidence that provides direct support for this feature of HeiDI in the context of CS-oriented behaviour and US-oriented behaviour: A lever CS that provokes sign-tracking can block the acquisition of goal-tracking to an auditory CS, and an auditory stimulus that provokes goal-tracking can block acquisition of sign-tracking to a lever CS (e.g., Derman, Schneider, Juarez, & Delamater, 2018). However, as I shall show, while Equations 1 and 2 have formally equivalent pooled error terms, their functional properties differ when a stimulus compound (AB) is paired with a US: Equation 1 functions as a pooled error term for the A-US and B-US, whereas Equation 2 functions as a separate error term for the US-A and US-B associations. Briefly, I will later show how this observation enables the use of a pooled error term to be reconciled with results showing that compound (AB) conditioning can result in unequal changes in conditioned responding to A and B depending on their prior training histories (e.g., Allman, Ward-Robinson, & Honey, 2004; Rescorla, 2000, 2001).

An important feature of Equation 1 is that the perceived salience of the US (β_{US}) sets the maximum perceived value of the US retrieved by the CS (i.e., V_{CS-US}). Similarly, the perceived salience of the CS in Equation 2 (α_{CS}) sets the maximum perceived value of the CS retrieved by the US (i.e., V_{US-CS}). The idea that the perceived salience of the directly activated and associatively activated representations of the stimuli (CS and US) determine associative change receives direct support from results reported by Dwyer, Figueroa, Gasalla, and López (2018). They examined the development of a flavour preference through pairing a flavour CS with an 8 % sucrose US. They observed that preceding this concentration of sucrose by either 2% sucrose (generating positive contrast) or 32% sucrose (generating negative contrast) affected the acquisition of the flavour preference: The flavour preference supported by 8% sucrose was larger when it was preceded by 2% sucrose than when it was preceded by 32% sucrose. Moreover, when the changes in the perceived salience of the US (8% sucrose) produced by contrast were directly assessed, through the analysis of licking microstructure, they directly correlated with the size of the resulting preference for the CS flavours.

3.4.1 Excitatory learning and error correction

Equations 1 and 2 are symmetrical rules governing the formation of CS-US and US-CS associations, respectively. Equation 1 represents a simplification to the Rescorla-Wagner learning rule (Equation 0) and determines the formation of CS-US associations; and Equation 2 provides the formally equivalent rule for US-CS associations. While Equations 1 and 2 are formally equivalent pooled error correcting rules, they have quite different functional properties in conventional conditioning procedures in which a compound of two CSs (AB) precedes a US. In short, Equation 1 functions as a pooled error term in conventional compound conditioning procedures (Rescorla & Wagner, 1972) whereas Equation 2 functions as a separate error term in such procedures (Bush & Mosteller, 1951; Hull, 1943). However, it is also worth noting that the model predicts that if a single CS were to be followed by a compound of two USs (US1 and US2), then the association of US1 with the CS would be weaker than if US1 had been paired with the CS in isolation. The prediction that there will be competition or overshadowing between the capacity of two USs to become associated with a single CS has received empirical support (see Miller & Matute, 1998).

In Equation 1, α_{CS} is a learning rate parameter confined to the unit interval $0 \leq \alpha_{CS} \leq 1$, and c*$\beta_{US}$ determines the asymptote for the CS-US association; whereas in Equation 2, β_{US} is a learning rate parameter confined to the unit interval $0 \leq \beta_{US} \leq 1$, and c*α_{CS} determines the asymptote for the US-CS association. Note that α_{CS} and β_{US} are dimensionless scalars, but when they serve as the asymptotes for associative strength they are multiplied by a constant of 1 in units of V (c). The requirement for c is that has units of V, but the numeric value is not fixed by this requirement. The constant c is assumed here to be 1 for simplicity. However, the asymptotic limits of learning need not be β_{US} in Equation 1 or α_{CS} in Equation 2, but some multiple of these values. When the CS is absent, α_{CS} and c*α_{CS} are set to 0 and when the US is absent β_{US} and c*β_{US} are set to 0. In keeping with the Rescorla-Wagner model, α_{CS} and β_{US} are assumed to reflect the perceived salience of the CS and US, respectively. According to Equation 1, the strength of the association from the CS to the US (i.e., V_{CS-US}) converges asymptotically on c*β_{US}. The change in the strength of the association between CS and the US on a given trial (ΔV_{CS-US}) is determined by the error or difference within the pooled error term (c*$\beta_{US} - \sum V_{Total-US}$); and $\sum V_{Total-US}$ denotes the net associative strength of all of the stimuli presented on that trial. During simple CS-US pairings, excitatory learning ceases when $\sum V_{Total-US} =$ c*β_{US}; and the learning rate parameter α_{CS} affects the rate at which V_{CS} approaches c*β_{US}. In this case, the pooled error term means that the acquisition of associative strength by a given stimulus will be influenced by the associative strength of other stimuli that accompany it; for example, when a compound of two stimuli (A and B) is paired with a US.

Equation 2 is the complementary learning rule governing the formation of the US-CS association. The change in the strength of this association ΔV_{US-CS} on a given trial is also determined by the discrepancy within the pooled error term (c*$\alpha_{CS} - \sum V_{Total-CS}$).

α_{CS} reflects the salience of the CS, and $\sum$V$_{Total-CS}$ denotes the associative strength of the US (in typical conditioning procedures). Learning ceases when $\sum$V$_{Total-CS}$ = c*α_{CS}, and the learning rate parameter β_{US} affects the rate at which V$_{US-CS}$ approaches c*α_{CS}. Because in a typical Pavlovian conditioning procedure there is only one US (cf. Miller & Matute, 1998), the c*α_{CS} value of each CS in a compound (e.g., A and B) sets the asymptote for the association from the US to that CS. This means that the US-CS associations will proceed independently for each of the components of a compound that is paired with a US. That is, while Equation 1 has both the formal and functional properties of Equation (0) and predicts the same phenomena as that model, Equation 2 has equivalent formal properties, but functions in the same way as a separate error term during compound conditioning (e.g., Bush & Mosteller, 1951; Hull, 1943). How could one test whether the analysis provided by Equations 1 and 2 is accurate?

Consider first the simple case in which two CSs (A and B) are presented together and paired with a US. Under these conditions, the associative strength accrued by A (V$_{A-US}$) and B (V$_{B-US}$) will be less than if these stimuli had been separately paired with the US. An effect known as overshadowing (e.g., Mackintosh, 1976). However, the state of affairs will be different for the reciprocal associations (i.e., V$_{US-A}$ and V$_{US-B}$). They will undergo the same change in associative strength as they would have done had conditioning with each occurred in isolation; because the c*α_A and c*α_B for stimulus A and B set separate asymptotes for the association between the US-A and US-B. Of course, the finding that overshadowing is observed under such conditions is uninformative; because Equations 0 and 1 prediction that V$_{A-US}$ will be lower when it has been conditioned in compound with B than when it has been conditioned alone. But, now imagine the same compound conditioning scenario, but that on this occasion a previous stage of training had established A as conditioned exciter (by pairing it with a US) and B had been established as a conditioned inhibitor (by pairing it paired with the absence of an otherwise predicted US). According to Equations 0 and 1, provided A and B are equally intense (i.e., α_A= α_B) then they should gain equivalent associative strength as a consequence of the AB compound being paired with the US. However, according to Equation 2, while the association between the US and A will not increase (having reached asymptote during the first stage) the association between the US and B will increase, because the US had not previously been paired with B. If the changes in the reciprocal associations were to be combined, then B should have gained greater combined associative strength than A. Rescorla (2000, 2001) has published a series of ingenious experiments that has confirmed this prediction under a variety of circumstances(see also Allman & Honey, 2005; Allman, Ward-Robinson, & Honey, 2004; Holmes, Chan, & Westbrook, 2019). I will provide a formal simulation of the analysis of these results, which have been taken to implicate separate error terms in Pavlovian conditioning, once the rules for combining the reciprocal associations has been described, and the way in which associative strength affects performance presented.

As just noted, CS-US pairings create a functional cell assembly through reciprocal associations between the CS and US. To capture this interaction and to simplify the performance rules, it is desirable to combine the net associative strengths of the CS-US association returned by Equation 1 (for V_{CS-US}) and the US-CS association returned by Equation 2 (for V_{US-CS}). The combined associative strength within this assembly (V_{COMB}) is given by Equation 3a[2]. The combined associative strength of a compound stimulus ($\mathrm{V}_{COMB-AB}$) composed of two CSs (A and B) is given by Equation 3b; in which $\sum \mathrm{V}_{AB-US}$ is the sum of V_{A-US} and V_{B-US}, and V_{US-A} and V_{US-B} are the strengths of the associations between the US and A, and the US and B.

$$V_{COMB} = V_{CS-US} + (\frac{1}{c} * V_{CS-US} * V_{US-CS}) \quad (3a)$$

$$V_{COMB-AB} = \sum V_{CS-US} + [\frac{1}{c} * \sum V_{AB-US} * (V_{US-A} + V_{US-B})] \quad (3b)$$

---

[2] The symmetrical combination rules can be used if the US (rather than the CS) was tested alone, e.g., V_{US-CS} + (1/c * V_{US-CS} * V_{CS-US})

This choice of combination rule recognizes the fact that while the CS directly activates the CS-US association, the US-CS association is only indirectly activated by the presentation of the CS. The rule has the general property that the directly activated link in a chain of associations will constrain the impact of the indirectly activated link on performance. This is true regardless whether the stimulus is a CS or a US. If the CS activates the representation of the US, V_{CS-US} will constrain the impact on performance of V_{US-CS}, however if the US activates the representation of the CS, V_{US-CS} will constrain the impact on performance of V_{CS-US}. For simplicity, when presenting a CS, if V_{CS-US} was ≈ 0 and V_{US-CS} was positive, then V_{COMB} (and hence both R_{CS} and R_{US}) ≈ 0 in spite of the fact that the relationship between the CS and US had been encoded (i.e., as V_{CS-US}). The significance of this property in the context of HeiDI will become apparent when the blocking phenomenon is considered in greater detail. (e.g., Kamin, 1969).

3.4.2 Extinction

When conditioning trials with a CS are followed by extinction trials where the CS is presented, but no US occurs, $c^{*}\beta_{US}$ is set to 0 and $\sum V_{Total-US}$ will be positive. Under these conditions, Equation 1 returns a negative value for ΔV_{CS-US}, but Equation 2 returns 0 for ΔV_{US-CS} (because $\beta_{US}= 0$). It is worth highlighting this asymmetry between what is learned during conditioning and extinction: excitatory learning involves changes to V_{CS-US} and V_{US-CS}, but conventional extinction procedures involve only changes to V_{CS-US}. The negative values returned by Equation 1 during extinction can be interpreted in two ways: First, they could denote the growth of negative associative strength (Konorski, 1948; Rescorla & Wagner, 1972; Wagner & Rescorla, 1972). Second, they could denote the formation of an excitatory association between the CS and a 'No US' node, which in turn inhibits the US node and thereby reduces conditioned behaviour (see Konorski, 1967; Pearce & Hall, 1980; Zimmer-Hart & Rescorla, 1974). In the first case, the negative values are directly reflected in the underpinning associative structure, and in the second case they reflect the product of an excitatory CS-No US association multiplied by an inhibitory No US-US association. However, according to both interpretations, the net associative strength of the forward association involving the CS (V_{CS-US}) is the sum of the positive and negative associative values returned by Equation 1; and the net associative strength of V_{CS-US} is the sum of the positive and negative values returned by Equation 2. Negative values of V_{US-CS} will be returned by Equation 2 when $\sum V_{Total-CS} > c^{*}\alpha_{CS}$. This situation would arise when the US is presented alone after conditioning has taken place or if additional USs were presented in the inter-trial intervals between CS-US pairings. Both of these manipulations result in a reduction in conditioned responding to the CS (see Rescorla, 1968, 1973)[3]. In fact, according to HeiDI while both of these manipulations will result in extinction of the US-CS association, adding US presentations during the intervals between CS-US trials allow the formation of a context-US association, which should block the development of the CS-US association. In keeping with this analysis, it has been argued that the effects of manipulating CS-US contingency, by adding US alone presentations during conditioning, might be multiply determined (Baker & Mercier, 1982).

Later simulations will confirm the description of the consequences of extinction presented in the previous paragraph. For the time being, it is important to note that according to Equations 1 and 2, extinction leaves a significant contribution to performance completely unchanged (i.e., the US-CS association, V_{US-CS}), rather than simply being counteracted by additional inhibitory learning, as is the case with the net CS-US association (V_{CS-US}). This feature of HeiDI is consistent with the general observation that post-extinction manipulations can reveal the presence of residual excitation in performance, which has represented an ongoing challenge to the Rescorla-Wagner model (e.g., Bouton, 2004). Manipulations that enable the US to be activated (or that disrupt

---

[3] There is also evidence that when CS-US pairings are followed by separate presentations of the same US but at a higher salience (called US inflation) the CR to the CS is amplified (Bouton, 1984; Rescorla, 1973). Under these conditions, in addition to any reduction in net V_{US-CS}, presentations of a higher salience US might change the response units activated by the US, which could affect later performance to the CS.

the CS-No US association) will result in a return in performance to the CS.

3.4.3 Inhibitory learning

If conditioning trials in which stimulus A is paired with a US are intermixed with trials on which A is presented with stimulus B and the US is not delivered, then non-reinforced AB trials will result in a reduction in the net associative strength of A, and B will become a net inhibitor. The net associative strength of AB is given by adding the positive and negative values returned by Equation 1 for stimulus A and B. The net associative strength of the US, V$_{US-CS}$, is the sum of the positive and negative associative values returned by Equation 2. According to Equation 2, on non-reinforced AB trials there will be no change in the US-A or US-B associations; again because β_{US} = 0. However, inhibitory learning can also be produced if AB is paired with a US that is smaller in magnitude than the US that is paired with A (e.g., Cotton, Goodall, & Mackintosh, 1982). Under these conditions, $\beta_{US} > 0$ and HeiDI predicts that there would be an increase in the excitatory strength of the US-A and US-B associations, which would contribute to the values of V$_{COMB}$ for A, B and AB. The prediction that conventional conditioned inhibition training and conditioned inhibition produced by a reduction in reinforcer magnitude result in different association structures has not been evaluated.

3.5 Performance rules

When a CS is presented there are two sources of information that are available to an animal, the perceived salience of the CS (related to α_{CS}) and the associative strength of that CS (V$_{CS-US}$), which can be considered an estimate of β_{US} given its relationship with relationship with c*β_{US}. These two sources of information are held to determine the nature of conditioned behaviour.

$$R_{CS} = \frac{c * \alpha_{CS}}{c * \alpha_{CS} + |V_{CS-US}|} V_{COMB} \quad (4)$$

$$R_{US} = \frac{|V_{CS-US}|}{c * \alpha_{CS} + |V_{CS-US}|} V_{COMB} \quad (5)$$

Equation 4 and Equation 5 generate two values, R$_{CS}$, and R$_{US}$, which reflect the proportions of α_{CS} and V$_{CS-US}$. As before, in these equations, α_{CS} is multiplied by c in order to translate it into units of V. To determine the extent to which these proportions affect the absolute levels of performance they are multiplied by the overall associative properties of the CS-US ensemble (i.e., V$_{COMB}$). R$_{CS}$ affects behaviour via connections from the CS to r1-r6, and R$_{US}$ affects behaviour via connections from the US to r1-r6 (see Figure 7). Because in the simulations presented here net V$_{CS-US}$ > 0, the real values of V$_{CS-US}$ can be used to determine R$_{CS}$ and R$_{US}$ in Equations 4 and 5. However, to address the fact that Equation 1 (and Equation 2) can return negative values, the use of absolute values ensures that the proportions in Equations 4 and 5 are ≤ 1. This choice also leaves open the possibility that a net inhibitor could provoke responding when presented alone (cf. Konorski, 1967; Pearce & Hall, 1980), rather than having no effect on performance unless it is presented with an excitor (Konorski, 1948; Wagner & Rescorla, 1972). For now, it is sufficient to note that Equation 4 returns a higher value for R$_{CS}$ as the value of c*α_{CS} increases relative to the value of V$_{CS-US}$, and Equation 5 returns a higher value for R$_{US}$ as V$_{CS-US}$ increases relative to c*α_{CS}. These two equations are readily extended to accommodate stimulus compounds (AB). To do so, the c*α values for A and B are simply combined (e.g., added) to form c*α_{AB}, and the net Vs of A and B are combined (e.g., added) to form V$_{AB-US}$. Similarly, a given stimulus (CS or US) can be conceived of as a set of elements with their own c*α values and net Vs, which could be entered into Equations 4 and 5 using the same approach (cf. Atkinson & Estes, 1963; Delamater, 2012; Wagner & Brandon, 1989).

While Equations 4 and 5 provide a simple basis for the distribution of the associative properties of the CS-US ensemble (i.e., V$_{COMB}$) to the response-generating units (r1-r6) though R$_{CS}$ and R$_{US}$, they do not specify how these response units once activated affect behaviour. One simple possibility is that a given value of R$_{CS}$, for example, results in the same amount of CS-oriented responding (r1-r3) irrespective of the value of R$_{US}$. This possibility equates to there being parallel activation of the response-generating units (r1-r6), and is formally expressed in Equation 6, where R$_{CS}$ and R$_{US}$ are translated into dimensionless values by being multiplied by the reciprocal of the constant, “c”. According to Equation 6, the activation of a given response unit (e.g., r1) is simply determined by adding the products of (i) multiplying R$_{CS}$ by the connection between the CS and r1 (CS-r1), and (ii) multiplying R$_{US}$ by the strength of connection between the US and the same response unit (e.g., US-r1). There are more complex possibilities involving the interaction between the activation values of the response-generating units (e.g., McClelland & Rumelhart, 1981), but for now this serves as simple place holder for future theoretical elaboration.

$$r1 = (\frac{1}{c} * R_{CS} * CS_{r1}) + (\frac{1}{c} * R_{US} * US_{r1}) \quad (6)$$

The simulations presented in later sections are derived from Equations 1-5. Equation 6 simply involves multiplying the resulting R$_{CS}$ and R$_{US}$ values by the fixed strength links (with values between 0 and 1) between the CS and US nodes and the response units (e.g., the CS-r1 and US-r1 links). If the two sets of links are equivalent (see Figure 7), then differences in activation of the response units will depend solely on R$_{CS}$ and R$_{US}$.

3.5.1 Individual differences in β_{US}

HeiDI assumes that α_{CS} and β_{US} are fixed for a given CS and US in a given animal, but propose that the perceived salience of the CS (relating to α_{CS}) and US (relating to β_{US}), and hence c*α_{CS} and V$_{CS-US}$ in Equations 4 and 5, can vary between animals. This assumption provides the basis for individual differences in R$_{CS}$ and R$_{US}$, because c*α_{CS} and V$_{CS-US}$ affect performance according to Equations 4 and 5 (remember V$_{CS-US}$ converges on c*β_{US} at asymptote)[4]. This analysis receives support from the observation that rodents who showed a strong liking for sucrose (as measured by licking microstructure; see Dwyer, 2012) are more likely to be goal-trackers (when sucrose was the US) than those who exhibited a weaker liking for sucrose (Patitucci et al., 2016). Individual variation in the palatability of sucrose can be readily aligned to differences in β_{US} that will affect both learning (i.e., the asymptotic value of V$_{CS-US}$) and the rate at which V$_{US-CS}$ reaches asymptote (through Equations 1 and 2) and the distribution of V$_{COMB}$ in performance (through V$_{CS-US}$ in Equations 3-6). Moreover, Dwyer et al. (2018) showed that individual differences in the palatability of sucrose (during their experiments involving contrast effects) were positively correlated with the flavour preference learning.

There is additional evidence that is consistent with the proposition that β_{US} for different USs varies between animals, and indeed within a given animal: When separate presentations of two levers are paired with the same US (e.g., food or sucrose) then the bias towards sign-tracking or goal-tracking on one lever correlates with the bias on the other, results presented in Experiment 2. However, when the presentation of one lever is paired with sucrose and the other lever is paired with food, then there is no correlation between the bias on the two levers (Patitucci et al., 2016). This pattern of results is consistent with the view that the β_{US} values for two USs (i.e., food and sucrose) can vary between animals and within a given animal.

Further evidence. A central proposition of HeiDI is that variation in V$_{CS-US}$ interacts with c*α_{CS} to determine performance. This proposition receives support from the results presented in Experiment 1, in which a CS is first paired with a US is then presented alone across a series of trials. Extinction trials should affect net V$_{CS-US}$, conditional on

---

[4] Equations 4 and 5 can be transformed for the case in which the US is presented alone: Under these conditions, c*β_{US} replaces c*α_A and V$_{US-CS}$ replaces V$_{CS-US}$

the reduction of c*β_{US} from a positive value to 0 in Equation 1, but not α_{CS}. The clear prediction is that while both R$_{CS}$ and R$_{US}$ should decrease during extinction (V$_{COMB}$ will reflect the reduction in V$_{CS-US}$; see Equation 3), Equations 4 and 5 predict that this decrease will be less marked for R$_{CS}$ than for R$_{US}$: α_{CS} will remain the same and V$_{CS-US}$ will be lower. This prediction was confirmed in rats that were designated as either sign-trackers or goal-trackers in Experiment 1. In both groups, the tendency for rats to interact with the lever (sign-tracking) extinguished less rapidly than interacting with the food-well (goal-tracking).

The results from a related conditioning preparation provide converging evidence from the proposed interaction between c*α_{CS} and V$_{CS-US}$ in determining R$_{CS}$ and R$_{US}$. Kaye and Pearce (1984) gave rats presentations of a localized light that was either paired with the delivery of a food pellet on every trial (in group continuous) or on a randomly scheduled 50% of occasions on which it is presented (in group partial). They observed that when the light was continuously reinforced it maintained a higher level of goal-tracking (food-well entries) and a lower level of sign-tracking (orienting and approach to the light) than when the light was partially reinforced (see also, Anselme, Robinson, & Berridge, 2013). According to Equation 1 and 2, net V$_{CS-US}$ will be higher during a continuous than a partial reinforcement schedule, and a continuous reinforcement should result in a greater bias towards goal-tracking (R$_{US}$) and a smaller bias towards sign-tracking (R$_{CS}$) than partial reinforcement, which could result in the opposite bias (see Equations 4 and 5). However, Kaye and Pearce (1984) also observed that sign-tracking was higher in absolute terms during partial than continuous reinforcement. This finding might reflect the fact that high levels of goal-tracking, during continuous reinforcement, were more likely to interfere (at the level of response output) with sign-tracking than the lower levels of goal-tracking engendered by partial reinforcement (see discussion of Equation 6). In any case, the fact that CS-oriented behaviour is maintained by partial reinforcement should also improve an animal's later ability to detect new relationships involving that CS (cf. Pearce & Hall, 1980; Wilson, Boumphrey, & Pearce, 1992).

3.6 Simulations of learning and performance

Simulations were done in a Shiny dashboard app (HeiDI) using R programming language. The code is available on the Open Science Framework (OSF; https://osf.io/j8tps/?view_only=c7e9286d80464cf1a244c9cface69b4e). The simulation figures presented in the thesis are based on the model functions created in the app ('model.R'), however the graphical representation has been adapted to include several simulations in the same figure. The code for the simulations presented in the thesis can be found on the OSF website (https://osf.io/h4fyg/), more specifically at 'Thesis/Simulations'.

The software used for the app was R (version 3.3.3) and RStudio (Version 1.1.463.). To reproduce the simulations download and install R software for the appropriate operating system. After successfully installing R, download and install RStudio. Next, open *'install_packages.R'* file and run the file which installs all the needed packages. The online live version of the app includes a 'Readme' (README folder) file with detailed instructions for reproducing the code. This document has a time stamp (every time it is compiled the date is updated) and any changes from a previous version are detailed in 'log' folder. Each update of the app has and will have a numbered log file. In the Appendix (HeiDI app) I include a more detailed explanation of the app structure and files.

3.6.1 Excitatory conditioning and extinction

Figure 8 shows simulations of acquisition of conditioning (left side of each panel, Trials 1-6) and extinction (right-side of each panel, Trials 7-12). Panels A and C depict simulations of the development of the CS-US association derived from Equation 1 ($\mathrm{V}_{\text{CS-US}}$), the US-CS association derived from Equation 2 ($\mathrm{V}_{\text{US-CS}}$) and their combined values (V_{COMB}). Maximum $\mathrm{V}_{\text{CS-US}}$ is determined by c*β_{US} and maximum $\mathrm{V}_{\text{US-CS}}$ is determined by c*α_{CS}. The learning rate at which $\mathrm{V}_{\text{CS-US}}$ reaches the asymptote is determined by c*α_{CS} (Equation 1). The learning rate at which $\mathrm{V_{US\text{-}CS}}$ reaches the asymptote is determined by c*β_{US} (Equation 2). Panels B and D show how Equations 4 and 5 translate learning into the performance values of $\mathrm{R_{CS}}$ (CS-oriented responses) and $\mathrm{R_{US}}$ (US-oriented responses) across a series of CS-US pairings.

Simulations were conducted in which either R_{CS} dominated performance during conditioning (Panel A and B; α_{CS} = .50 and β_{US} = .30; this bias would be consistent with a sign-tracker) or R_{US} dominated performance (panels C and D; α_{CS} = .30 and β_{US} = .50; this bias would be consistent with a goal-tracker). When $\alpha_{CS} > \beta_{US}$, $\mathrm{R_{CS}}$ dominated $\mathrm{R_{US}}$, but when $\beta_{US} > \alpha_{CS}$ then the reverse is the case. The general conclusion is that if c*$\alpha_{CS} = \mathrm{V_{CS\text{-}US}}$, then Equations 4 and 5 return the same or similar contributions for $\mathrm{R_{CS}}$ and $\mathrm{R_{US}}$; but if c*$\alpha_{CS} \neq \mathrm{V_{CS\text{-}US}}$ then the component with the largest value (c*α_{CS} or $\mathrm{V_{CS\text{-}US}}$) contributes proportionately more to performance.

In all of the simulations that follow, it is assumed that the constant (c) is 1 in units of V. Therefore, the values of α_{CS} and β_{US} are the same as those of c*α_{CS} and c*β_{US}, respectively. For ease of presentation, I will only refer to α_{CS} and β_{US} without explicit reference to the transformations that appear in Equations 1-6. Starting with panels A and B, when α_{CS} and β_{US} have been arranged for R_{CS} to dominate performance, it is clear from panel A that during conditioning $\mathrm{V}_{US-CS} > \mathrm{V}_{CS-US}$ (when α_{CS} = .50 and β_{US} = .30), and that V_{COMB} is similar to V_{US-CS}. During extinction, inspection of panel A shows that V_{CS-US} and V_{COMB} decline, but V_{US-CS} does not (because β_{US} = 0). Panel B shows that the reduction in R_{CS} is (numerically) less marked than R_{US}. Moving to panels C and D, when α_{CS} and β_{US} have been arranged for R_{US} to dominate performance, during conditioning $\mathrm{V}_{CS-US} > \mathrm{V}_{US-CS}$ and $\mathrm{V}_{COMB} > \mathrm{V}_{CS-US}$. During extinction, V_{CS-US} and V_{COMB} decline, but V_{US-CS} does not. Panel D shows that the reduction in R_{CS} occurs much less rapidly than the reduction in R_{US}. For both types of simulated rats, STs (Panel A and B) and GTs (Panel C and D), HeiDI predicts that US-oriented behaviour (goal-tracking) decline more rapidly than CS oriented behaviour (sign-tracking).

In Experiment 1, rats received training in which one lever, (L_1) was followed by a reinforcer and another lever (L_2) was non-reinforced. During the reversal stage the contingencies on the lever were reversed: L_1 was non-reinforced and L_2 was now reinforced. The results of Experiment 1 confirm HeiDI's prediction, goal-tracking changed more rapidly than sign-tracking when the contingencies were reversed. Note that in the simulations, the effects of reversing contingencies at the beginning of trial 6, in the simulations are evident in trial 7. This is because trial 7 shows the changes in associative strength derived from events during trial 6.

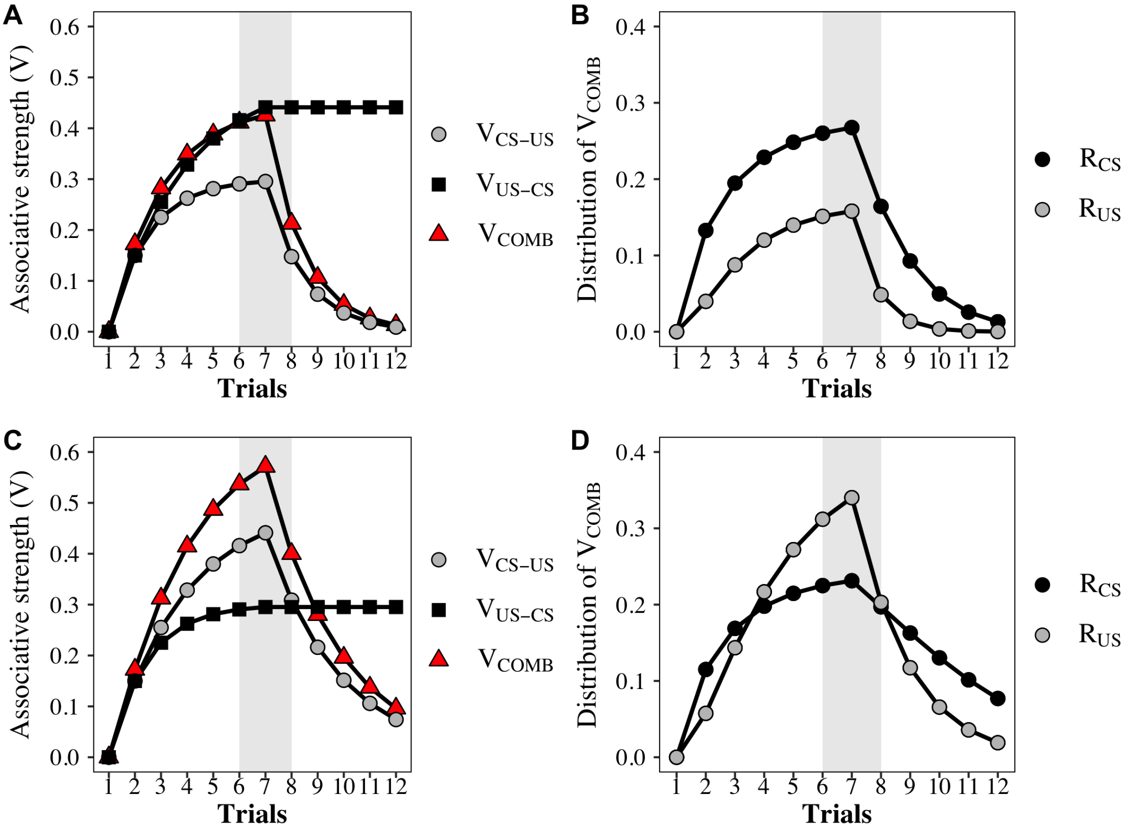


Figure 8: Simulations of the effects of an extinction procedure on sign-tracking and goal-tracking: Simulation for conditioning (trials 1-6) and extinction (trials 7-12) for the reinforced (+) lever. Panels A and C depict the output values for $V_{\text{CS-US}}$, $V_{\text{US-CS}}$, V_{COMB}, and panels B and D show the corresponding output values for R_{CS} and R_{US}. The parameters during conditioning were chosen to result in a bias towards R_{CS} (i.e., $\alpha_{CS} = .50$ and $\beta_{US} = .30$; panels A and B) or a bias towards R_{US} (i.e., $\alpha_{CS} = .30$ and $\beta_{US} = .50$; panels C and D). During extinction, β_{US} was set to 0.

3.6.2 Change in US value

In Experiment 2 rats received during training separate presentations of two levers (L_1 and L_2), that were both paired with the same reinforcer during training (either food pellets or sucrose). Work in our laboratory indicated that food pellets maintain higher levels of both lever pressing and food-well entries than does sucrose; which should be evident in the first stage of training. During the second stage, the reinforcers associated with the two levers were switched: the rats given pellets during training received sucrose during the switch and those given sucrose during training received food pellets during the switch. Because the transition between reinforcers does not affect the stimulus (α_{CS}) in any obvious way, changes in performance must reflect the increase or decrease in β_{US} affects performance. The change from sucrose (low β_{US}) to food pellets (high β_{US}) increase β_{US}, whereas the change from food pellets (high β_{US}) to sucrose (low β_{US}) decrease β_{US}. Results from Experiment 2 confirm this prediction as it has been seen that goal-tracking was more affected by a change in US value than sign-tracking.

Figure 9 shows simulations of acquisition of conditioning (left side of each panel, Trials 1-6) and switch in US value (right-side of each panel, Trials 7-12). Panels A and C depict the output values for $V_{\text{CS-US}}$, $V_{\text{US-CS}}$, V_{COMB}, and panels B and D show the corresponding output values for R_{CS} and R_{US}. The parameters were chosen to result in either an increase in US value after the switch ($\alpha_{CS} = .50$, β_{US} (before) $= .30$ and β_{US} (after) $= .50$; Panel A and B) or a decrease in US value ($\alpha_{CS} = .50$, β_{US} (before) $= .50$ and β_{US} (after) $= .30$; Panel C and D). For HeiDI manipulations that change β_{US} should affect goal-tracking and, as a secondary consequence, sign-tracking. Inspection of panels B and D shows that R_{US} is more affected by the change in US value than R_{CS}. These simulations are implemented in an open source app HeiDI app. More details about how I conducted the simulations can be found in the Appendix, section Simulations.

In Chapter 4 I present and discuss simulations derived from HeiDI for inhibitory conditioning and individual differences in US value. For now, I discuss how using a pooled error term in Equations 1 and 2 enables HeiDI to explain effects that have posed a challenge for Rescorla-Wagner model.

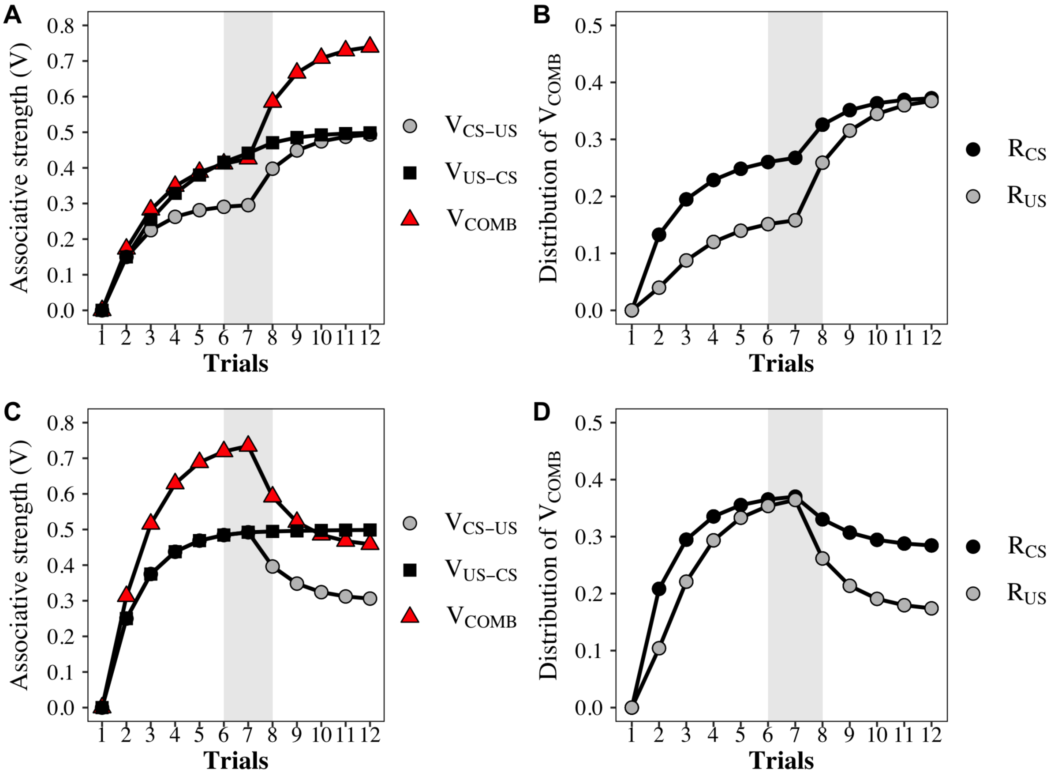


Figure 9: Simulations of the effects of a change in US value on sign-tracking and goal-tracking: Simulation for conditioning (trials 1-6) and after the switch in US value (trials 7-12). Panels A and C depict the output values for $V_{\text{CS-US}}$, $V_{\text{US-CS}}$, V_{COMB}, and panels B and D show the corresponding output values for R_{CS} and R_{US}. The parameters were chosen to result in either an increase in US value after the switch ($\alpha_{CS} = .50$, β_{US} (before) $= .30$ and β_{US} (after) $= .50$; Panel A and B) or a decrease in US value ($\alpha_{CS} = .50$, β_{US} (before) $= .50$ and β_{US} (after) $= .30$; Panel C and D).

3.6.3 Compound conditioning and the pooled error term

HeiDI provides a potential reconciliation of the use of a pooled error terms with the observation that stimuli with different associative histories appear to undergo unequal change when they are conditioned in compound; an observation that was taken to be inconsistent with the Rescorla-Wagner model and its successors (see Holmes et al., 2019). To recap: In one set of experiments, Rescorla (2000) initially trained rats with two excitatory stimuli (A and C), each followed by reinforcement (A+, C+) as well as two inhibitory stimuli (B and D) each non-reinforced in the presence of X (X+; XB-; XD-). Then a compound consisting of one excitor and one inhibitor is reinforced (AB+).

Testing is conducted with two compounds each containing one excitor and one inhibitor. The test compounds differ in whether they contain the excitor (AD) or inhibitor (BC) from the AB reinforcement phase.

Let assume that A and C both had excitatory associative strength of .50, and B and D both had inhibitory associative strength of -.50 before the compound, AB, was paired with the US (i.e., AB -> US). According to Equations 0 and 1, the associative strength of both should increase an equivalent amount: A from .50 to .75 and B from -.50 to -.25. This would mean that the AD compound should have an associative strength of .25 (.75 + -.50) and the BC compound should also have an associative strength of .25 (.50 + -.25). However, according to HeiDI one also needs to consider the fate of the backward associations during compound conditioning: between the US and A, and between the US and B. If it is assumed that α for all stimuli is .30, then V_{US-A} will be .30 by the end of the first stage of training, but V_{US-B} will be 0, because B has not been paired with the US. This will mean that while V_{US-A} will not change during pairings of AB with the US (the asymptote for V_{US-A} determined by $\alpha = .30$ will have been reached as a result of the first stage of training), V_{US-B} can increase (e.g., from 0 to .30). This will mean that during the test, $V_{COMB-BC}$ will be higher than $V_{COMB-AD}$. This analysis retains a pooled error term for all associations, but recognizes the fact – hitherto unacknowledged – that associations from the US to A and B will proceed independently of one another in conventional conditioning procedures (i.e., when there is only a single US).

Simulations confirm the accuracy of this analysis across a broad range of parameters[5], but in the interests of consistency the parameters were set in the way described in the previous paragraph: The α values of A, B, C and D were set at .30; by the end of stage 1, V_{A-US} and V_{C-US} were .50 (i.e., $\beta_{US} = .50$) while V_{B-US} and V_{D-US} were $-.50$; and V_{US-A} and V_{US-C} were .30, whereas V_{US-B} and V_{US-D} were 0. Having set these parameters, I then simulated how the CS-US associations involving A and B changed during conditioning with the AB compound (see Figure 10). Inspection of Panel A in Figure 10 confirms that V_{A-US} and V_{B-US} increased by equivalent amounts, and that

---

[5] These simulations are implemented in the HeiDI app as "Pooled error term".

while V_{US-A} remained the same, V_{US-B} increased to .30. Panel B in Figure 10 shows how the associative strengths of AD and BC change when the changes involving A and B were added to the existing strengths of D and B, respectively. Inspection of Panel B in Figure 10 confirms that the net V_{AD-US} and V_{BC-US} increase equivalently as a consequence of AB conditioning trials (the black symbols overlap with one another). However, while V_{US-BC} increases, V_{US-AD} does not. Panel C in Figure 10 shows that the $V_{COMB-BC}$ is greater than $V_{COMB-AD}$, reflecting the greater contribution of V_{US-BC} to BC than V_{US-AD} to AD. Finally, Panel D in Figure 10 reveals that the difference between BC and AD is evident in both R_{US} and R_{CS}; but in absolute terms is most evident for R_{CS}. This difference reflects the fact that with the parameters employed in the illustrative simulation, the combined alpha scores (α_{AD} and α_{BC} = .60) are greater than the V_{AD-US} and V_{BC-US} (both = .25). When other aspects of the simulation are held constant, but α for all stimuli was set at .10 (i.e., α_{AD} and α_{BC} = .20), the absolute difference between BC and AD is (approximately) equally evident for R_{US} and R_{CS}.

Rescorla and Wagner (1972) is unable to account for the fact that reinforcement of a compound (AB+) affects differently a previous excitor (A) or an inhibitor (B) without weakening assumptions governing the learning rules or adding more complexity to the model. HeiDI combines a pooled error term for the forward associations and separate error terms for backwards associations. This way, HeiDI is able to accommodate differential increases or decreases in associative strength for excitors and inhibitors, which models based on purely combined error term are unable to account for.

3.6.4 Combining stimuli with different associative histories

Rescorla and Wagner (1972) made the simplifying assumption that the associative strength of a compound stimulus (V_{AB-US}) is simply the sum of the individual associative strengths of A and B (i.e., $V_{A-US} + V_{B-US}$). Together with the assumption that V bears an ordinal relationship to performance, the model is constrained to predicted that there will be an ordinal relationship between performance to A, B and AB. For exam-

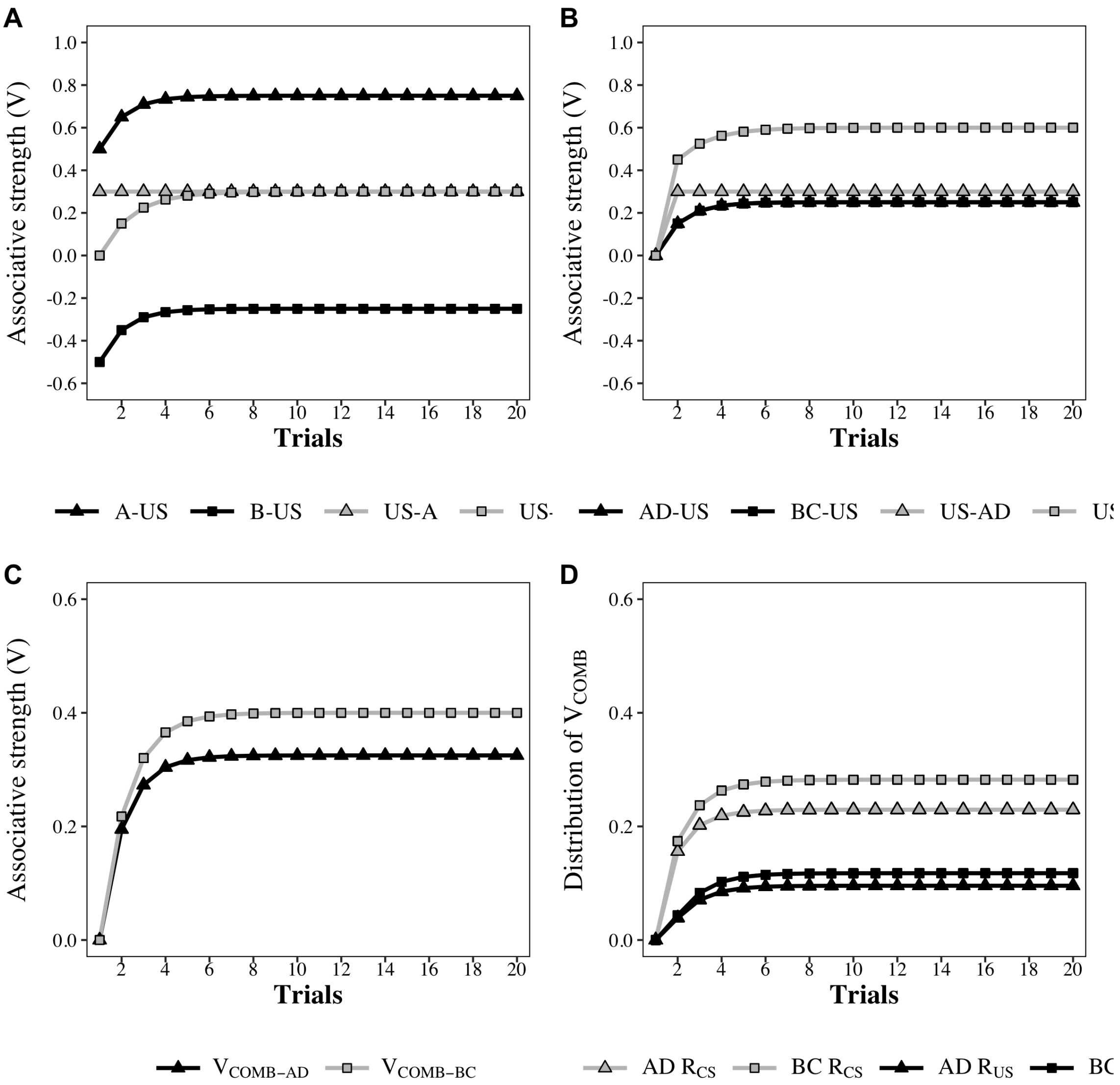


Figure 10: Simulations for HeiDI of the changes in associative strength when combining stimuli with different associative histories. A conditioned excitor (A) and an inhibitor (B) are conditioned in compound (AB), and tested with an inhibitor (D) and excitor (C) in compounds AD and BC. Panel A shows the output values for changes in associative strength of the components (A and B) of a stimulus compound (AB) that is paired with a US. Stimulus A (and C) begin compound conditioning with a V_{CS-US} of .50, and V_{US-CS} of .30; whereas B (and D) begin with a V_{CS-US} of -.50 and V_{US-CS} of 0. Panel B depicts the output values for the test compounds: V_{AD-US}, V_{US-AD}, V_{BC-US} and V_{US-BC}. Panel C shows the output values for the combination of the forward and backward associations for AD ($V_{COMB-AD}$) and BC ($V_{COMB-BC}$), while panel D illustrates how the differences in $V_{COMB-AD}$ and $V_{COMB-BC}$ are reflected in the output values for R_{CS} (CS-oriented behaviour) and R_{US} (US-oriented behaviour) during the test compounds AD and BC.

ple, if two stimuli with excitatory associative strength are combined then performance to the compound AB should exceed both A and B; whereas if one stimulus is excitatory

(A) and the other (B) is untrained (and without associative strength) then performance to AB should match A, and both should exceed B. Finally, if A is excitatory and B inhibitory then performance to AB should be less than A and greater than B, unless the excitatory value of A was less than or equal to the inhibitory value of B. While the predictions of HeiDI and the Rescorla-Wagner model mirror one another in some of these cases, they diverge in others.

Summation

The analysis begins with the first example, where two CSs (A and B) that have been separately paired with US are predicted to summate when they are combined at test. Equations 1 and 2 were used to generate the requisite individual Vs for stimulus A and B, and Equations 4 and 5 to determine performance. First, it was confirmed that summation was evident in both R_{CS} and R_{US} irrespective of whether the parameters were chosen to result in a bias towards the R_{CS} (e.g., α_A and $\alpha_B = .50$, and $\beta_{US} = .30$), or R_{US} (e.g., α_A and $\alpha_B = .50$, and $\beta_{US} = .70$). However, at an empirical level, summation is not an inevitable consequence of presenting two excitatory stimuli in compound. The circumstances under which summation does and does not occur have yet to be fully determined (Pearce, Aydin, & Redhead, 1997; Pearce, Redhead, & George, 2002), with theoretical analyses tending to focus on how the combination or configuration of stimuli changes the way in which they are processed (e.g., Brandon, Vogel, & Wagner, 2000; Pearce, 1994). I will not comment further here on the nature of such "configural" processes, but the aforementioned theoretical analyses make an important assumption: Separate conditioning trials with A and B results in them acquiring associative strength (relatively) independently of one another (see Brandon et al., 2000; Pearce, 1994). HeiDI does not make this assumption, and this fact has important implications for the conditions under which summation will be observed.

HeiDI assumes that associations form from the US to the CS. Unlike the development of A-US and B-US associations, which proceeds independently, the net US-A association will be weakened on a trial on which B is paired with the US and the net US-B association will be weakened on a trial when A is paired with the US. This fact does not in itself affect the prediction that summation will be observed (simulations included these reciprocal associations), however it does raise the possibility that another form of learning will occur that could constrain summation. To the extent that the A-US and US-B associations enable the B node to become active on a trial with A, and the B-US and US-A enable the A node to become active on a trial with B, there is the potential for inhibition to develop between A and B (see McLaren, Kaye, & Mackintosh, 1989; McLaren & Mackintosh, 2000).

$$\Delta V_{A-B} = \alpha_A(c * \alpha_B - \sum V_{Total-B}) \quad (7)$$

$$\Delta V_{B-A} = \alpha_B(c * \alpha_A - \sum V_{Total-A}) \quad (8)$$

First it is assumed that the change in the strength of the association between nodes A and B is governed by Equation 7, and the reciprocal B-A association is governed by Equation 8. These equations are formally equivalent to Equations 1 and 2. They provide a basis for the formation of associations between the elements of a compound (AB), allowing behaviour established to one stimulus (e.g., A) to transfer to the other (e.g., B). I will return to these CS-CS associations in the context of a potential analysis of features of blocking. The equations also provide the basis for the development of inhibition between A and B when both have been paired with the same US. According to Equations 7 and 8, net inhibition will develop between the A and B nodes, to the extent that the combined effect of the forward (e.g., A-US) and backward associations (e.g., US-B) provide an indirect basis for V$_{A-B}$ to be positive when B is absent. Thus, on a trial when A is presented, $\alpha_B = 0$ and the ability of A to activate B (i.e., V$_{A-B}$) will depend on multiplying the strengths of the A-US and US-B associations:

$1/c^*V_{A-US} * V_{US-B}$; and on a trial when B is presented, $\alpha_A = 0$ and V_{B-A} will depend on: $1/c^*V_{B-US} * V_{US-A}$. The development of this inhibition will mean that when A and B are presented together (e.g., for a summation test) their corresponding nodes will be less likely to become active than if either had been presented alone: Performance to an AB compound will be constrained by the development of inhibition between the nodes activated by A and B when both are followed by the same US. It is worth noting that such a constraint on summation would be less likely if A and B were to be followed by different reinforcers during conditioning; reinforcers with the same tendency to provoke conditioned responding but with distinct sensory properties (e.g., A-food and B-sucrose).

In keeping with the analysis outlined in the previous paragraph (i.e., that HeiDI predicts summation between two stimuli separately paired with the same US will be constrained by inhibition between them), Watt and Honey (1997) observed that a compound (AB) was more likely to provoke conditioned responding at test if its components had been separately paired with different appetitive reinforcers (food and sucrose) that support the same conditioned response, than if they had been paired with the same reinforcer (food or sucrose; or both food and sucrose, on different trials). It predicts this because there is less inhibition between A and B when they are followed by different appetitive USs than when they are followed by the same appetitive US: The shared sensory properties of the shared USs provide a basis for inhibition between A and B, and this means that they can suppress one another's activation. In general terms, differences in the development of inhibition between A and B engendered by different training procedures should affect the likelihood of summation being observed. The development of inhibition between A and B, when both are paired with the same outcome, has not been directly assessed in studies of summation or considered at a theoretical level (cf. Brandon, Vogel, & Wagner, 2000; Pearce, 1994). However, there is evidence that is consistent with this suggestion from studies of categorization (Aitken, Bennett, McLaren, & Mackintosh, 1996) and perceptual learning (e.g., Dwyer & Mackintosh, 2002; Mundy, Dwyer, & Honey, 2006).

External inhibition

When an associatively neutral stimulus (B) is presented with a stimulus with associative strength (A) the conditioned response to that stimulus is often disrupted; an effect known as external inhibition. For example, Pavlov (1927, p. 44) originally observed that the amount of conditioned responding to a CS (in his case the amount of salivation in dogs) was reduced when a stimulus with no associative properties was presented with the CS. This effect is not predicted by the Rescorla-Wagner model, and has been interpreted in terms of a decrease in attention to the CS (Mackintosh, 1974, p. 16). In a set of simulations in which the associative strength of V_{B-US} was set to zero and it was presented with a stimulus with excitatory associative strength (A) that possessed excitatory associative strength ($V_{A-US} > 0$), the presence of B increased R_{CS} and reduced R_{US}, for AB relative to A; which would increase CS-oriented behaviour relative to US-oriented behaviour. There is also some evidence that is consistent with this analysis from studies of a related effect, known as disinhibition, where conditioned responding (e.g., instrumental lever pressing for food) can be augmented by the presentation of a stimulus (e.g., a light or white noise Brimer, 1970; Brimer & Kamin, 1963). In fact, this effect appears to be most apparent when the level of lever pressing is low (e.g., at the onset of a fixed interval, Flanagan & Webb, 1964; Hinrichs, 1968). Unfortunately, none of these studies measured ongoing goal-tracking, which should be the mirror image of behaviour directed towards the lever.

Summation tests for conditioned inhibition

Finally, combining a stimulus with strong excitatory properties (A) and a stimulus with modest net inhibitory properties (B) will mean that V_{AB-US} will take a lower value than V_{A-US}. Equations 1 and 2 were used to generate the individual Vs for a reinforced stimulus (A) and a stimulus (B) that was non-reinforced in the presence of A. Equations

3-5 were used to determine performance. Whether the parameters were chosen to result in a bias towards R$_{CS}$ (e.g., α_A and α_B = .50, and β_{US} = .30), or R$_{US}$ (e.g., α_A and α_B = .50, and β_{US} = .70), combining A with B resulted in lower levels of both. The values for R$_{CS}$ and R$_{US}$ for the AB compound would remain positive (albeit lower than those for A alone) because V$_{COMB}$ will still be positive. However, if A had modest excitatory properties and B had strong inhibitory properties, then V$_{COMB}$ would be negative - and as a result R$_{CS}$ and R$_{US}$ would also be negative. Adopting Equation 6 would mean that r1 would be negative (unless either CS-r1 or US-r1 were also negative). In this case, an example of a positive r1 might be to approach the lever and a negative r1 to withdraw from the lever. If the negative values returned by Equations 1 and 2 were construed as involving the activation of a No US node (Konorski, 1967; Pearce & Hall, 1980), then the excitatory CS-No US association would result in R$_{CS}$ and R$_{NO-US}$ being positive, and R$_{NO-US}$ could then directly generate different forms of responding not supported by either the CS or US.

3.6.5 Blocking: Learning and performance

It was noted in the introduction that one of the key features of the Rescorla-Wagner model was its ability to explain how the associative strength of one stimulus within a compound affects the associative strength gained by another stimulus within the compound (e.g., blocking, Kamin, 1969). The formal similarity between Equation 1 and the Rescorla-Wagner model is clear, and like this model, Equation 1 generates these important effects on the development of the CS-US association. However, other features of HeiDI mean that blocking is not – as the Rescorla-Wagner model predicts – inevitable.

In extremis, Equations 1-3 in concert with Equations 4 and 5 provide an account of blocking that is clearly related to the Rescorla-Wagner model: If V$_{A-US}$ $\approx$ c*β_{US} at the end of a period of training where A has been paired with a US, then conditioning with a compound (AB) will result in little or no increase in the B-US association (i.e., V$_{B-US}$ $\approx$ 0). However, according to HeiDI, the reciprocal US-B association (V$_{US-B}$) will be unaffected by the fact that A has a reciprocal association with the US (V_{US-A}), because the c*α_A and c*α_B values of A and B provide a separate basis for the formation of these associations. The prediction that the US-B association is not blocked will ordinarily be without consequence because Equation 3 will return a V_{COMB} for B ≈ 0 (i.e., if $\mathrm{V}_{B-US} \approx 0$ then $\mathrm{V}_{B-US} + (1/\mathrm{c}^*\mathrm{V}_{B-US} * \mathrm{V}_{US-B}) \approx 0$). According to Equations 4 and 5, R_{CS} and $\mathrm{R}_{US} \approx 0$ because $\mathrm{V}_{COMB} \approx 0$. However, one clear implication of this analysis is that treatments that enable the US-B association to influence performance should reduce the blocking effect; and there is evidence that the performance to a blocked stimulus can be augmented under some conditions (for a review, see Miller et al., 1995).

Both HeiDI and the Rescorla-Wagner model predict that V_{B-US} (and V_{A-US}) will increase during the compound conditioning phase of a blocking procedure if $\mathrm{V}_{A-US} <$ c*β_{US}. However, unlike the Rescorla-Wagner model, HeiDI predicts that the pattern of performance when B is tested will reflect the values of c*α_B and V_{B-US}. Under these conditions, A might generate US-oriented behaviour (i.e., $\mathrm{V}_{A-US} >$ c*α_A), but the associative strength gained by B might be evident as CS-oriented behaviour (i.e., c*$\alpha_B > \mathrm{V}_{B-US}$). This simple observation has an important implication: A blocking effect might not be evident if the experimental assay was more sensitive to CS-oriented behaviour than to US-oriented behaviour. The fact that V_{B-US} is low will reduce V_{COMB} in Equation 3, but it will simultaneously increase the contribution to performance of the CS-oriented component (i.e., R_{CS}) and reduce the US-oriented component (i.e., R_{US}). While it would be tendentious to argue that failures to observe blocking (e.g., Maes et al., 2016) provide support for the analysis presented above – grounds for such failures abound – there can little doubt that blocking effects can be less complete than a simple rendering of the Rescorla-Wagner model would predict (for a recent review and analysis, see Urcelay, 2017).

However, perhaps the most serious challenge to the account of blocking offered by the Rescorla-Wagner model involves the conditions under which “unblocking” occurs. Conventional procedures for blocking involve two stages in which the reinforcer is the same: A->US and then AB->US. The fact that increasing the number of USs between stage 1 (e.g., A->US1) and stage 2 (AB->US1-US2) results in unblocking (i.e., learning about B) is perfectly consistent with the model, because this change introduces a positive discrepancy in the pooled error term (see Equations 0 and 1). The problematic result is the fact that reducing the reinforcer (i.e., A->US1-US2 and then AB->US1) can also result in responding to B (i.e., unblocking, Dickinson, Hall, & Mackintosh, 1976). Taken in isolation, Equations 0 and 1 predict that the reduction in the number of reinforcers should have resulted in B acquiring inhibitory properties (e.g., Cotton et al., 1982). 'Downshift unblocking', as it is known, has been taken as evidence that the reduction in the US prevents the reduction in attention to B that would ordinarily result from the fact that the US was predicted by A; and allows B to be learnt about (e.g., Mackintosh, 1975; Pearce & Hall, 1980). While there has been some progress in understanding the conditions under which downshift unblocking occurs (Holland, 1988) there is no consensus about its explanation. Many have simply adopted the view that downshift unblocking is prima facie evidence that attention can change as a result of experience (Pearce & Mackintosh, 2010). However, a speculative explanation for this effect can be derived from application of HeiDI, without appealing to changes in attention (see Figure 11).

The essence of the analysis is that the removal of the second shock allows a within-compound B-A association to form more effectively during downshift unblocking than during standard blocking; and this association allows B to "borrow" the associative properties of A. Consider a blocking procedure in which A is first followed by successive presentations of the same nominal US. Each US can be treated as having partially separate representations (US1 and US2). Under these conditions, A will become linked to both US1 and US2 until each link reaches the asymptote determined by c*β_{US1} and c*β_{US2}; and critically links will be strengthened between US1 and A, and US2 and A, until their combined associative strength = c*α_A. When AB is paired with US1 and US2, the associations between B and both US1 and US2 will be blocked; and the

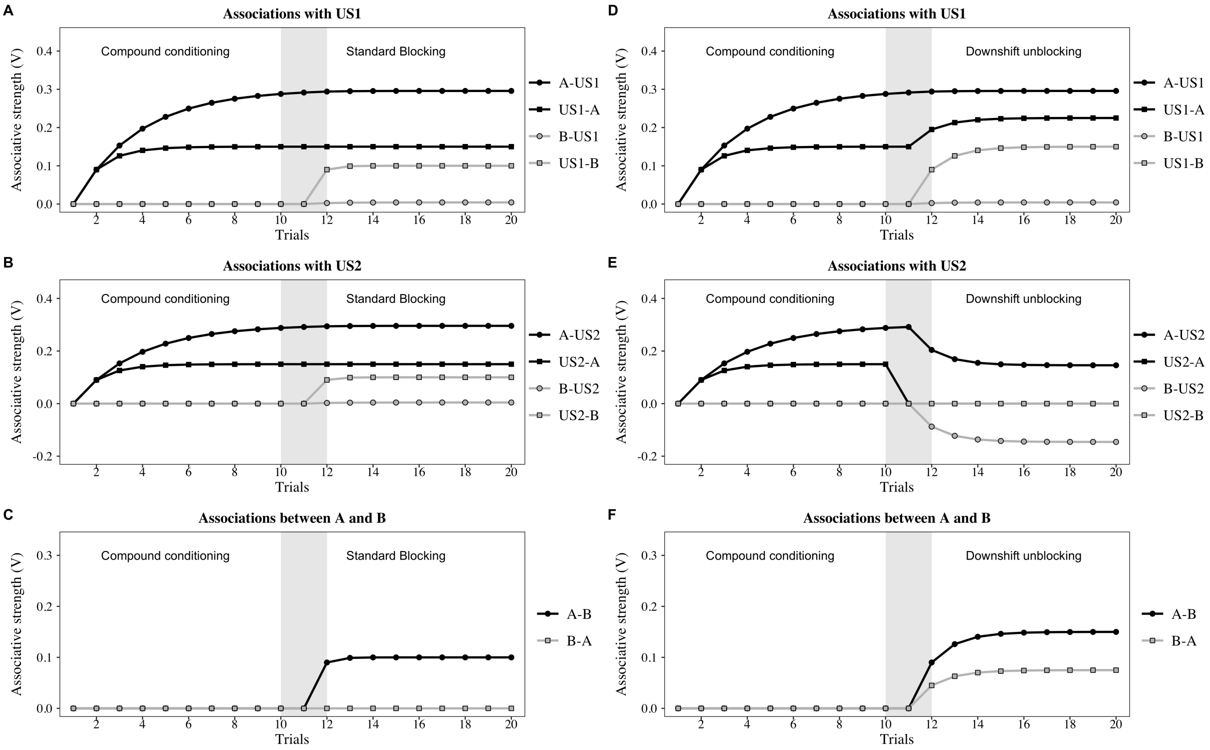


Figure 11: Simulations for HeiDI of the changes in associative strength during standard blocking (panels A-C) and downshift unblocking procedures (panels D-E). The parameters used were: $\alpha_A = \alpha_B = .30$, and $\beta_{US1} = \beta_{US2} = .30$. At the outset of compound conditioning, A-US1 and A-US2 were set to .30, and US1-A and US2-A were both set to .15. Panels A and D show the output values for the strengths of the A and B with US1 associations (A-US1, US1-A, B-US1, US1-B) and Panels B and E show the output values for the strengths of the A and B with US2 associations (A-US2, US2-A, B-US2, US2-B), returned by Equations 1 and 2 combined with Equations 7 and 8. Note that US2-A and US2-B is set to 0 in panel E to reflect the fact that the US2 is absent; but these associations will not change during unblocking. Panels C and F show the strength of the A-B and B-A associations. A key observation is that the B-A association gains strength during downshift unblocking (panel F), but not standard blocking (panel C).

combined effect of the US1-A and US2-A associations will mean that B will not be able to enter association with A. However, this will not be the case when US2 is omitted. If the change in the B-A association is assumed to be determined by α_B * (c*α_A − $\sum$ V$_{TOTAL-A}$), with $\sum$V$_{TOTAL-A}$ = V$_{US1-A}$ + V$_{US2-A}$ + V$_{B-A}$, then the removal of US2 will enable the strengthening of the B-A association (and further increases in the US1-A association). Under these conditions, downshift unblocking will occur to the extent that the influence of the B-A association in retrieving the associative properties of A (stronger following downshift unblocking than standard blocking) outweighs the fact that A-US2 is weaker and B-US2 is negative after downshift unblocking. This account is necessarily speculative, mirroring the fact that the understanding of the conditions under which downshift unblocking occurs remains incomplete (see Holland, 1988). However, it receives support from studies conducted by Rescorla and Colwill (1983), which showed that manipulations that should disrupt B-A associations also reduce the difference in performance to B between standard blocking and downshift unblocking[6].

The simulations[7] presented in Figure 11 for the compound conditioning stage are based – in the interests of simplicity – on the following parameters: $\alpha_A = \alpha_B =$ c*α_A $=$ c*α_B $=$ .30, and $\beta_{US1} = \beta_{US2} =$ c*β_{US1} $=$ c*β_{US2} $=$.30[8]. However, the critical difference in the B-A association during standard blocking and downshift unblocking is a general one. At the outset of simulated compound conditioning, for both standard blocking (panels A-C) and downshift unblocking (panels D-F), V_{A-US1} was set to .30 and V_{A-US2} was set to .30 to reflect the assumption that $\beta_{US1} = \beta_{US2} =$ c*β_{US1} $=$ c*β_{US2} $=$ .30. Critically, V_{US1-A} and V_{US2-A} were set at .15 for standard blocking, whereas for downshift unblocking V_{US1-A} was set at .15 and V_{US2-A} was set to 0 (to reflect the fact that US2 is absent). For the same reason, V_{US2-B} was also set to 0.

Panels A-C (standard blocking) and panels D-F (downshift unblocking) depict the values returned by the combination of Equations 1 and 2 with Equations 7 and 8. Panels A and D show associations between A and B with US1 (A-US1, US1-A, B-US1, US1-B) and Panels B and E show associations between A and B with US2 (A-US2, US2-A, B-US2, US2-B) and Panels C and F show associations between A and B (V_{A-B} and

---

[6] It is worth noting that within-compound (A-B) associations could also form during the experiments demonstrating unequal change in the associative strength of the elements of a compound (AB). However, in this case, there was evidence that these associations were not responsible for the effects that were observed (see Allman et al., 2004; Rescorla, 2000)

[7] These simulations are implemented in the HeiDI app as “Downshift unblocking”.

[8] The simulations that I report do not include associations between US1 and US2, because they would not influence the formation the excitatory B-A association upon which the analysis rests. Moreover, while the formation of US2-US1 and US1-US2 associations would tend to reduce respectively the A-US1 and A-US2 associations during conditioning with A, the absence of US2 during downshift unblocking would allow increases in both the A-US1 and B-US1 associations. Furthermore, the reductions in the net associative strength of the A-US2 and B-US2 associations produced by the absence of US2 would be less marked than those depicted in Figure 8D and 8E, because US1 would gain a proportion the overall net reduction. In summary, the inclusion of US1-US2 associations increases the likelihood that downshift unblocking would be observed.

V_{B-A}). Inspection of panels A-C confirms that during standard blocking associations involving A remained the same, and that associations from US1, US2 and A to B all increased by equivalent amounts. Critically, the B-A association did not develop, and this association can provide no basis upon which B could provoke conditioned responding; and the reciprocal US1-B and US2-B associations cannot – in isolation – contribute to performance. In contrast, during downshift unblocking because US2 is absent the US1-A and B-A associations can strengthen. This will mean both that V_{COMB-A} will be higher following downshift unblocking than standard blocking and that B will be able to access V_{COMB-A} through the B-A association. In order for this state of affairs to generate more performance to B it would need to outweigh the fact that the A-US2 and B-US2 are weaker or inhibitory after downshift unblocking than standard blocking. In Chapter 5, I will consider how the associative strength (V_{COMB-A}) borrowed by one stimulus (B) from another stimulus (A), with which it has an association (V_{B-A}), is manifest in performance. For now, it is sufficient to note that HeiDI provides one formal analysis of how within-compound associations might affect the outcome of blocking and unblocking procedures (cf. Urcelay, 2017).

3.6.6 Latent inhibition: An alternative associative analysis

Rescorla and Wagner (1972) recognized the fact that while their model provided a ready account for blocking, it did not address the fact that simple preexposure to a CS retards later excitatory and inhibitory conditioning (for a review, see Hall, 1991; Lubow, 1989). That is, the original model did not provide an account of latent inhibition (Lubow & Moore, 1959). But, why should repeated presentation of a to-be-conditioned stimulus affect the rate at which (excitatory and inhibitory) conditioned performance emerges to that stimulus? This observation in particular, as well as downshift unblocking, has prompted theorists to conclude that models of Pavlovian conditioning need to include another process that changes as a function of experience: attention, associability of CS processing (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Wagner, 1981).

However, a critical feature of latent inhibition, which provides a potential theoretical link with an associative analysis of blocking, is that latent inhibition is context specific. If preexposure to the CS occurs in one context (defined by the cues present in one experimental chamber) and conditioning takes place in another context, then latent inhibition is much reduced (e.g., Hall & Honey, 1989; Honey & Good, 1993). The general significance of this observation is that it suggests that – during the preexposure stage – animals encode where the stimulus has been presented; for example, by forming a context-CS association (cf. Wagner, 1981). This observation enables HeiDI to provide a simple analysis of latent inhibition: the blocking of the US-CS association by the context-CS association[9]

It was argued that during excitatory conditioning, performance is determined by both a CS-US association and a US-CS association, and that during inhibitory conditioning, performance could reflect the status of both a CS-No US and a No US-CS association (Konorski, 1967). While a context-CS association will not block the CS-US and CS-No US associations, it will block the development of the US-CS and No US-CS associations. Thus, the simple inclusion of a US-CS association (and No US-CS association) enables an account of latent inhibition that does not require a separate attentional or associability process (e.g., Mackintosh, 1975; Pearce & Hall, 1980) or changes in CS processing of the form envisaged by Wagner (1981).

In addition to this novel analysis of latent inhibition, the presence of a US-CS association means that the effective salience of CSs that are good predictors can be augmented (cf. Mackintosh, 1975). It has been shown that the α value of a stimulus affects the rate at which CS-oriented and US-oriented components of performance develop (see Figure 8). The US-CS association provides a natural way in which activation of the US might be reflected back to the CS and maintain its activation. Moreover, I have already

---

[9] It should be acknowledged that while the context specificity of latent inhibition is consistent with the view that context-CS associations provide a potential explanation for latent inhibition (and habituation), the fact that attempts to extinguish the context-CS association have often had no effect on latent inhibition is inconsistent with this account (see Baker & Mercier, 1982; Hall & Minor, 1984). However, the interpretation of failures of this kind is not straightforward (see Honey, Iordanova, & Good, 2010).

noted that when a CS is followed by a reduction in US magnitude (e.g., during extinction or partial reinforcement), CS-oriented responding increases relative to US-oriented responding, which could also affect the subsequent learning involving that CS. HeiDI thereby provides a simple analysis of phenomena that are routinely taken to indicate that the associability of stimuli (their α value) or their processing changes as a result of experience (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Pearce & Mackintosh, 2010; Wagner, 1981). According to this analysis, these phenomena are another product of the reciprocal associations that form between the CS and US, and between the components of stimulus compounds.

In this chapter I examined how HeiDI is able to accommodate some of phenomena that are beyond of the scope of Rescorla-Wagner Model, and how learning would translate into conditioned behaviour directed towards the CS or the US. In the following chapter, Chapter 4, I discuss and test new predictions derived from the model. Chapter 5, the General Discussion, will then return to the wider theoretical implications of the HeiDI model.

3.7 List of equations

$$\Delta V_{CS-US} = \alpha_{CS}(c * \beta_{US} - \sum V_{Total-US}) \ \ (1)$$

$$\Delta V_{US-CS} = \beta_{US}(c * \alpha_{CS} - \sum V_{Total-CS}) \ \ (2)$$

$$V_{COMB} = V_{CS-US} + (\frac{1}{c} * V_{CS-US} * V_{US-CS}) \ \ (3a)$$

$$V_{COMB-AB} = \sum V_{CS-US} + [\frac{1}{c} * \sum V_{AB-US} * (V_{US-A} + V_{US-B})] \ \ (3b)$$

$$R_{CS} = \frac{c * \alpha_{CS}}{c * \alpha_{CS} + |V_{CS-US}|} V_{COMB} \ \ (4)$$

$$R_{US} = \frac{|V_{CS-US}|}{c * \alpha_{CS} + |V_{CS-US}|} V_{COMB} \ \ (5)$$

$$r1 = (\frac{1}{c} * R_{CS} * CS_{r1}) + (\frac{1}{c} * R_{US} * US_{r1}) \ \ (6)$$

$$\Delta V_{A-B} = \alpha_{A}(c * \alpha_{B} - \sum V_{Total-B}) \ \ (7)$$

$$\Delta V_{B-A} = \alpha_{B}(c * \alpha_{A} - \sum V_{Total-A}) \ \ (8)$$

Chapter 4

Testing predictions of HeiDI

4.1 Individual differences in β_{US}

4.1.1 Introduction

As discussed in Chapter 3 (Section 3.2), recent evidence suggests that performance is affected by how the outcome is valued or perceived. Dwyer et al. (2018) showed that individual differences in the palatability of sucrose (during their experiments involving contrast effects) were positively correlated with the flavour preference learning. Moreover, rodents who showed a strong liking for sucrose (as measured by licking microstructure; see Dwyer, 2012) are more likely to be goal-trackers (when sucrose was the US) than those who exhibited a weaker liking for sucrose (Patitucci et al., 2016).

HeiDI assumes that the perceived salience of the CS (relating to α_{CS}) and US (relating to β_{US}; Equations 4 and 5) varies between animals, however the values of α_{CS} and β_{US} are fixed for a given animal. This variation in how the US is valued affects both learning (i.e., the asymptotic value of V_{CS-US} and the rate at which V_{US-CS} reaches asymptote, through Equations 1 and 2) and the distribution of V_{COMB} in performance (through V_{CS-US} in Equations 3-6).

Simulations of HeiDI are presented in Figure 13. The parameters during conditioning were chosen to result either in a higher US value (α_{CS} = .30 and β_{US} = .70; panels A and B) or a lower US value (α_{CS} = .30 and β_{US} = .50; panels C and D). Simulation for conditioning (left hand side of each panel; trials 1-6) and extinction (right-hand side of each panel; trials 7-12) for a reinforced (+) lever. Panels A and C depict the output values for $\mathrm{V_{CS\text{-}US}}$, $\mathrm{V_{US\text{-}CS}}$, $\mathrm{V_{COMB}}$, and panels B and D show the corresponding output values for $\mathrm{R_{CS}}$ and $\mathrm{R_{US}}$. It can be observed from Figure 13 that a higher US value will result in higher levels of goal-tracking or $\mathrm{R_{US}}$ (panel B), in contrast with lower US value

(Panel D). Note that, in contrast with the analysis presented for downshift unblocking in Chapter 3 (section 3.6.5), a two pellet reward is presented as a higher β_{US} value instead of two separate β_{US} values for US1 and US2.

To explore this prediction, rats were given training with one reinforced lever (e.g., left lever) and one non-reinforced lever (e.g., right lever). Half the animals were reinforced with one pellet reward and the other half with two pellets reward, after each presentation of the reinforced lever. HeiDI predicts that this manipulation will be shown in the rat's tendency to interact with the food-well, in other words, animals in the two pellet group will have higher overall goal-tracking levels than animals in the one pellet group.

4.1.2 Experiment 3

Method

Animals and apparatus

Twenty-three naive male Sprague Dawley rats were used (supplied by Charles River, UK). The animals were split into two groups: one or two pellets reward (11 rats in one pellet group, 12 rats in the two pellet group). The rats were housed as described in Experiment 1. Their mean ad libitum weight before the start of the experiment was 384g (range: 391-577g) and they were maintained at between 85 and 95 % of these weights by giving them restricted access to food at the end of each day. The rats had continuous access to water when they were in their cages. The experimental chambers were those used in Experiment 1. The design of the experiment is similar to Experiment 1, with the addition that in Experiment 3, half of the rats received two pellets as a reward, instead of one (see Table 3).

Procedure

The procedure used for Experiment 3 was the same as for Experiment 1 with the addition of another experimental group were the reinforcer was a two pellet reward instead of one.

Table 3: Design Of the Experiment 3 and 4

Classification	**Training**	**Reversal**	**Notation**
1pel	L_1 +	L_1 -	L_1 (+ ‖ -) 1pel
	L_2 -	L_2 +	L_2 (- ‖ +) 1pel
2pel	L_1 +	L_1 -	L_1 (+ ‖ -) 2pel
	L_2 -	L_2 +	L_2 (- ‖ +) 2pel

Note: 1pel and 2pel refer to the reward magnitude (one group of rats was reinforced with one pellet and the other group with two pellets); L_1 and L_2 refer to two levers (left and right, counterbalanced). During training, L_1 was reinforced ("+"; one food pellet in the 1pel group or two food pellets in the 2pel group) and L_2 was non-reinforced ("-"). During the reversal stage L_1 was non-reinforced and L_2 was reinforced (one or two pellets).

Data Analysis

Data analysis was conducted as described in Experiment 1. Rats were split into STs and GTs taking a median split (as described in Experiment 1) for each reward group (one and two pellets). For the 1 pellet, group bias scores above .67 were classified as GTs and below .58 as STs. For the two pellet group, bias scores above .23 were classified as GTs and bias scores below .19 as STs. The distribution of sign-tracking and goal-tracking behaviour is shown in Figure 12. Mixed ANOVAs were conducted separately for lever pressing and food-well entries with within-subjects factors block (training 6 levels: T1-T6; transition blocks 2 levels: T6-R1; reversal 6 levels: R1-R6;) and lever (2 levels: reinforced "+" vs non-reinforced "-") and between-subjects factor classification (2 levels: ST vs GT) as in Experiment 1, with the addition of reward as between-subjects factor (2 levels: one vs two pellets).

Results

Simulations of HeiDI are presented in Figure 13. The main results from Experiment 3 are shown in Figure 14. The analysis will begin with results from the training stage (left hand panels of Figure 14), before moving to the transition between training and reversal (identified by the grey section), and finally the reversal stage as a whole (right-hand

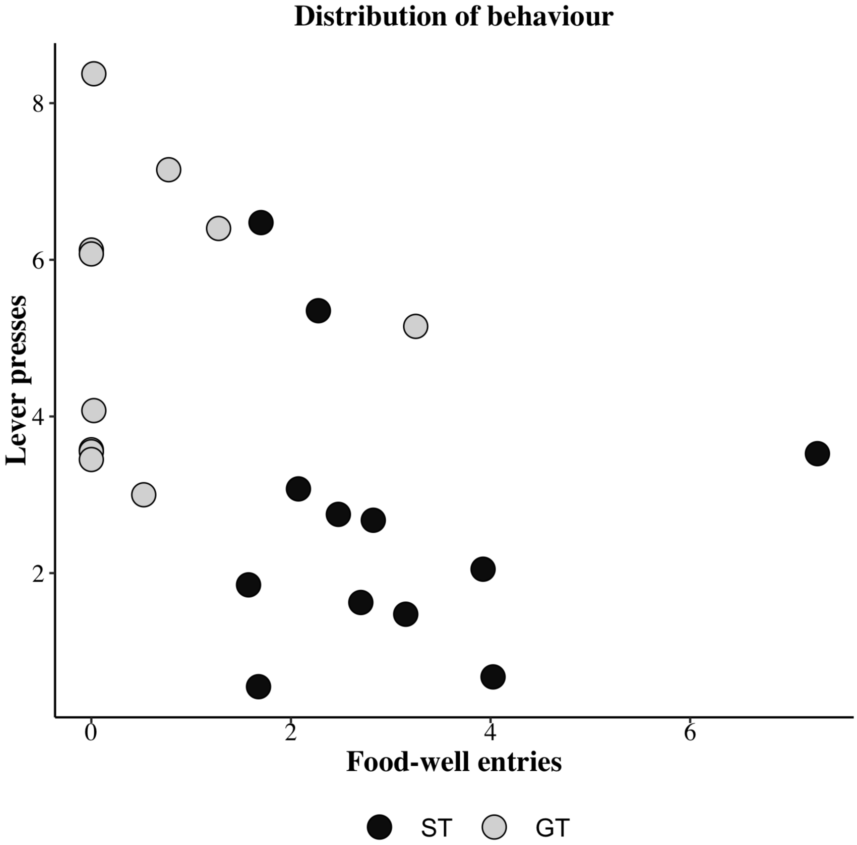


Figure 12: Distribution of sign-tracking (lever presses) and goal-tracking (food-well entries) behaviour for Experiment 3 per (10-s) trial during last block of training (T6) for L_1 (+ || -). The black symbols correspond to sign-trackers (ST) and the clear symbols to goal-trackers (GT).

pannels).

Training

Inspection of the results from the first stage of training (left-hand side of the upper and lower panels of Figure 14) suggest that as training progressed rats in both groups (ST and GT) showed more lever presses and food-well entries during the reinforced L_1 than the nonreinforced L_2. For lever presses, both STs and GTs show higher levels of response for two pellets reward, however the difference was numerically larger in STs than GTs. Food-well entries did not show a difference between one and two pellets reward for either STs or GTs. The description of the training results is supported by separate analyses of lever presses and food-well entries.

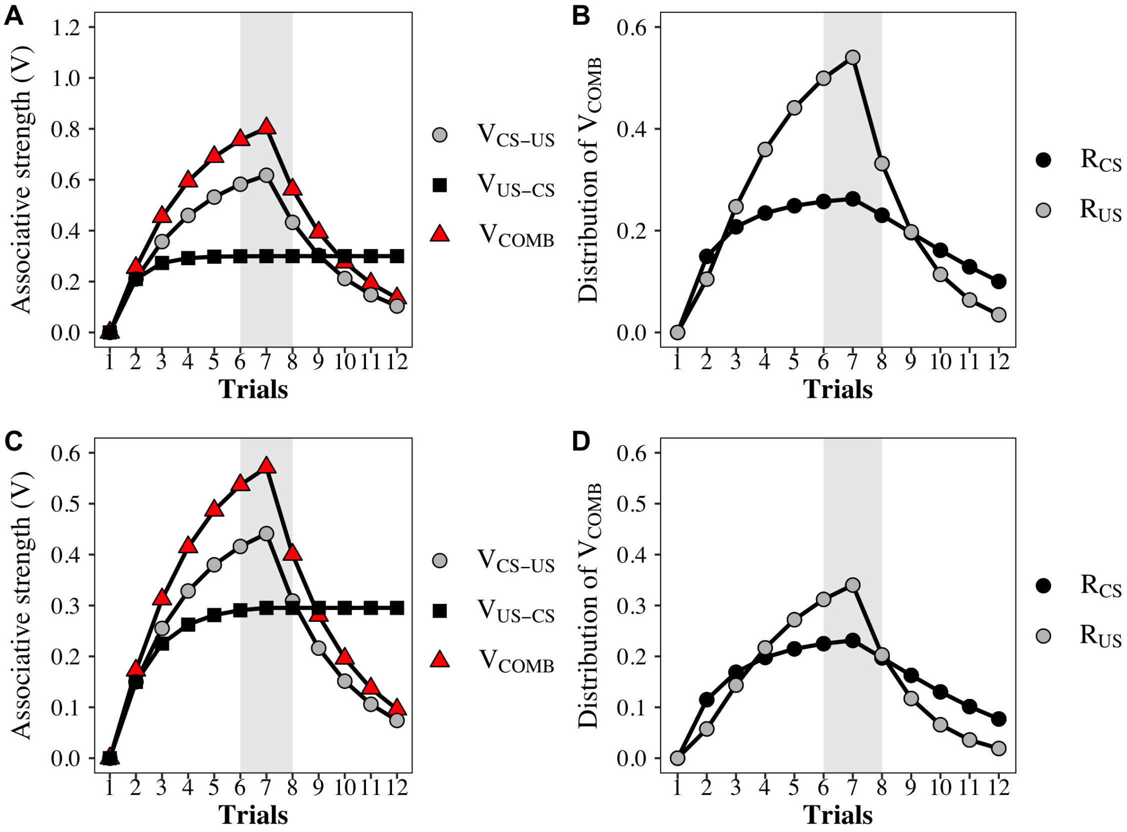


Figure 13: Simulation of HeiDI for individual differences in β_{US} on sign and goal-tracking: Simulation for conditioning (left-hand side of each panel; trials 1-6) and extinction (right-hand side of each panel; trials 7-12) for the reinforced (+) lever. Panels A and C depict the output values for $\mathrm{V_{CS\text{-}US}}$, $\mathrm{V_{US\text{-}CS}}$, $\mathrm{V_{COMB}}$, and panels B and D show the corresponding output values for $\mathrm{R_{CS}}$ and $\mathrm{R_{US}}$. The parameters during conditioning were chosen to result either in a higher US value (i.e., $\alpha_{CS} = .30$ and $\beta_{US} = .70$; panels A and B) or a lower US value (i.e., $\alpha_{CS} = .30$ and $\beta_{US} = .50$; panels C and D). During extinction of reinforcement, β_{US} was set to 0.

Lever presses

An ANOVA conducted for lever presses revealed main effects of reward, $F(1, 19) = 6.45$, $p = .020$, $\eta_p^2 = .25$, classification, $F(1, 19) = 40.41$, $p < .001$, $\eta_p^2 = .68$, block, $F(2.43, 46.20) = 8.96$, $p < .001$, $\eta_p^2 = .32$, and lever, $F(1, 19) = 65.83$, $p < .001$, $\eta_p^2 = .78$. There was an interaction between classification and block, $F(2.43, 46.20) = 4.73$, $p = .009$, $\eta_p^2 = .20$, and critically between reward and lever, $F(1, 19) = 6.75$, $p = .018$, $\eta_p^2 = .26$, as well as classification and lever, $F(1, 19) = 45.42$, $p < .001$, $\eta_p^2 = .71$ and between block and lever, $F(2.36, 44.78) = 14.22$, $p < .001$, $\eta_p^2 = .43$, but not between reward and classification, $F(1, 19) = 2.28$,

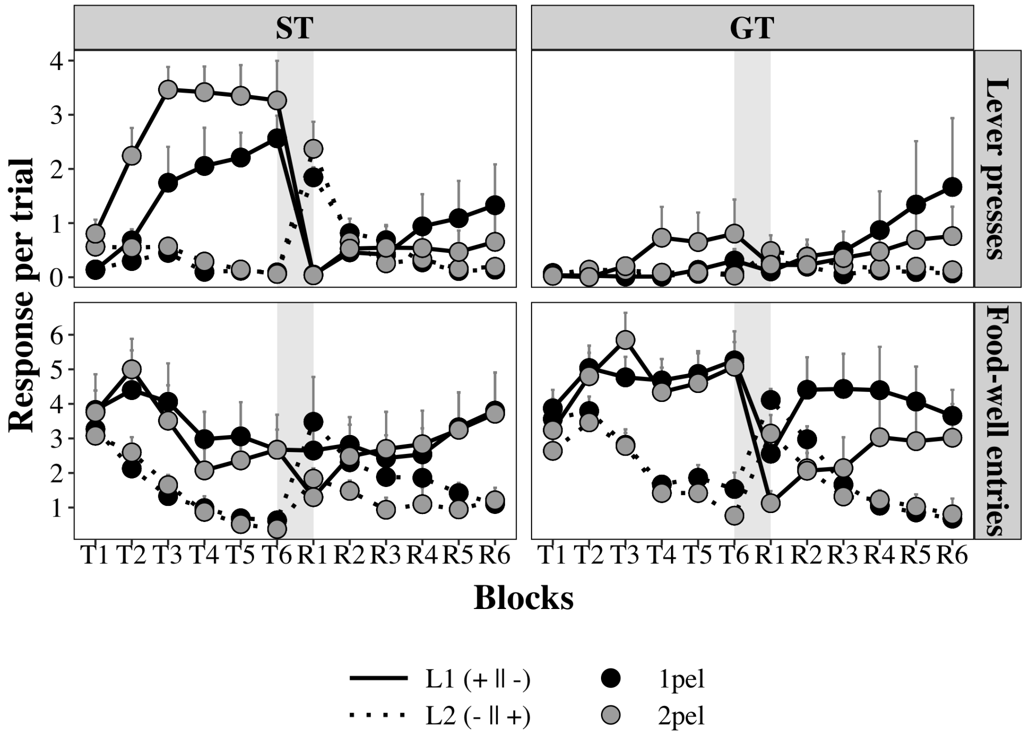


Figure 14: Results from Experiment 3: the effect of different values of β_{US} on sign-tracking and goal-tracking. Mean (+SEM) lever presses (upper panel) and food-well entries (lower panel) per (10-s) trial across the two stages: training (T1–T6) and reversal (R1–R6). During training, rats received presentations of one lever paired with food pellets (L_1 (+ || -)) and nonreinforced presentations of a second lever (L_2 (- || +)); rats were classified as sign-trackers (STs) and goal-trackers (GTs) on the basis of their behaviour during the final block of training (T6). They then received a reversal: L_1 non-reinforced and L_2 reinforced. The grey section indicates transition between initial training and the reversal of the contingencies.

p = .147, η_p^2 = .11 or between reward and block, $F(2.43, 46.20) = 0.52$, p = .631, η_p^2 = .03. There was a triple interaction between classification, block and lever, $F(2.36, 44.78) = 6.83$, p = .002, η_p^2 = .26, but no interaction between reward, classification and block, $F(2.43, 46.20) = 0.67$, p = .544, η_p^2 = .03, reward, classification and lever, $F(1, 19) = 2.18$, p = .156, η_p^2 = .10, reward, block and lever, $F(2.36, 44.78) = 0.71$, p = .520, η_p^2 = .04. There was no four-way interaction, $F(2.36, 44.78) = 0.70$, p = .522, η_p^2 = .04.

Food-well entries

An ANOVA conducted for food-well entries revealed main effects of classification, $F(1, 19) = 7.16$, $p = .015$, $\eta_p^2 = .27$, block, $F(2.31, 43.84) = 16.86$, $p < .001$, $\eta_p^2 = .47$, lever, $F(1, 19) = 43.07$, $p < .001$, $\eta_p^2 = .69$, but not of reward, $F(1, 19) = 0.26$, $p = .616$, $\eta_p^2 = .01$. There was an interaction between classification and block, $F(2.31, 43.84) = 4.28$, $p = .016$, $\eta_p^2 = .18$, and between block and lever, $F(2.62, 49.70) = 14.28$, $p < .001$, $\eta_p^2 = .43$, but not between reward and classification, $F(1, 19) = 0.03$, $p = .855$, $\eta_p^2 < .01$, reward and block, $F(2.31, 43.84) = 0.81$, $p = .468$, $\eta_p^2 = .04$, between reward and lever, $F(1, 19) = 0.00$, $p = .950$, $\eta_p^2 < .01$ or between classification and lever, $F(1, 19) = 0.68$, $p = .419$, $\eta_p^2 = .03$. There was a triple interaction between classification, block and lever, $F(2.62, 49.70) = 5.24$, $p = .005$, $\eta_p^2 = .22$, but not between reward, classification and block, $F(2.31, 43.84) = 0.73$, $p = .505$, $\eta_p^2 = .04$, reward, classification and lever, $F(1, 19) = 0.25$, $p = .625$, $\eta_p^2 = .01$, or between reward, block and lever, $F(2.62, 49.70) = 0.43$, $p = .708$, $\eta_p^2 = .02$. There was no four-way interaction between reward, classification, block and lever, $F(2.62, 49.70) = 0.60$, $p = .598$, $\eta_p^2 = .03$.

Transition blocks

Inspection of the results for the transition blocks (grey section in Figure 14) for lever presses suggests no difference between one pellet and two pellets reward for either STs or GTs (upper panels). The same is true for food-well entries, no difference is observed between the one and two pellets in either STs or GTs. The reversal in contingencies affected more the levels of food-well entries (lower panels) than lever-presses (upper panels) in both groups (1pel and 2pel). This description of the training results is supported by separate analyses of lever presses and food-well entries. Unlike in Experiment 1, the larger and faster effect of reversing contingencies for food-well entries than lever-presses was a difference of degree, as reversal did impact on lever press to some degree during the transition blocks.

Lever presses

An ANOVA conducted for lever presses for the transition blocks revealed main effects of classification, $F(1,\ 19) = 19.38$, $p < .001$, $\eta_p^2 = .50$, block, $F(1,\ 19) = 4.57$, $p = .046$, $\eta_p^2 = .19$, lever, $F(1,\ 19) = 7.73$, $p = .012$, $\eta_p^2 = .29$, but not of reward, $F(1,\ 19) = 1.15$, $p = .297$, $\eta_p^2 = .06$. There was an interaction between classification and lever, $F(1,\ 19) = 1.73$, $p = .204$, $\eta_p^2 = .08$, and critically between block and lever, $F(1,\ 19) = 43.37$, $p < .001$, $\eta_p^2 = .70$, but not between reward and classification, $F(1,\ 19) = 0.06$, $p = .813$, $\eta_p^2 < .01$, critically reward and block, $F(1,\ 19) = 0.15$, $p = .702$, $\eta_p^2 < .01$, classification and block, $F(1,\ 19) = 3.79$, $p = .066$, $\eta_p^2 = .17$, or reward and lever, $F(1,\ 19) = 0.80$, $p = .383$, $\eta_p^2 = .04$. There were no triple interaction, reward, classification and block, $F(1,\ 19) = 0.00$, $p = .996$, $\eta_p^2 < .01$, reward, classification and lever, $F(1,\ 19) = 0.16$, $p = .697$, $\eta_p^2 < .01$, reward, block and lever, $F(1,\ 19) = 1.18$, $p = .292$, $\eta_p^2 = .06$, classification, block and lever, $F(1,\ 19) = 23.69$, $p < .001$, $\eta_p^2 = .55$. There was no four-way interaction, reward, classification, block and lever, $F(1,\ 19) = 0.16$, $p = .695$, $\eta_p^2 < .01$.

Food-well entries

An ANOVA conducted for food-well entries revealed main effects of classification, $F(1,\ 19) = 5.67$, $p = .028$, $\eta_p^2 = .23$, and lever, $F(1,\ 19) = 45.03$, $p < .001$, $\eta_p^2 = .70$, but not of reward, $F(1,\ 19) = 3.92$, $p = .062$, $\eta_p^2 = .17$ or block, $F(1,\ 19) = 0.49$, $p = .494$, $\eta_p^2 = .02$. Critically, there were interactions between reward and block, $F(1,\ 19) = 5.82$, $p = .026$, $\eta_p^2 = .23$, classification and block, $F(1,\ 19) = 7.05$, $p = .016$, $\eta_p^2 = .27$, classification and lever, $F(1,\ 19) = 1.81$, $p = .195$, $\eta_p^2 = .09$, and critically between block and lever, $F(1,\ 19) = 59.50$, $p < .001$, $\eta_p^2 = .76$, but not between reward and classification, $F(1,\ 19) = 0.00$, $p = .974$, $\eta_p^2 < .01$, or between reward and lever, $F(1,\ 19) = 0.37$, $p = .552$, $\eta_p^2 = .02$. There was no triple interaction, reward, classification and block, $F(1,\ 19) = 0.55$, $p = .467$, $\eta_p^2 = .03$, reward, classification and lever, $F(1,\ 19) = 0.14$, $p = .713$, $\eta_p^2 < .01$, reward, block and lever, $F(1,\ 19) = 0.20$, $p = .660$, $\eta_p^2 = .01$, classification, block and lever, $F(1,\ 19) = 6.91$, $p = .017$, $\eta_p^2 = .27$ or a four-way interaction between reward, classification, block and lever, $F(1,\ 19) = 0.24$, $p = .631$, $\eta_p^2 = .01$. In contrast to Experiment 1 and Experiment 2 there was an interaction between block and lever for lever presses ($F(1,\ 19) = 43.37$, $p < .001$, $\eta_p^2 = .70$), however the effect size for food-well entries is considerably larger ($F(1,\ 19) = 59.50$, $p < .001$, $\eta_p^2 = .76$).

Reversal

Across the blocks of reversal training (right-hand side of the upper and lower panels of Figure 14), no differences were observed between the one and two pellet groups in neither lever presses nor food-well entries. Both STs and GTs increased responding for L_2 and decreased for L_1 for both lever presses and food-well entries.

Lever presses An ANOVA conducted for lever presses for the reversal stage revealed no main effects of reward, $F(1,\ 19) = 0.21$, $p = .653$, $\eta_p^2 = .01$, classification, $F(1,\ 19) = 0.62$, $p = .440$, $\eta_p^2 = .03$, block, $F(1.52,\ 28.84) = 1.58$, $p = .225$, $\eta_p^2 = .08$, or lever, $F(1,\ 19) = 0.68$, $p = .420$, $\eta_p^2 = .03$. There was an interaction between classification and block, $F(1.52,\ 28.84) = 3.45$, $p = .057$, $\eta_p^2 = .15$ and between block and lever, $F(1.29,\ 24.60) = 16.44$, $p < .001$, $\eta_p^2 = .46$, but not between reward and classification, $F(1,\ 19) = 0.00$, $p = .997$, $\eta_p^2 < .01$, reward and block, $F(1.52,\ 28.84) = 1.30$, $p = .281$, $\eta_p^2 = .06$, between reward and lever, $F(1,\ 19) = 0.69$, $p = .415$, $\eta_p^2 = .04$ or between classification and lever, $F(1,\ 19) = 1.25$, $p = .278$, $\eta_p^2 = .06$. There was a triple interaction between classification, block and lever, $F(1.29,\ 24.60) = 3.10$, $p = .082$, $\eta_p^2 = .14$, but not between reward, classification and block, $F(1.52,\ 28.84) = 0.09$, $p = .867$, $\eta_p^2 < .01$, reward, classification and lever, $F(1,\ 19) = 0.05$, $p = .822$, $\eta_p^2 < .01$ or between reward, block and lever, $F(1.29,\ 24.60) = 1.01$, $p = .347$, $\eta_p^2 = .05$. There was no four-way interaction between reward, classification, block and lever, $F(1.29,\ 24.60) = 0.38$, $p = .595$, $\eta_p^2 = .02$.

Food-well entries

An ANOVA conducted for food-well entries for the reversal stage revealed a main effect of lever, $F(1,\ 19) = 12.41$, $p = .002$, $\eta_p^2 = .40$, but no effect of reward, $F(1,\ 19) = 1.75$, $p = .202$, $\eta_p^2 = .08$, classification, $F(1,\ 19) = 0.18$, $p = .674$, $\eta_p^2 < .01$ or block, $F(2.51,\ 47.72) = 1.20$, $p = .318$, $\eta_p^2 = .06$. Critically, there was an interaction between block and lever, $F(2.08,\ 39.60) = 26.60$, $p < .001$, $\eta_p^2 = .58$, but no interactions between reward and classification, $F(1, 19) = 0.16$, $p = .694$, $\eta_p^2 < .01$, reward and block, $F(2.51, 47.72) = 2.27$, $p = .102$, $\eta_p^2 = .11$, classification and block, $F(2.51, 47.72) = 1.48$, $p = .237$, $\eta_p^2 = .07$, between reward and lever, $F(1, 19) = 0.23$, $p = .635$, $\eta_p^2 = .01$, or between classification and lever, $F(1, 19) = 0.08$, $p = .775$, $\eta_p^2 < .01$. There were no triple interactions between, reward, classification and block, $F(2.51, 47.72) = 0.64$, $p = .568$, $\eta_p^2 = .03$, reward, classification and lever, $F(1, 19) = 1.49$, $p = .237$, $\eta_p^2 = .07$, reward, block and lever, $F(2.08, 39.60) = 0.09$, $p = .919$, $\eta_p^2 < .01$, classification, block and lever, $F(2.08, 39.60) = 2.31$, $p = .111$, $\eta_p^2 = .11$ or four-way interaction between reward, classification, block and lever, $F(2.08, 39.60) = 0.87$, $p = .432$, $\eta_p^2 = .04$.

Discussion

The purpose of this experiment was to explore how reinforcers of different values affect conditioned behaviour. HeiDI predicts higher levels of US-oriented behaviour with USs of higher value. In contrast to this prediction, the results showed that a US with a higher β_{US} value (two pellets) produced more sign-tracking compared to a US with a lower β_{US} value (one pellet), and no differences were observed in goal-tracking. This may be the result of manipulating reward value through increasing the number of rewards rather than their size or value more directly. Regardless, the results from Experiment 3, provide further evidence for the results obtained in Experiment 1 and Experiment 2, replicating the finding that lever pressing behaviour is more resistant to contingency changes than food-well behaviour.

4.1.3 Experiment 4

While Experiment 3 provided results that were not consistent with the predictions of the HeiDI model regarding the effects of increasing the value of the US. Experiment 4 repeated the same experimental design.

Method

Animals and apparatus

Twenty-four Sprague Dawlley female rats were used (supplied by Charles River, UK). The rats were housed as described in Experiment 1. Their mean ad libitum weight before the start of the experiment was 238.17 (range: 205 - 309 g) and they were maintained at between 85 and 95 % of these weights by giving them restricted access to food at the end of each day. The rats had continuous access to water when they were in their cages. The apparatus is the same as described in Experiment 1.

Procedure

The same procedure was used as in Experiment 3.

Data analysis

The data was analysed in the same way as in Experiment 3. Rats were split into STs and GTs taking a median split (as described in Experiment 1) for each reward group (one and two pellets). The distribution of lever-presses and food-well entries behaviour is shown in Figure 16 where it is clear that very few rats pressed the lever. For the one pellet group bias scores below .99 were classified as ST and all scores of 1 were classified as GT. The median split in this experiment meant that some animals were classified as STs when they were barely pressing the lever. For the two pellet group bias scores below .66 were classified as STs and scores above .86 as GTs. This implies a distribution skewed towards goal-tracking but less so than for the one pellet group.

Results

Training

Inspection of the results from the first stage of training (left-hand side of the upper and lower panels of Figure 15) suggests that only the animals in ST group are pressing

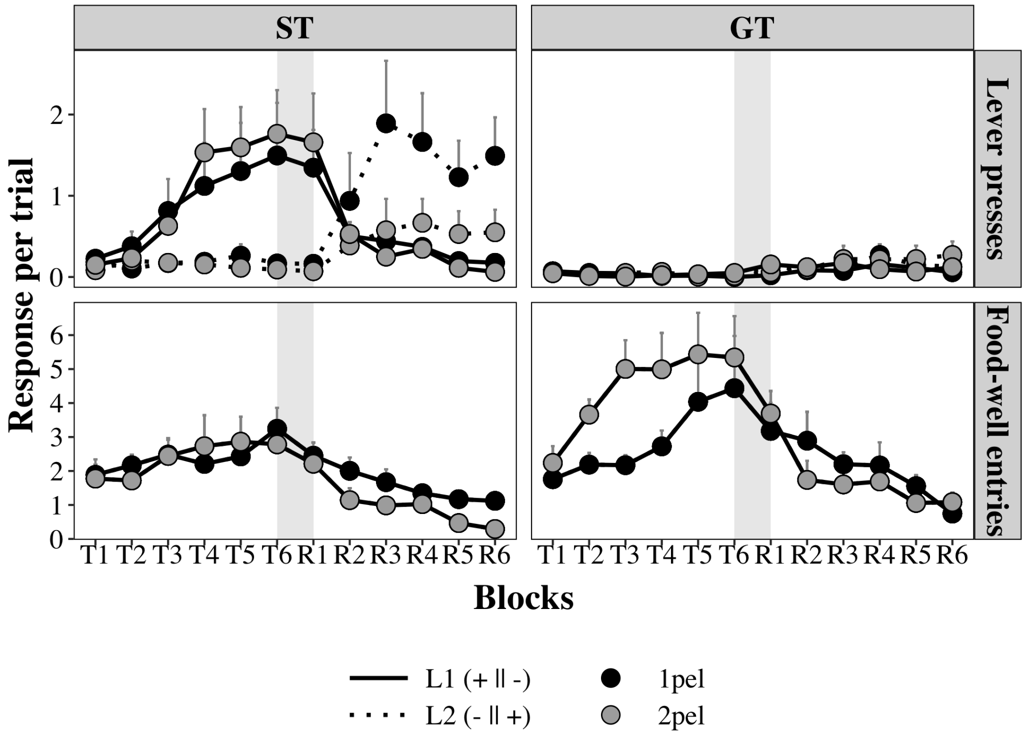


Figure 15: Results from Experiment 4 (replication of Experiment 3 in female rats): the effect of different values of β_{US} on sign-tracking and goal-tracking. Mean (+SEM) lever presses (upper panel) and food-well entries (lower panel) per (10-s) trial across the two stages: training (T1–T6) and reversal (R1–R6). During training, rats received presentations of one lever paired with food pellets (L_1 (+ || -)) and nonreinforced presentations of a second lever (L_2 (- || +)); rats were classified as sign-trackers (STs) and goal-trackers (GTs) on the basis of their behaviour during the final block of training (T6). They then received a reversal: L_1 non-reinforced and L_2 reinforced. The grey section indicates transition between initial training and the reversal of the contingencies.

the lever, with no differences between one or two pellets as a reinforcer. In contrast with previous experiments, the GT animals show no lever-oriented behaviour. The distribution of lever-presses and food-well entries behaviour shown in Figure 16 suggests that compared to other cohorts, relatively few animals show lever press behaviour and the GT group display especially little lever press behaviour. The implications of this issue are addressed in the discussion section. Differences in the magnitude of reward were not observed in either lever presses or food-well entries. This description of the training results is supported by separate analyses of lever presses and food-well entries.

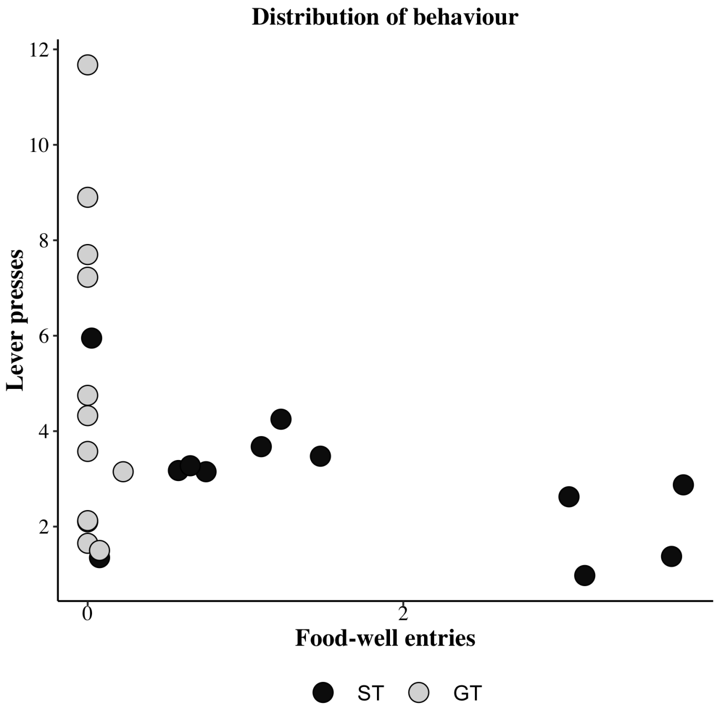


Figure 16: Distribution of sign-tracking (lever presses) and goal-tracking (food-well entries) behaviour for Experiment 4 per (10-s) trial during last block of training (T6) for L_1 (+ || -). The black symbols correspond to sign-trackers (ST) and the clear symbols to goal-trackers (GT).

Lever presses

An ANOVA conducted for lever presses for the training stage revealed main effects of classification, $F(1,\ 44) = 0.02$, $p = .888$, $\eta_p^2 < .01$, block, $F(2.46,\ 108.12) = 0.55$, $p = .616$, $\eta_p^2 = .01$, lever, $F(1,\ 44) = 0.14$, $p = .711$, $\eta_p^2 < .01$, but critically no effect of reward, $F(1,\ 44) = 2.16$, $p = .148$, $\eta_p^2 = .05$. There was a two-way interaction between classification and block, $F(2.46,\ 108.12) = 0.31$, $p = .777$, $\eta_p^2 < .01$, classification and lever, $F(1,\ 44) = 0.16$, $p = .691$, $\eta_p^2 < .01$, as well as between block and lever, $F(2.58,\ 113.36) = 0.46$, $p = .683$, $\eta_p^2 = .01$, but not between reward and classification, $F(1,\ 44) = 2.46$, $p = .124$, $\eta_p^2 = .05$, and critically not between reward and block, $F(2.46,\ 108.12) = 3.26$, $p = .033$, $\eta_p^2 = .07$ or reward and lever, $F(1,\ 44) = 2.86$, $p = .098$, $\eta_p^2 = .06$. There was a three-way interaction between classification, block and lever, $F(2.58,\ 113.36) = 0.38$, $p = .740$, $\eta_p^2 < .01$, but not between reward, classification and block, $F(2.46,\ 108.12) = 2.57$, $p = .070$, $\eta_p^2 = .06$, reward, classification and lever, $F(1,\ 44) = 3.08$, $p = .086$, $\eta_p^2 = .07$ or between reward, block and lever, $F(2.58,\ 113.36) = 3.46$, $p = .024$, $\eta_p^2 = .07$. There was no four-way interaction between reward, classification, block and lever, $F(2.58,\ 113.36) = 3.09$, $p = .037$, $\eta_p^2 = .07$.

Food-well entries

Parallel analysis conducted for food-well entries for the training stage revealed main effects of classification, $F(1,\ 44) = 0.01$, $p = .913$, $\eta_p^2 < .01$ and lever, $F(1,\ 44) = 4.69$, $p = .036$, $\eta_p^2 = .10$, but no effect of reward, $F(1,\ 44) = 0.40$, $p = .528$, $\eta_p^2 < .01$ or block, $F(2.47,\ 108.72) = 1.64$, $p = .192$, $\eta_p^2 = .04$. Critically, there was an interaction between block and lever, $F(2.58,\ 113.43) = 2.07$, $p = .117$, $\eta_p^2 = .04$, but no other two-way interactions between reward and classification, $F(1,\ 44) = 1.41$, $p = .241$, $\eta_p^2 = .03$, critically reward and block, $F(2.47,\ 108.72) = 0.55$, $p = .615$, $\eta_p^2 = .01$, classification and block, $F(2.47,\ 108.72) = 1.33$, $p = .269$, $\eta_p^2 = .03$, reward and lever, $F(1,\ 44) = 2.78$, $p = .102$, $\eta_p^2 = .06$, or classification and lever, $F(1,\ 44) = 0.82$, $p = .370$, $\eta_p^2 = .02$. There were no three-way interactions between reward, classification and block, $F(2.47,\ 108.72) = 0.93$, $p = .414$, $\eta_p^2 = .02$, reward, classification and lever, $F(1,\ 44) = 3.36$, $p = .073$, $\eta_p^2 = .07$, reward, block and lever, $F(2.58,\ 113.43) = 1.83$, $p = .154$, $\eta_p^2 = .04$, classification, block and lever, $F(2.58,\ 113.43) = 0.60$, $p = .593$, $\eta_p^2 = .01$, or a four-way interaction between reward, classification, block and lever, $F(2.58,\ 113.43) = 1.78$, $p = .162$, $\eta_p^2 = .04$.

Transition blocks

Inspection of the results for the transition blocks (grey section of Figure 15) show that lever presses (when they are at material levels) are less affected by the change in contingencies than food-well entries.

Lever presses

ANOVA conducted for the transition blocks for lever presses revealed main effects of classification, $F(1,\ 44) = 0.07$, $p = .792$, $\eta_p^2 < .01$, and lever, $F(1,\ 44) = 0.01$, $p = .906$, $\eta_p^2 < .01$, but not of reward, $F(1,\ 44) = 3.51$, $p = .068$, $\eta_p^2 = .07$, or block, $F(1,\ 44) = 0.20$, $p = .654$, $\eta_p^2 < .01$. There was a two-way interaction between classification and lever, $F(1,\ 44) = 0.01$, $p = .910$, $\eta_p^2 < .01$, but not between reward and classification, $F(1,\ 44) = 1.75$, $p = .193$, $\eta_p^2 = .04$, reward and block, $F(1,\ 44) = 0.71$, $p = .405$, $\eta_p^2 = .02$, classification and block, $F(1,\ 44) = 0.04$, $p = .852$, $\eta_p^2 < .01$, reward and lever, $F(1,\ 44) = 2.63$, $p = .112$, $\eta_p^2 = .06$ or critically between block and lever, $F(1,\ 44) = 0.10$, $p = .752$, $\eta_p^2 < .01$. There were no three-way between interactions reward, classification and block, $F(1,\ 44) = 0.27$, $p = .604$, $\eta_p^2 < .01$, reward, classification and lever, $F(1,\ 44) = 3.20$, $p = .080$, $\eta_p^2 = .07$, reward, block and lever, $F(1,\ 44) = 0.81$, $p = .374$, $\eta_p^2 = .02$, classification, block and lever, $F(1,\ 44) = 0.00$, $p = .955$, $\eta_p^2 < .01$ or a four-way interaction reward, classification, block and lever, $F(1,\ 44) = 0.04$, $p = .842$, $\eta_p^2 < .01$.

Food-well entries

ANOVA conducted for the transition blocks for food-well entries revealed a main effect of lever, $F(1,\ 44) = 5.35$, $p = .025$, $\eta_p^2 = .11$, but no main effects of reward, $F(1,\ 44) = 0.13$, $p = .720$, $\eta_p^2 < .01$, classification, $F(1,\ 44) = 0.00$, $p = .968$, $\eta_p^2 < .01$ or block, $F(1,\ 44) = 0.35$, $p = .555$, $\eta_p^2 < .01$. Critically, there was an interaction between block and lever, $F(1,\ 44) = 2.42$, $p = .127$, $\eta_p^2 = .05$, but not between reward and classification, $F(1,\ 44) = 1.33$, $p = .255$, $\eta_p^2 = .03$, between reward and block, $F(1,\ 44) = 0.24$, $p = .629$, $\eta_p^2 < .01$, classification and block, $F(1,\ 44) = 0.79$, $p = .380$, $\eta_p^2 = .02$, reward and lever, $F(1,\ 44) = 1.35$, $p = .252$, $\eta_p^2 = .03$, or classification and lever, $F(1,\ 44) = 0.01$, $p = .938$, $\eta_p^2 < .01$. There were no three-way interactions between reward, classification and block, $F(1,\ 44) = 0.00$, $p = .997$, $\eta_p^2 < .01$, reward, classification and lever, $F(1,\ 44) = 1.07$, $p = .307$, $\eta_p^2 = .02$, reward, block and lever, $F(1,\ 44) = 1.50$, $p = .228$, $\eta_p^2 = .03$, or classification, block and lever, $F(1,\ 44) = 5.13$, $p = .028$, $\eta_p^2 = .10$. There was no four-way interaction between reward, classification, block and lever, $F(1,\ 44) = 6.71$, $p = .013$, $\eta_p^2 = .13$.

Reversal

Across the blocks of reversal training (right-hand side of the upper panels of Figure 14), lever presses increased for L_2 and decreased during L_1 in STs, however the increase was larger for the 1pel group. Lever-press behaviour remained negligible in the GT groups. Food-well behaviour (right-hand side of the lower panels of Figure 14) increased for L_2 and decreased during L_1. The description of these results is supported by the following analysis.

Lever presses

ANOVA conducted for lever presses for the reversal stage revealed main effects of classification, $F(1,\ 44) = 1.23$, $p = .273$, $\eta_p^2 = .03$ and lever, $F(1,\ 44) = 2.23$, $p = .143$, $\eta_p^2 = .05$, but no effect of reward, $F(1,\ 44) = 0.07$, $p = .795$, $\eta_p^2 < .01$, or block, $F(3.48,\ 153.07) = 3.29$, $p = .017$, $\eta_p^2 = .07$. There was a two-way interaction between reward and lever, $F(1,\ 44) = 0.43$, $p = .515$, $\eta_p^2 < .01$ and between block and lever, $F(1.66,\ 72.91) = 0.81$, $p = .427$, $\eta_p^2 = .02$, but not between reward and classification, $F(1,\ 44) = 0.08$, $p = .779$, $\eta_p^2 < .01$, not between reward and block, $F(3.48,\ 153.07) = 4.53$, $p = .003$, $\eta_p^2 = .09$, classification and block, $F(3.48,\ 153.07) = 1.28$, $p = .282$, $\eta_p^2 = .03$, or between classification and lever, $F(1,\ 44) = 1.73$, $p = .195$, $\eta_p^2 = .04$. There was a triple interaction between reward, classification and lever, $F(1,\ 44) = 1.49$, $p = .228$, $\eta_p^2 = .03$ and between classification, block and lever, $F(1.66,\ 72.91) = 0.49$, $p = .582$, $\eta_p^2 = .01$, but not between reward, classification and block, $F(3.48,\ 153.07) = 2.95$, $p = .028$, $\eta_p^2 = .06$, or reward, block and lever, $F(1.66,\ 72.91) = 0.32$, $p = .688$, $\eta_p^2 < .01$. There was not a four-way interaction between reward, classification, block and lever, $F(1.66,\ 72.91) = 0.30$, $p = .697$, $\eta_p^2 < .01$.

Food-well entries

ANOVA conducted for food-well entries for the reversal stage revealed main effects of block, $F(3.06,\ 134.81) = 0.62$, $p = .605$, $\eta_p^2 = .01$ and lever, $F(1,\ 44) = 2.21$, $p = .144$, $\eta_p^2 = .05$, but not of reward, $F(1,\ 44) = 2.87$, $p = .097$, $\eta_p^2 = .06$ or classification, $F(1,\ 44) = 0.39$, $p = .536$, $\eta_p^2 < .01$. There was a two-way interaction between block and lever, $F(2.27,\ 99.98) = 5.05$, $p = .006$, $\eta_p^2 = .10$, but not between reward and classification, $F(1,\ 44) = 2.29$, $p = .137$, $\eta_p^2 = .05$, not between reward and block, $F(3.06,\ 134.81) = 1.37$, $p = .253$, $\eta_p^2 = .03$, classification and block, $F(3.06,\ 134.81) = 0.32$, $p = .813$, $\eta_p^2 < .01$, reward and lever, $F(1,\ 44) = 0.11$, $p = .746$, $\eta_p^2 < .01$ or between classification and lever, $F(1,\ 44) = 4.27$, $p = .045$, $\eta_p^2 = .09$. There was a triple interaction between reward, classification and lever, $F(1,\ 44) = 6.09$, $p = .018$, $\eta_p^2 = .12$, but between reward, classification and block, $F(3.06, 134.81) = 0.36$, $p = .786$, $\eta_p^2 < .01$, reward, block and lever, $F(2.27, 99.98) = 0.83$, $p = .454$, $\eta_p^2 = .02$, or classification, block and lever, $F(2.27,\ 99.98) = 0.56$, $p = .594$, $\eta_p^2 = .01$. There was not a four-way interaction between reward, classification, block and lever, $F(2.27,\ 99.98) = 1.40$, $p = .250$, $\eta_p^2 = .03$.

Discussion

The purpose of this experiment was to replicate the findings in Experiment 3. However, the results from this experiment are not able to confirm or disprove these results, as very few animals showed sign-tracking behaviour. The median split used to classify the rats in ST and GT, labels animals as GT even if they have very few lever-presses. This has not been an issue in the past, as the distribution of ST and GT was equal, with animals that exclusively lever-press, others that only engage with the food-well, and a majority that display both types of behaviour despite a bias to one or other. Differences in cohorts in the distribution of behaviour are documented (e.g., Fitzpatrick et al., 2013), however for this experiment the distribution materially limits the conclusions that can be drawn.

4.1.4 Discussion Experiments 3 and 4

The purpose of Experiment 3 and Experiment 4 was to investigate the prediction from HeiDI that increasing reward value, and thus presumably increasing β_{US}, would selectively increase the rates of goal-tracking behaviour relative to sign-tracking behaviour. Experiment 3 suggested that increasing US value by providing two reward pellets rather than one results in higher levels of sign-tracking, contrary to HeiDI predictions. However, Experiment 4 did not replicate these findings. Instead, Experiment 4 showed no differences in conditioned behaviour for the manipulation of reward magnitude. The results from Experiment 4, however, need to be treated with caution because the distribution of behaviour, heavily skewed towards goal-tracking behaviour, is probably not sensitive enough to provide any insight into the patterns of lever press and food-well entry behaviours across both ST and GT animals.

Notwithstanding the potential issues with Experiment 4, it is relatively clear across these experiments that increasing the number of reward pellets from one to two did not - as may have been predicted by the HeiDI model - produce an increase in goal-tracking behaviour relative to sign-tracking behaviour. That said, the simulations provided in Chapter 3 relating to downshift unblocking do suggest that changing the number of rewards may have effects beyond simply increasing the overall value of reward. Thus a more diagnostic test of the predictions drawn from HeiDI regarding changes in reward magnitude (and presumably of Bus) may involve manipulations that do not confound reward number with magnitude (e.g. using different types of pellet where one is preferable to the other, or different volumes of sucrose reward). Given the importance of US value in the control of behaviour - especially potentially maladaptive behaviour (e.g., Flagel et al., 2009; Meyer et al., 2012) - such future studies would be important.

In addition, while Experiments 3 and 4 did not conclusively address one of the novel predictions from HeiDI, they did reinforce the findings from Experiment 1 that reversal of the reward contingencies across levers had a faster/larger impact on goal-tracking behaviour than on sign-tracking behaviour in both predominantly ST or GT animals. The relative difference in sensitivity to contingency reversal is accounted for by HeiDI.

4.2 Feature positive effect

4.2.1 Introduction

Jenkins and Sainsbury (1970) investigated discrimination learning in pigeons using paradigms in which exposed pigeons to either a feature positive design (FP; e.g., AB+/B-) or feature negative design (FN; e.g., B+/AB-). For the feature positive design, a compound of visual stimuli (e.g., AB) would be followed by a reward, but the presentation of only one of the elements (e.g., B) would not be followed by a reward. For the feature negative design (B+, AB-) a single element is followed by a reward (e.g., B), but the presentation of a compound which contains the unique element would not be followed by a reward (e.g., AB). They noticed that the birds in the FP group acquired the discrimination rapidly and ceased responding to B when presented alone, however the birds in the FN group exhibited little evidence of such learning. Another interesting finding was that the birds started directing most of the pecks in the FP group towards A, the unique feature. The superiority in learning a feature positive discrimination compared to a feature negative discrimination is a robust effect and has been seen in rats (e.g., Crowell & Bernhardt, 1979), honeybees (e.g., Abramson et al., 2013), as well as humans (e.g., Lotz, Uengoer, Koenig, Pearce, & Lachnit, 2012). In the study of Jenkins and Sainsbury (1970), pigeons showed no sign of learning in the feature negative group, however there are instances where this discrimination can be acquired. The evidence suggests that learning in a feature negative discrimination is slower than in a matched feature positive discrimination, and may happen only under particular conditions (e.g., more likely to observe learning in a sequential but not simultaneous presentation of the stimuli, see Gokey & Collins, 1980).

Simulations of HeiDI's predictions are shown in Figure 17 and Figure 18. Figure 17 shows simulations of conditioning for a FP design (Panel A and B) and FN design (Panel C and D) for when parameters are arranged for a bias towards R_{CS} (α_A and $\alpha_B =$ .50, β_{US} = .30.). Examinations of Figure 17 (Panels A and C) shows that the difference in associative strength between AB and B is larger for the FP than FN design - a clear feature positive effect. However, when these associative strengths are decomposed into R_{CS} and R_{US} (Panels B and D) it is apparent that the difference in response tendency between AB and B between FP and FN designs is particularly apparent for R_{CS} and negligible for $\mathrm{R_{US}}$. Thus, HeiDI predicts that in animals biased to sign-tracking, the feature positive effect should be more apparent in their sign-tracking behaviour than their goal-tracking behaviour.

Figure 18 shows simulations of conditioning for a FP design (Panel A and B) and FN design (Panel C and D) for when parameters are arranged for a bias towards $\mathrm{R_{US}}$ (α_A and $\alpha_B = .30$, $\beta_{US} = .50$). Examination of Panels A and C of Figure 18 also shows a clear difference in associative strength between AB and A consistent with a feature positive effect when the parameters reflect a bias towards goal-tracking behaviour. However, when these associative strengths are decomposed into $\mathrm{R_{CS}}$ and $\mathrm{R_{US}}$ (Panels B and D) it is apparent that the difference in response tendency between AB and B between FP and FN designs is larger for $\mathrm{R_{CS}}$ than $\mathrm{R_{US}}$. Thus, HeiDI predicts that even in animals biased to goal-tracking, the feature positive effect should be more apparent in their sign-tracking behaviour than their goal-tracking behaviour.

To test these predictions rats were exposed to either a feature positive (FP) or a feature negative (FN) discrimination learning. The purpose of this experiment was to see how a feature positive and negative discrimination learning translated into performance. Moreover, it is known that different types of stimuli lead to different sign-tracking behaviour (e.g., Hearst & Franklin, 1977). The simulations predict different levels of observed differences when the bias is arranged towards the CS or US and for this reason, for half of the animals, A is a lever (design Table 4) and for the other half is a light (design Table 5), which potentially imply different levels of α_{CS}. The animals in the FP group were exposed to two types of trials: a compound AB, which was followed by a reinforcer (AB+) and single presentations of B, which was non-reinforced (B-). The animals in the FN group received presentations of the compound AB which were not reinforced (AB-) and presentations of B, which were followed by a reinforcer (B+). B was always a lever and A was for half the animals a lever (Experiment 5: Lever condition; e.g., for a particular animal A is left lever and B is right lever) and for the other half was a light (Experiment 6: Light condition; e.g., for a particular animal A is left light and B is left lever). When A is a lever, Experiment 5, I record lever presses and I am able to look at how the animals interact both with A and B, but for when A is a light, the animals can approach and orient towards the light, however this behaviour was not recorded.

4.2.2 Experiment 5: Lever condition

Method

Animals and apparatus

Thirty-two Lister Hooded naïve male rats were used (supplied by Charles River, UK). Their mean ad libitum weight before the start of the experiment was 286g (range: 244-326g) and they were maintained at between 85 and 95 % of these weights by giving them restricted access to food at the end of each day. The rats had continuous access to water when they were in their cages. The conditioning boxes used in this experiment are described in Experiment 1.

Procedure

The rats had two 24-min pre-training sessions where food pellets were delivered on a variable-time (VT) 60-s schedule (range: 40–80 s). Rats then received a single session of training on each of the next 12 days of training, which occurred at the same time of day for a given rat. The experiment consisted in two conditions: feature positive (FP) and feature negative (FN). The animals in the FP group received two types of trials: AB, which was followed by a reinforcer (AB+) and B, which was non-reinforced (B-). The animals in the FN group received presentations of AB which were not reinforced (AB-) and presentations of B, which was followed by a reinforcer (B+). A and B were two levers (counterbalanced, for half the animals A was the left lever and B the right lever, and for the remaining animals A was the right lever and B was the left lever). On AB trials, the two levers were presented simultaneously. All the animals had 20 trials of the compound AB and 20 presentation of B per session, one session per day. The order in which type of trials were presented was random with the constraint that there could not be more than three same type trials in succession. The trials were delivered on a variable-time (VT) 60-s schedule (range: 40–80 s). Rats were randomly assigned to the experimental groups.

Table 4: Design Of the Experiment 5: Lever Condition

Classification	Group	Training
ST	Feature positive (FP)	AB + / B -
	Feature negative (FN)	AB - / B +
GT	Feature positive (FP)	AB + / B -
	Feature negative (FN)	AB - / B +

Note: ST refers to a sign-tracker and GT refers to goal-tracker. Rats were classified as ST or GT on the basis of their bias towards lever pressing or entering the food-well during the final block of training on the reinforced lever (or compound). Both "A" and "B" were levers. "+" denotes the stimulus or compound was reinforced, and "-" denotes that it was non-reinforced.

Data Analysis

Successive sessions during the training were combined into 6×2-day blocks (6 × training: T1-T6). At the end of the training phase, the rats were split into two groups, sign-trackers (ST) and goal-trackers (GT), based on their tendency to engage with the lever and the food-well. The rats were split into two groups, STs and GTs, using the bias score described in Experiment 1. Bias scores above -.12 were classified as GTs and bias scores below -.18 as STs. The distribution of sign-tracking and goal-tracking behaviour is shown in Figure 19.

Mixed ANOVAs were conducted separately for lever pressing and food-well entries with within-subjects factors of block (6 levels: T1-T6) and lever (2 levels: reinforced vs non-reinforced), and between-subjects factors feature (2 levels: FP vs FN) and classification (2 levels: ST vs GT). To look at how rats respond differently to A and B when learning occurrs a separate analysis is also presented, where comparisons are made for responding to A, B and B(AB) (when B is part of the compound AB).

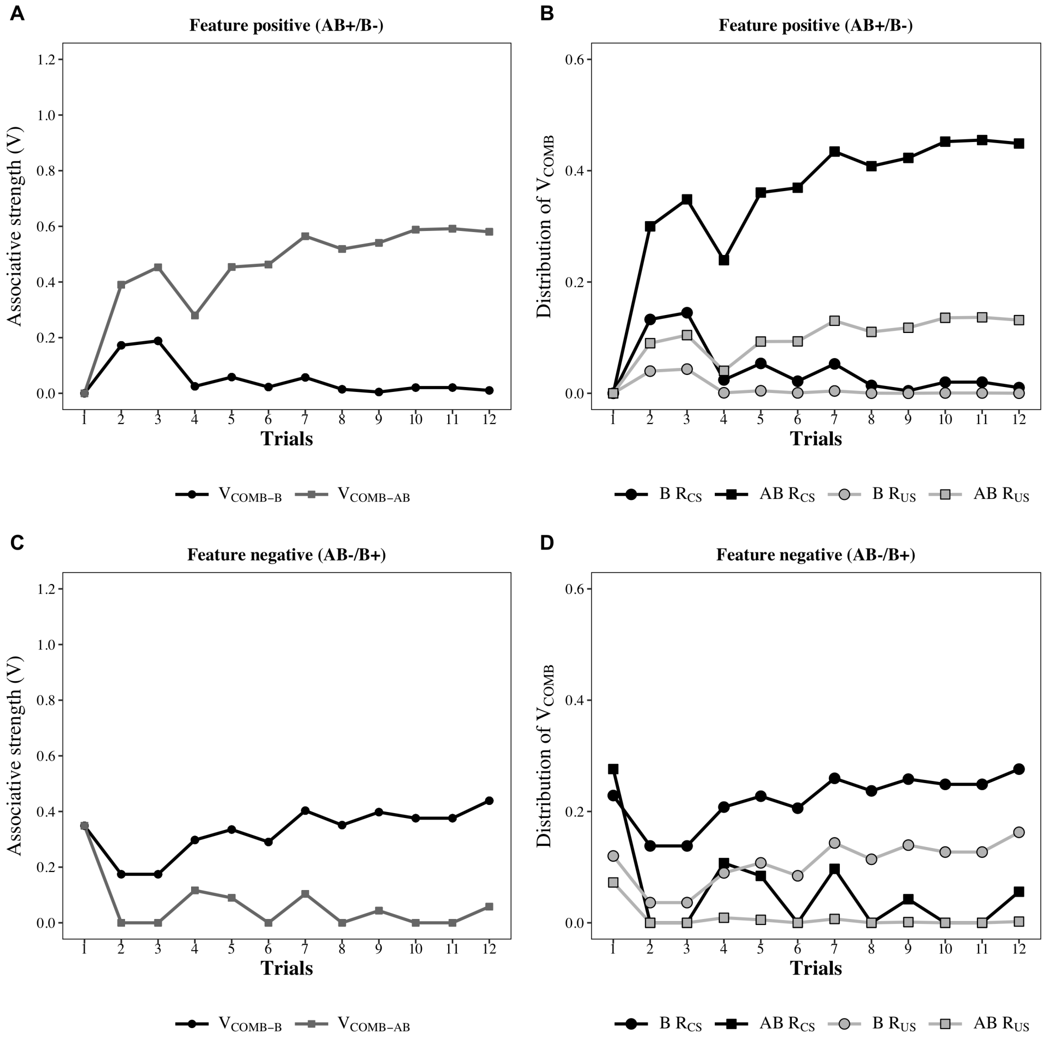


Figure 17: Simulations of conditioning for a feature positive design (FP; Panel A and B) and a feature negative design (FN; Panel C and D). Panels A and C depict the output values for $V_{\text{COMB-B}}$ and $V_{\text{COMB-AB}}$ and panels B and D show the corresponding output values for R_{CS} and R_{US} for both B and AB. The parameters were chosen to result in a bias towards the CS with α_A and $\alpha_B = .50$, $\beta_{US} = .30$.

Results

The main results from Experiment 5 are shown in Figure 20 and in Figure 21. My analysis will begin with the lever condition, where I separate by response (lever pressing and food-well entries) and condition (feature positive and feature negative). This analysis also includes separating between “A”, “B” and “B(AB)” for lever pressing. For the light condition, I separate by response (lever pressing and food-well entries) and condition (feature positive and feature negative).

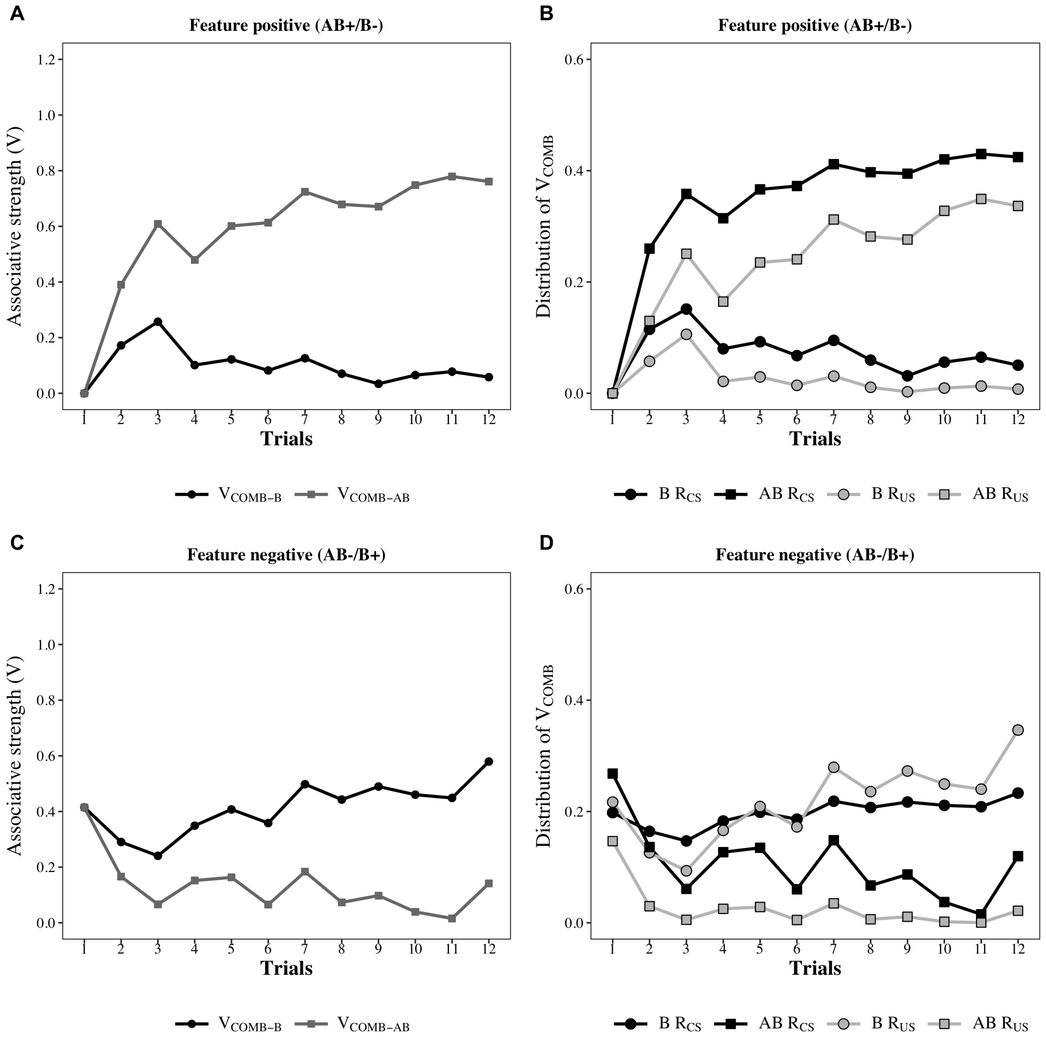


Figure 18: Simulations of conditioning for a feature positive design (FP; Panel A and B) and a feature negative design (FN; Panel C and D). Panels A and C depict the output values for $V_{\text{COMB-B}}$ and $V_{\text{COMB-AB}}$ and panels B and D show the corresponding output values for R_{CS} and R_{US} for both B and AB. The parameters were chosen to result in a bias towards the US with α_A and $\alpha_B = .30$, $\beta_{US} = .50$.

Lever presses

Inspection of the left and right-hand side of the upper panel Figure 20, which depicts lever presses for the lever condition, suggests that the animals in the FP group learn the discrimination between reinforced and non-reinforced stimuli better than the FN group. For the FP condition both STs and GTs show a discrimination between reinforced and non-reinforced lever, which is not observed in the FN group. This description is supported by the following analysis.

An ANOVA conducted for lever presses showed main effects of feature (feature

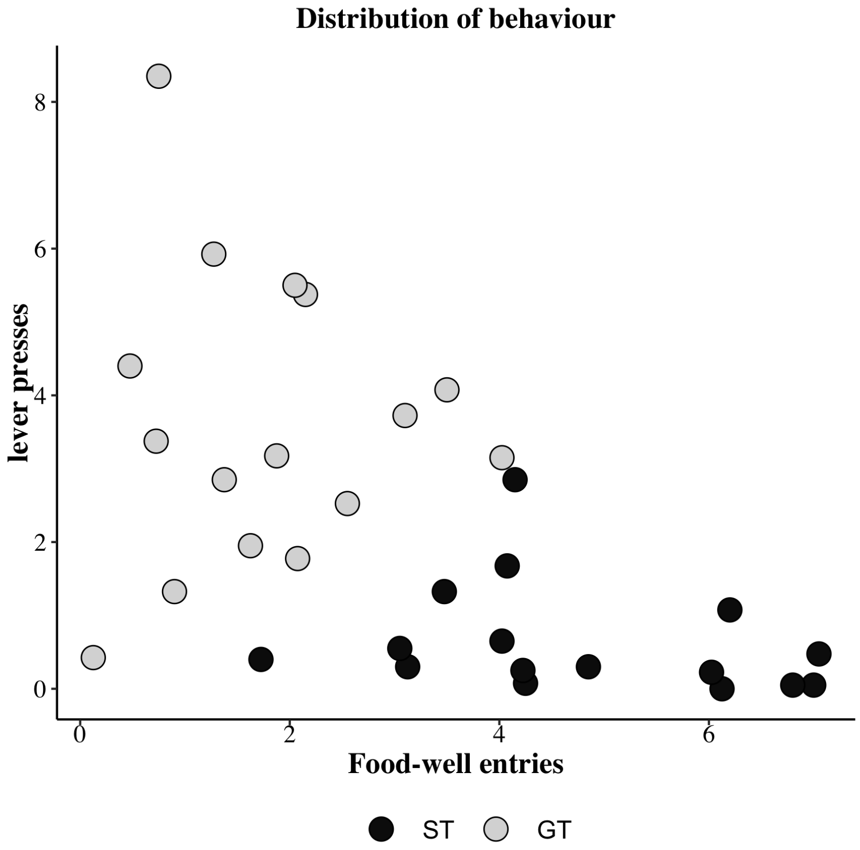


Figure 19: Distribution of sign-tracking (lever presses) and goal-tracking (food-well entries) behaviour for Experiment 5 per (10-s) trial during last block of training (T6) for L_1 (+ || -). The black symbols correspond to sign-trackers (ST) and the clear symbols to goal-trackers (GT).

positive vs feature negative), $F(1,\ 28) = 11.56$, $p = .002$, $\eta_p^2 = .29$, classification, $F(1,\ 28) = 26.81$, $p < .001$, $\eta_p^2 = .49$, block, $F(3.55,\ 99.43) = 41.71$, $p < .001$, $\eta_p^2 = .60$, and lever, $F(1,\ 28) = 68.61$, $p < .001$, $\eta_p^2 = .71$. There was also an interaction between feature and classification, $F(1,\ 28) = 8.14$, $p = .008$, $\eta_p^2 = .23$, feature and block, $F(3.55, 99.43) = 7.05$, $p < .001$, $\eta_p^2 = .20$, classification and block, $F(3.55, 99.43) = 13.83$, $p < .001$, $\eta_p^2 = .33$, critically between feature and lever, $F(1,\ 28) = 95.95$, $p < .001$, $\eta_p^2 = .77$ (reflecting the better discrimination in the feature positive groups), classification and lever, $F(1,\ 28) = 12.62$, $p = .001$, $\eta_p^2 = .31$ and between block and lever, $F(2.15,\ 60.27) = 11.87$, $p < .001$, $\eta_p^2 = .30$. There were triple interactions between feature, classification and block, $F(3.55,\ 99.43) = 5.36$, $p < .001$, $\eta_p^2 = .16$, between feature, classification and lever, $F(1,\ 28) = 10.79$, $p = .003$, $\eta_p^2 = .28$, and between

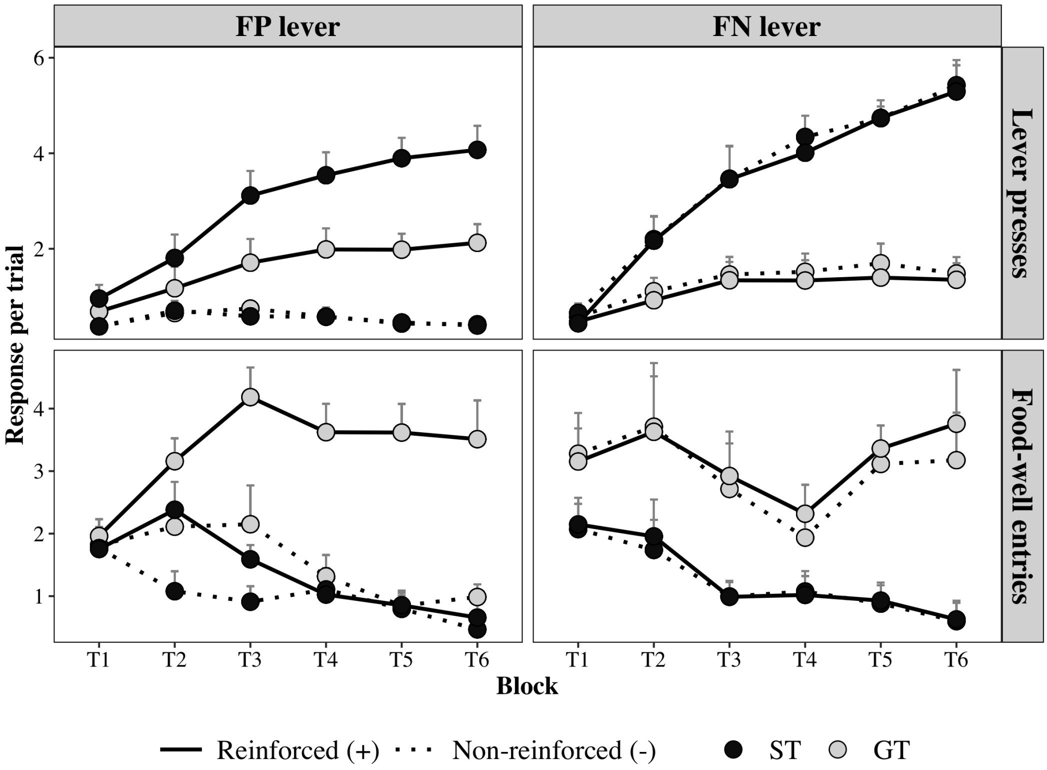


Figure 20: Results from Experiment 5: the effects of a feature positive design on sign-tracking and goal-tracking. Mean (+SEM) lever presses (upper panels) and food-well entries (lower panels) per (10-s) trial across training (T1–T6) separated by condition (left panels: feature positive; right panels: feature negative). Rats in group FP received reinforced trials of a compound AB "Reinforced (+)" and non-reinforced trials of B "Non-reinforced (-)" and rats in group FN received non-reinforced trials of a compound AB "Non-reinforced (-)" and reinforced trials of B "Reinforced (-)". The black symbols correspond to sign-trackers (ST) and the grey symbols to goal-trackers (GT).

feature, block and lever, $F(2.15,\ 60.27) = 12.48$, $p < .001$, $\eta^2_p = .31$, but not between classification, block and lever, $F(2.15,\ 60.27) = 2.19$, $p = .117$, $\eta^2_p = .07$. The four-way interaction was not significant, $F(2.15,\ 60.27) = 1.65$, $p = .198$, $\eta^2_p = .06$.

Th critical interaction between feature and lever, ($F(1,\ 28) = 95.95$, $p < .001$, $\eta^2_p = .77$) is further analysed to address whether the discrimination is acquired at all in the FN group.

Feature positive: An ANOVA conducted for lever presses for feature positive design showed main effects of classification, $F(1,\ 14) = 6.58$, $p = .022$, $\eta^2_p = .32$, block, $F(1.97,\ 27.61) = 11.82$, $p < .001$, $\eta^2_p = .46$, and lever, $F(1,\ 14) = 89.35$, $p < .001$, $\eta^2_p = .86$. There were interactions between classification and lever, $F(1,\ 14) = 12.78$,

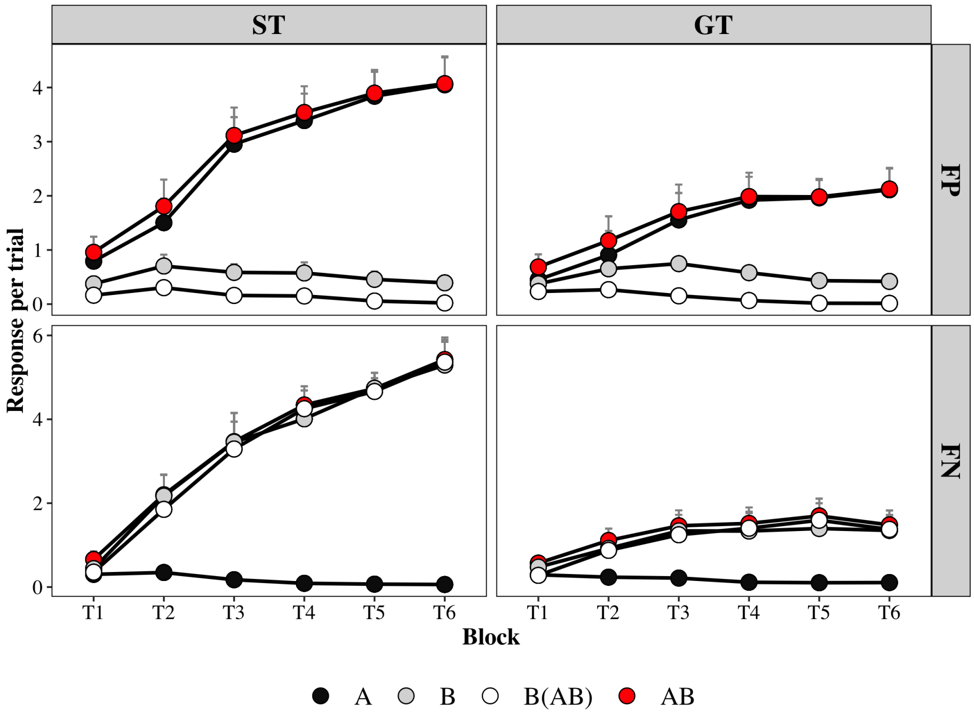


Figure 21: Results from Experiment 5 split by type of stimuli: A vs B vs B(AB). Mean (+SEM) lever presses trial across training (T1–T6) separated by classification (left panels: ST, sign-trackers; right panels: GTs, goal-trackers) and design (upper panels: FP; lower panels: FN). Rats in group FP received reinforced trials of a compound AB, and non-reinforced trials of B, and rats in group FN received non-reinforced trials of a compound AB, and reinforced trials of B

p = .003, η_p^2 = .48 and between block and lever, $F(1.78,\ 24.93) = 14.10$, $p < .001$, η_p^2 = .50, but not between classification and block, $F(1.97,\ 27.61) = 1.62$, p = .216, η_p^2 = .10, and there was no triple interaction between classification, block and lever, $F(1.78,\ 24.93) = 2.11$, p = .147, η_p^2 = .13. Next, I analyse the interaction between classification and lever by conducting separated ANOVAs for the ST and the GT groups. An ANOVA conducted for the ST group for lever presses (FP) shows main effects of block, $F(1.56,\ 9.38) = 6.87$, p = .018, η_p^2 = .53 lever, $F(1,\ 6) = 156.78$, $p < .001$, η_p^2 = .96, and an interaction between block and lever, $F(5,\ 30) = 8.26$, $p < .001$, η_p^2 = .58. A parallel analysis for the GT group shows main effects of block, $F(5,\ 40) = 4.17$, p = .004, η_p^2 = .34, lever, $F(1,\ 8) = 14.21$, p = .005, η_p^2 = .64 and an interaction between block and lever, $F(1.95,\ 15.58) = 4.72$, p = .026, η_p^2 = .37. Both ST and GT learn the discrimination between the reinforced and the non-reinforced lever, however the effect size is bigger for ST.

Feature negative: A parallel analysis was conducted for the feature negative design which showed main effects of classification, $F(1,\ 14) = 20.29$, $p < .001$, $\eta_p^2 = .59$, block, $F(5,\ 70) = 30.19$, $p < .001$, $\eta_p^2 = .68$, and lever, $F(1,\ 14) = 6.69$, $p = .022$, $\eta_p^2 = .32$. There was an interaction between classification and block, $F(5,\ 70) = 13.28$, $p < .001$, $\eta_p^2 = .49$, but not between classification and lever, $F(1,\ 14) = 0.21$, $p = .654$, $\eta_p^2 = .01$, or critically between block and lever, $F(5,\ 70) = 0.41$, $p = .838$, $\eta_p^2 = .03$. There was not a triple interaction between classification, block and lever, $F(5,\ 70) = 0.79$, $p = .562$, $\eta_p^2 = .05$. It is worth noting that there were more responses on the non-reinforced lever than the reinforced lever, the opposite of an accurate discrimination.

Food-well entries

Inspection of the left and right-hand side of the lower panel in Figure 20, which depicts food-well entries for the lever condition, suggests that the GT animals in the FP group learn the discrimination between reinforced and non-reinforced stimuli better than ST. The animals in the FN group do not learn the discrimination. This description is supported by the following analysis.

An ANOVA conducted for food-well entries revealed main effects of classification, $F(1,\ 28) = 31.07$, $p < .001$, $\eta_p^2 = .53$, block, $F(2.58,\ 72.32) = 3.50$, $p = .025$, $\eta_p^2 = .11$, and lever, $F(1,\ 28) = 28.88$, $p < .001$, $\eta_p^2 = .51$, but no effect of feature, $F(1,\ 28) = 1.62$, $p = .214$, $\eta_p^2 = .05$. There were interactions between classification and block, $F(2.58,\ 72.32) = 2.98$, $p = .044$, $\eta_p^2 = .10$, critically between feature and lever, $F(1,\ 28) = 18.05$, $p < .001$, $\eta_p^2 = .39$, and between classification and lever, $F(1,\ 28) = 12.70$, $p = .001$, $\eta_p^2 = .31$, but not between feature and classification, $F(1, 28) = 1.17$, $p = .290$, $\eta_p^2 = .04$, or between feature and block, $F(2.58, 72.32) = 1.93$, $p = .140$, $\eta_p^2 = .06$. There were triple interactions between feature, classification and lever, $F(1,\ 28) = 8.37$, $p = .007$, $\eta_p^2 = .23$ and between classification, block and lever, $F(2.77,\ 77.65) = 7.90$, $p < .001$, $\eta_p^2 = .22$, but not between feature, classification and block, $F(2.77, 77.65) = 4.55$, $p = .007$, $\eta_p^2 = .14$ $F(2.58, 72.32) = 0.87$, $p = .449$, $\eta_p^2 = .03$, or between feature, block and lever, $F(2.77, 77.65) = 2.50$, $p = .070$, $\eta_p^2 = .08$. There was a four-way interaction, $F(2.77, 77.65) = 2.98$, $p = .040$, $\eta_p^2 = .10$.

Feature positive: An ANOVA conducted for food-well entries in the feature positive design showed main effects of classification, $F(1, 14) = 15.25$, $p = .002$, $\eta_p^2 = .52$, block, $F(2.89, 40.52) = 4.40$, $p = .010$, $\eta_p^2 = .24$ and lever, $F(1, 14) = 24.15$, $p < .001$, $\eta_p^2 = .63$. There were interactions between classification and block, $F(2.89, 40.52) = 4.19$, $p = .012$, $\eta_p^2 = .23$, classification and lever, $F(1, 14) = 10.87$, $p = .005$, $\eta_p^2 = .44$ and critically block and lever, $F(2.29, 32.04) = 4.26$, $p = .019$, $\eta_p^2 = .23$, and a triple interaction between classification, block and lever, $F(2.29, 32.04) = 6.42$, $p = .003$, $\eta_p^2 = .31$. The results suggest that animals learn the FP discrimination (block and lever interaction, $F(2.29, 32.04) = 4.26$, $p = .019$, $\eta_p^2 = .23$) however the interaction between classification and lever ($F(1, 14) = 10.87$, $p = .005$, $\eta_p^2 = .44$) suggests learning is better for GTs animals. Next, I analyse the interaction between classification and lever by conducting separated ANOVAs for the ST and the GT groups. An ANOVA conducted for the ST group for food-well entries (FP) shows a main effect of block, $F(2.38, 14.26) = 6.50$, $p = .008$, $\eta_p^2 = .52$ and interaction between block and lever, $F(5, 30) = 8.31$, $p < .001$, $\eta_p^2 = .58$, however there is no main effect of lever, $F(1, 6) = 1.27$, $p = .303$, $\eta_p^2 = .17$. A parallel analysis for the GT group, shows main effects of block, $F(5, 40) = 3.49$, $p = .010$, $\eta_p^2 = .30$, lever, $F(1, 8) = 36.23$, $p < .001$, $\eta_p^2 = .82$ and an interaction between block and lever, $F(2.04, 16.31) = 6.09$, $p = .010$, $\eta_p^2 = .43$. These results suggest that the GT group shows learning between the reinforced and the non-reinforced lever. STs show a difference between the reinforced and non-reinforced lever suggested by the block and lever interaction, however this difference is less clear with no overall effect of lever.

Feature negative: A parallel analysis conducted for the feature negative design revealed main effects of classification, $F(1, 14) = 16.55$, $p = .001$, $\eta_p^2 = .54$ and lever, $F(1, 14) = 7.63$, $p = .015$, $\eta_p^2 = .35$, but not an effect of block, $F(2.12, 29.74) = 2.26$, $p = .119$, $\eta_p^2 = .14$. There was not an interaction between classification and block, $F(2.12, 29.74) = 1.31$, $p = .286$, $\eta_p^2 = .09$, classification and lever, $F(1, 14) = 2.72$,

$p = .121$, $\eta_p^2 = .16$, critically block and lever, $F(5,\ 70) = 0.75$, $p = .591$, $\eta_p^2 = .05$, and no triple interaction between classification, block and lever, $F(5,\ 70) = 1.77$, $p = .130$, $\eta_p^2 = .11$. These results suggest the animals might discriminate between the reinforced and non-reinforced stimuli, but the effect is not large enough to show a block and lever interaction. Numerically this seems to be true only for GTs, possibly due to a floor effect in STs, however there was no interaction between lever and classification.

Analysis by stimulus: A vs B vs B(AB)

As previously mentioned in the data analysis section, for Experiment 5, both responding to A and B was measured, either alone, or within AB compound. As the stimuli A, B and B(AB) have different meaning depending of the condition this will be analysed separated for feature positive and negative, respectively. Inspection of the FP group (upper panel of Figure 21), where the compound AB was reinforced and the presentation of B was not, suggests that responding to the reinforced compound is mainly driven by responding to A. Neither, STs or GTs learn the discrimination between the presentation of B as part of the compound (reinforced) and non-reinforced presentations of B, interacting more with B when presented alone. Inspection of the FN group (lower panel of Figure 21), where the compound AB was non-reinforced and the presentation of B was reinforced, suggests that the responding is mainly driven by B, however the animals do not learn when B predicts or not a reinforcer, showing no evidence of learning. This description is supported by the following analysis.

Feature positive: An ANOVA conducted for the feature positive design on lever presses in the lever condition separated by stimulus, A vs B vs B(AB), showed main effects of classification, $F(1, 14) = 6.58$, $p = .022$, $\eta_p^2 = .32$, block, $F(1.97, 27.61) = 11.82$, $p < .001$, $\eta_p^2 = .46$ and stimulus, $F(1.07,\ 15.04) = 91.27$, $p < .001$, $\eta_p^2 = .87$. There was an interaction between classification and stimulus, $F(1.07,\ 15.04) = 11.11$, $p = .004$, $\eta_p^2 = .44$, and between block and stimulus, $F(2.17,\ 30.31) = 16.35$, $p < .001$, $\eta_p^2 = .54$, but not between classification and block, $F(1.97,\ 27.61) = 1.62$, $p = .216$, $\eta_p^2 = .10$.

There was not a three way interaction, $F(2.17,\ 30.31) = 1.63$, $p = .211$, $\eta_p^2 = .10$. Next, I analyse separately A vs B and B vs B(AB), to see which stimuli the animals discriminate between.

An ANOVA conducted the for feature positive design, with levels for stimulus A vs B, revealed main effects of classification, $F(1,\ 14) = 7.18$, $p = .018$, $\eta_p^2 = .34$, block, $F(1.97,\ 27.61) = 14.43$, $p < .001$, $\eta_p^2 = .51$ and stimulus, $F(1,\ 14) = 77.04$, $p < .001$, $\eta_p^2 = .85$. There was an interaction between classification and stimulus, $F(1,\ 14) = 12.60$, $p = .003$, $\eta_p^2 = .47$, and between block and stimulus, $F(2.00,\ 28.06) = 15.81$, $p < .001$, $\eta_p^2 = .53$, but not between classification and block, $F(1.97,\ 27.61) = 1.54$, $p = .233$, $\eta_p^2 = .10$. There was not a three-way interaction. $F(2.00,\ 28.06) = 1.81$, $p = .182$, $\eta_p^2 = .11$.

A parallel comparison between B and B(AB), revealed main effects of block, $F(2.73,\ 38.16) = 3.64$, $p = .024$, $\eta_p^2 = .21$, and stimulus (B vs B(AB)), $F(1,\ 14) = 79.69$, $p < .001$, $\eta_p^2 = .85$, but not classification, $F(1,\ 14) = 0.00$, $p = .991$, $\eta_p^2 < .01$. There was an interaction between block and stimulus, $F(5,\ 70) = 2.97$, $p = .017$, $\eta_p^2 = .18$, but not between classification and block, $F(2.73,\ 38.16) = 0.20$, $p = .881$, $\eta_p^2 = .01$, or between classification and stimulus, $F(1,\ 14) = 0.19$, $p = .673$, $\eta_p^2 = .01$, or a three-way interaction, $F(5,\ 70) = 0.38$, $p = .858$, $\eta_p^2 = .03$. The interaction between block and stimulus for the B vs B(AB) analysis ($F(5,\ 70) = 2.97$, $p = .017$, $\eta_p^2 = .18$), shows responding to B being higher when presented alone than in the compound, B(AB). This is inconsistent learning about B part of part of the compound being reinforced, suggesting that all the responding for the AB compound is driven by responding to A.

Feature negative: A parallel analysis for the feature negative design on lever presses in the lever condition separated by stimulus, showed main effects of classification, $F(1,\ 14) = 20.29$, $p < .001$, $\eta_p^2 = .59$, block, $F(5,\ 70) = 30.19$, $p < .001$, $\eta_p^2 = .68$ and stimulus, $F(1.05,\ 14.69) = 77.79$, $p < .001$, $\eta_p^2 = .85$. There was an interaction between classification and block, $F(5,\ 70) = 13.28$, $p < .001$, $\eta_p^2 = .49$, classification and stimulus, $F(1.05,\ 14.69) = 22.26$, $p < .001$, $\eta_p^2 = .61$, block and stimulus, $F(4.02,\ 56.25) = 31.61$,

$p < .001$, $\eta_p^2 = .69$ and a three-way interaction between classification, block and stimulus, $F(4.02,\, 56.25) = 11.75$, $p < .001$, $\eta_p^2 = .46$. As for FP, next, I analyse separately A vs B and B vs B(AB), to see between which stimuli the animals discriminate.

An ANOVA conducted for feature positive negative, with levels for stimulus A vs B, revealed main effects classification, $F(1,\, 14) = 17.42$, $p < .001$, $\eta_p^2 = .55$, block, $F(2.87,\, 40.14) = 23.27$, $p < .001$, $\eta_p^2 = .62$ and stimulus (A vs B), $F(1,\, 14) = 71.23$, $p < .001$, $\eta_p^2 = .84$. There was an interaction between classification and block, $F(2.87,\, 40.14) = 12.16$, $p < .001$, $\eta_p^2 = .46$, classification and stimulus, $F(1,\, 14) = 20.51$, $p < .001$, $\eta_p^2 = .59$, and critically between block and stimulus, $F(5,\, 70) = 33.41$, $p < .001$, $\eta_p^2 = .70$. There was a triple interaction between classification, block and stimulus, $F(5,\, 70) = 13.57$, $p < .001$, $\eta_p^2 = .49$.

A parallel comparison between B and B(AB), revealed main effects of classification, $F(1,\, 14) = 21.27$, $p < .001$, $\eta_p^2 = .60$, block, $F(5,\, 70) = 33.02$, $p < .001$, $\eta_p^2 = .70$, but not stimulus, $F(1,\, 14) = 0.36$, $p = .560$, $\eta_p^2 = .02$. There was an interaction between classification and block, $F(5,\, 70) = 13.67$, $p < .001$, $\eta_p^2 = .49$, but not between classification and stimulus, $F(1,\, 14) = 0.20$, $p = .660$, $\eta_p^2 = .01$, block and stimulus, $F(5,\, 70) = 1.91$, $p = .104$, $\eta_p^2 = .12$ or a three-way interaction, $F(5,\, 70) = 0.95$, $p = .454$, $\eta_p^2 = .06$. These results show that even if there is no responding for A as the non-reinforced stimulus (A vs B comparison, block and stimulus interaction, $F(5,\, 70) = 33.41$, $p < .001$, $\eta_p^2 = .70$), the animals do not learn about the fact that B is reinforced when presented alone (B vs B(AB) analysis, block and stimulus interaction, $F(5,\, 70) = 1.91$, $p = .104$, $\eta_p^2 = .12$).

Discussion

Overall, these results show a feature positive effect, animals learn FP discrimination, however there is little evidence of learning the FN discrimination. The analysis split by stimulus shows that learning in the FP design is driven by learning about the feature positive, A. Responding to B, does not show learning about B, when it is or non-reinforced. It is worth noting that both ST and GT show learning in FP as lever presses, however for food-well entries it is only evident for GT animals (upper left panel of Figure 20). This supports HeiDI's prediction that a feature positive effect would be more evident in CS-oriented responses.

4.2.3 Experiment 6: Light condition

Method

Animals and apparatus

Thirty-two male rats were used (supplied by Charles River, UK). Their mean ad libitum weight before the start of the experiment was 270g (range: 220-312g) and they were maintained at between 85 and 95 % of these weights by giving them restricted access to food at the end of each day. The rats had continuous access to water when they were in their cages. The conditioning boxes used in this experiment are described in Experiment 1.

Procedure

The procedure was identical to Experiment 5 with the exception that B was always a lever and A was a light above the stimulus B (counterbalanced, for half the animals B was the left lever and A was the left light and for the other half, B was the right lever and A was the right light).

Data analysis

The analysis has been conducted in a similar way to Experiment 5, however I did not record responding to the light in any way. This means that the comparison between reinforced and non-reinforced lever for the light condition is only based on responding to "B", the lever which was either presented alone or in compound with the light. For the FP design, B was reinforced when presented with the light and non-reinforced when presented alone and for FN design, B was reinforced when presented alone and non-reinforced when presented with the light.

Successive sessions during the training were combined into 6×2-day blocks (6 ×

Table 5: Design Of the Experiment 6: Light Condition

Classification	Group	Training
ST	Feature positive (FP)	AB + / B -
	Feature negative (FN)	AB - / B +
GT	Feature positive (FP)	AB + / B -
	Feature negative (FN)	AB - / B +

Note: ST denotes a sign-tracker; GT denotes a goal-tracker. Rats were classified as ST or GT on the basis of their bias towards lever pressing or entering the food-well during the final block of training on the reinforced lever (or compound). "A" was a light and "B" was a lever. "+" denotes the stimulus or compound were reinforced, and "-" denotes they were non-reinforced.

training: T1-T6). At the end of the training phase, the rats were split into two groups, sign-trackers (ST) and goal-trackers (GT), based on their tendency to engage with the lever and the food-well, as described in Experiment 1. Bias scores above .08 were classified as GTs and bias scores below -.08 as STs. The distribution of sign-tracking and goal-tracking behaviour is shown in Figure 22.

Results

Lever presses

Inspection of the left and right-hand side of the upper panel in Figure 23, which depicts lever presses for the light condition, suggests that neither the animals in the FP group or in the FN group learn the discrimination between reinforced and non-reinforced stimuli. In both FP and FN designs, STs show higher number of lever presses than GTs. This description is supported by the following analysis.

An ANOVA conducted for the light condition on lever presses revealed main effects of classification, $F(1,\ 28) = 11.24$, $p = .002$, $\eta_p^2 = .29$, and block, $F(2.14,\ 60.02) = 8.55$, $p < .001$, $\eta_p^2 = .23$, but no effect of feature, $F(1,\ 28) = 0.04$, $p = .838$, $\eta_p^2 < .01$ or lever, $F(1,\ 28) = 0.27$, $p = .607$, $\eta_p^2 < .01$. There were no interactions between feature and classification, $F(1,\ 28) = 0.18$, $p = .678$, $\eta_p^2 < .01$, or between feature and block,

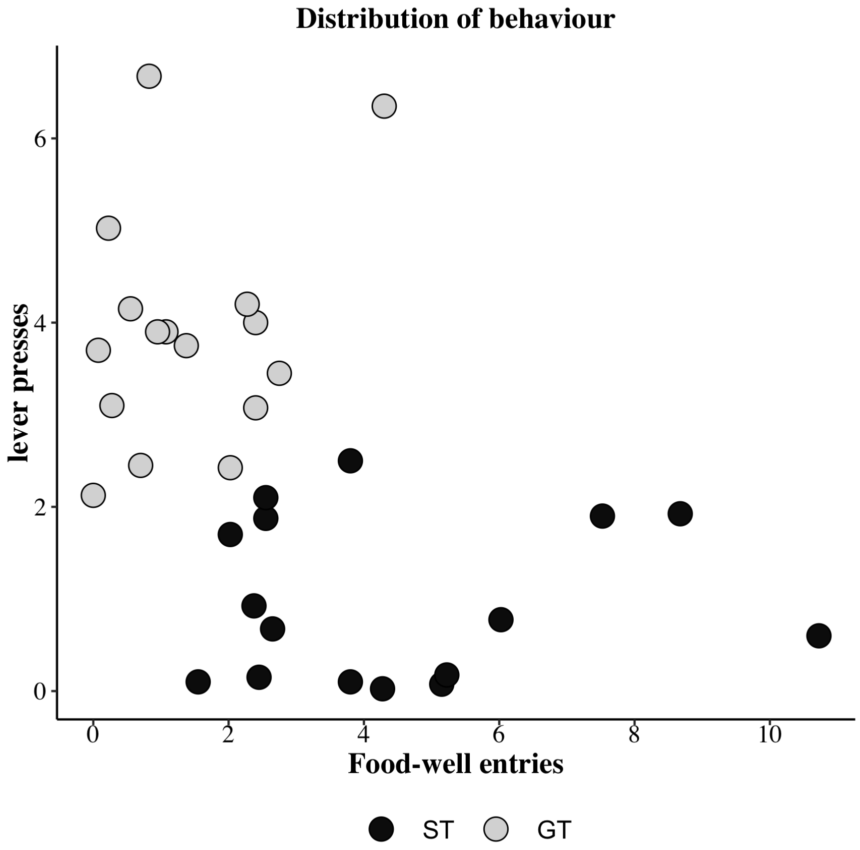


Figure 22: Distribution of sign-tracking (lever presses) and goal-tracking (food-well entries) behaviour for Experiment 6 per (10-s) trial during last block of training (T6) for L_1 (+ || -). The black symbols correspond to sign-trackers (ST) and the clear symbols to goal-trackers (GT).

$F(2.14, 60.02) = 0.20$, $p = .833$, $\eta_p^2 < .01$, classification and block, $F(2.14, 60.02) = 3.03$, $p = .052$, $\eta_p^2 = .10$, critically feature and lever, $F(1,\ 28) = 1.27$, $p = .269$, $\eta_p^2 = .04$, classification and lever, $F(1,\ 28) = 0.59$, $p = .449$, $\eta_p^2 = .02$ or between block and lever, $F(5,\ 140) = 1.39$, $p = .231$, $\eta_p^2 = .05$. There were no triple interactions between feature, classification and block, $F(2.14,\ 60.02) = 0.17$, $p = .855$, $\eta_p^2 < .01$, between feature, classification and lever, $F(1,\ 28) = 0.52$, $p = .477$, $\eta_p^2 = .02$, feature, block and lever, $F(5,\ 140) = 2.02$, $p = .080$, $\eta_p^2 = .07$ or between classification, block, lever, $F(5,\ 140) = 1.03$, $p = .403$, $\eta_p^2 = .04$. There was no four-way interaction, $F(5,\ 140) = 0.51$, $p = .772$, $\eta_p^2 = .02$. As there was no interaction between feature and lever ($F(1,\ 28) = 1.27$, $p = .269$, $\eta_p^2 = .04$) no further analysis are presented for lever presses. These results suggest that the animals do not show learning in either the FP

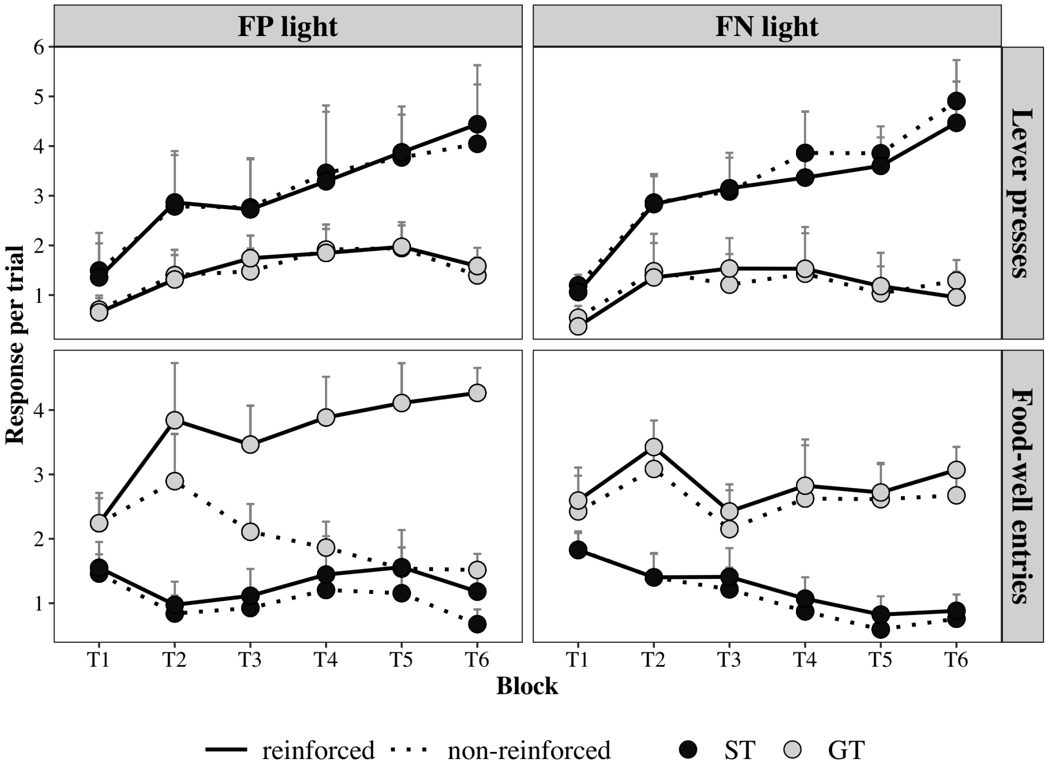


Figure 23: Results from Experiment 6: the effects of a feature positive design on sign-tracking and goal-tracking. Mean (+SEM) lever presses (upper panels) and food-well entries (lower panels) per (10-s) trial across training (T1–T6) separated by condition (left panels: feature positive; right panels: feature negative). Rats in group FP received reinforced trials of a compound AB "Reinforced (+)" and non-reinforced trials of B "Non-reinforced (-)" and rats in group FN received non-reinforced trials of a compound AB "Non-reinforced (-)" and reinforced trials of B "Reinforced (-)". The black symbols correspond to sign-trackers (ST) and the grey symbols to goal-trackers (GT).

or FN design.

Food-well entries

Inspection of the left and right-hand side of the lower panel in Figure 23, which depicts food-well entries for the light condition, suggests that the GTs animals in the FP group learn the discrimination between reinforced and non-reinforced stimuli. The animals in the FN group show no evidence of learning. This description is supported by the following analysis.

ANOVA conducted for the light condition on food-well entries revealed main effects of classification, $F(1,\ 28) = 18.88$, $p < .001$, $\eta_p^2 = .40$, and lever, $F(1,\ 28) = 17.86$, $p < .001$, $\eta_p^2 = .39$, but no main effects of feature, $F(1,\ 28) = 0.02$, $p = .881$, $\eta_p^2 < .01$

or block, $F(3.09,\ 86.63) = 0.65$, $p = .593$, $\eta_p^2 = .02$. There were interactions between feature and lever, $F(1,\ 28) = 7.99$, $p = .009$, $\eta_p^2 = .22$, between classification and lever, $F(1,\ 28) = 7.70$, $p = .010$, $\eta_p^2 = .22$ and between block and lever, $F(3.47,\ 97.22) = 5.30$, $p = .001$, $\eta_p^2 = .16$, but not between feature and classification, $F(1,\ 28) = 0.02$, $p = .881$, $\eta_p^2 < .01$, feature and block, $F(3.09,\ 86.63) = 0.55$, $p = .653$, $\eta_p^2 = .02$, or between classification and block, $F(3.09,\ 86.63) = 2.14$, $p = .099$, $\eta_p^2 = .07$. There were triple interactions between feature, block and lever, $F(3.47,\ 97.22) = 3.95$, $p = .008$, $\eta_p^2 = .12$ and between classification, block and lever, $F(3.47,\ 97.22) = 2.28$, $p = .075$, $\eta_p^2 = .08$, not between feature, classification and block, $F(3.09,\ 86.63) = 0.49$, $p = .698$, $\eta_p^2 = .02$, or between feature, classification and lever, $F(1,\ 28) = 5.32$, $p = .029$, $\eta_p^2 = .16$. There was a four-way interaction between feature, classification, block and lever, $F(3.47,\ 97.22) = 3.05$, $p = .026$, $\eta_p^2 = .10$.

Feature positive: An ANOVA conducted for food-well entries for the feature positive design revealed main effects of lever, $F(1,\ 14) = 13.02$, $p = .003$, $\eta_p^2 = .48$, but no effect of classification, $F(1,\ 14) = 7.81$, $p = .014$, $\eta_p^2 = .36$ or block, $F(2.32,\ 32.48) = 0.17$, $p = .871$, $\eta_p^2 = .01$. There was an interaction between classification and lever, $F(1,\ 14) = 6.76$, $p = .021$, $\eta_p^2 = .33$, and critically between block and lever, $F(3.07,\ 43.04) = 5.46$, $p = .003$, $\eta_p^2 = .28$, but not between classification and block, $F(2.32,\ 32.48) = 1.01$, $p = .386$, $\eta_p^2 = .07$. There was a three-way interaction between classification, block and lever, $F(3.07,\ 43.04) = 3.07$, $p = .036$, $\eta_p^2 = .18$. Next, I conduct separate analysis for the ST and GT groups, to see which animals have learnt the discrimination between the reinforced and non-reinforced stimulus. An ANOVA conducted for the ST group shows a main effect of lever, $F(1,\ 4) = 4.02$, $p = .115$, $\eta_p^2 = .50$, but no effect of block, $F(5,\ 20) = 1.30$, $p = .303$, $\eta_p^2 = .25$ and no interaction between block and lever, $F(2.46,\ 9.85) = 1.14$, $p = .370$, $\eta_p^2 = .22$. A parallel analysis for the GT group, shows a main effect of lever, $F(1,\ 10) = 22.61$, $p < .001$, $\eta_p^2 = .69$ and an interaction between block and lever, $F(5,\ 50) = 9.89$, $p < .001$, $\eta_p^2 = .50$, but no block effect, $F(2.27,\ 22.70) = 0.96$, $p = .408$, $\eta_p^2 = .09$. These results suggest that GTs show better discrimination between the reinforced and non-reinforced stimulus than STs.

Feature negative: A parallel analysis conducted for the feature negative condition revealed a main effect of classification, $F(1,\ 14) = 12.45$, $p = .003$, $\eta_p^2 = .47$, but not lever, $F(1,\ 14) = 11.00$, $p = .005$, $\eta_p^2 = .44$, or an effect of block, $F(2.12,\ 29.74) = 1.40$, $p = .262$, $\eta_p^2 = .09$. There were no interactions between classification and block, $F(2.12,\ 29.74) = 1.89$, $p = .166$, $\eta_p^2 = .12$, classification and lever, $F(1,\ 14) = 1.23$, $p = .286$, $\eta_p^2 = .08$, critcially between block and lever, $F(3.68,\ 51.50) = 0.30$, $p = .864$, $\eta_p^2 = .02$, or a three-way interaction, $F(3.68,\ 51.50) = 0.58$, $p = .663$, $\eta_p^2 = .04$. These results suggests that the animals do not learn the FN discrimination.

Discussion

These results show some evidence of FP discrimination learning, mainly driven by the GTs food-well entries, but again no evidence of learning the FN discrimination. Experiment 5 showed that the feature positive effect is mainly explained by the animals' responding to the unique feature, which in the light condition is not measured.

4.2.4 Discussion Experiments 4 and 5

The results from Experiment 5 and 6 show overall that the animals learn a feature positive better than a feature negative discrimination, where they showed little evidence of learning. There was a difference between the lever and light conditions, as the animals showed learning the discrimination by mainly interacting with A (Figure 21), and in the light condition (Experiment 6) behaviour towards A was not measured. Directing behaviour towards the unique feature is consistent with the results found by Jenkins and Sainsbury (1970). Sign-trackers expressed learning more in lever presses and goal-trackers in food-well entries, which was more obvious in the Experiment 5 (Lever experiment, (Figure 23). In Experiment 6, lever presses showed little learning in the FP design, however this is probably due to the fact that behaviour towards A was not measured (Figure 23). Goal-tracking rats show learning of a feature positive discrimination in food-well entries.

This highlights the importance of how learning is assessed, as different measurements could lead to different conclusions. For example, in Experiment 6 (Light condition), only measuring behaviour towards B would lead us to conclude either that STs have not learnt a feature positive discrimination or that learning is mainly showed by interacting with the food-well. Experiment 5 (Lever condition) reveals the fact that learning is observed in both lever presses and food-well entries in both STs and GTs, however it is mainly driven by the unique feature.

HeiDI predicted a feature positive effect, however it predicted that while a FN discrimination is harder to learn but it still should be learnt. Other novel predictions derived from HeiDI are difficult to assess as they rely on the animals learning a feature negative discrimination, however the general prediction that it should be more obvious in lever presses than food-well entries was not present in the overall results because such an effect was apparent in Experiment 5 but not Experiment 6 (albeit, sign-tracking or orientation to the positive feature light was not assessed). Most associative learning models predict a faster acquisition of a feature positive discrimination, however they also fail to predict no learning in a feature negative discrimination (e.g., Mackintosh, 1975; McLaren, Kaye, & Mackintosh, 1989; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Wagner, 1981). The conditions under which the animals are able to learn a FN discrimination are still to be elucidated.

Chapter 5

General Discussion

Theories of associative learning assume that pairings of a CS with a US result in the formation of an association between the central representations of the events (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Wagner, 1981). This association allows the CS to activate the representation or idea of the US and thereby behaviour. These theories assume a monotonic relationship between a single construct that represents learning and acquired behaviour. This view is challenged by the fact that a simple auto-shaping procedure produces marked individual differences in behaviour: some rats predominantly interact with the CS (e.g., lever), others investigate the location where the US is about to be delivered , and the remainder show patterns of behaviour in between these two extremes (e.g., Patitucci et al., 2016). Focusing on one measure (e.g., lever presses) leads to the conclusion that one group (sign-trackers) better acquire the CS-US association, while focusing on another measure (e.g., food-well entries) leads to the opposite conclusion.

The purpose of this thesis was to increase our understanding of how learning is translated into performance, and the origin of individual differences in how learning is expressed. The empirical results from Chapter 2 contributed to the development of a new model which integrates learning and performance, HeiDI (Chapter 3). This model provides new approach to known phenomena, which in the past have posed a challenge for many learning theories, but it also provides new testable predictions about how learning is translated in conditioned behaviour. In Chapter 4 I tested some of these novel predictions. I will first summarise the novel empirical findings, from Chapter 2 and 4, before moving on to discuss the implications of the theoretical advances provided by HeiDI.

5.1 Empirical findings

The few attempts at explaining individual differences in conditioned behaviour have appealed to the assumption of two learning systems (e.g., S-S/S-R, model-based/model-free), each of them generating one type of behaviour (e.g., Lesaint et al., 2014). There are at least two forms that this analysis could take. One possibility is that the two learning systems operate differently across rats, the balance favouring one system or another. If a single system governed all behaviour in a given rat then both food-well and lever-press responses should exhibit the characteristic property of that system: When governed by an S-S system (or model based), activity directed toward both the lever and the food-well will change rapidly in the face of a change in contingencies; whereas when governed by an S-R (model-free) system both will change relatively slowly. Another possibility is that both systems operate in parallel in a given rat, however each system generates one type of behaviour. This analysis predicts that a given form of response will exhibit the same characteristics independently of whether the animal in which it is observed is classified as a ST or a GT; with food-well activity being derived from the operation of a S-S system (or model based) and lever-oriented behaviour being derived from a S-R system (or model-free) that operate to different degrees in all rodents. The dominant response might be toward the food-well in one rodent and lever in another, but in both rats food-well activity should more rapidly track changes in reinforcement contingencies than should lever-oriented activity.

These predictions were assessed in Experiment 1 and Experiment 2. In Experiment 1 rats received discrimination training where the presentation of one lever (L_1) was paired with food pellets and another lever (L_2) was non-reinforced. When rats learnt the discrimination, the contingencies on the levers were reversed with L_1 now non-reinforced and L_2 reinforced. This procedure resulted in marked individual differences in conditioned responding; with some rats interacting with L1 (but not L2) and others approaching the site of food delivery during L_1 (but not L_2). Rats were classified as ST and GT on the basis of their responding in the last block of training. Results showed that when the contingencies were reversed, the different levels of lever pressing to L_1 (and L_2) in Groups ST and GT remained remarkably stable during the first block of reversal. In contrast, the levels of food-well entries changed more rapidly in both Groups ST and GT (Figure 2).

In Experiment 2 rats were presented with two levers (L_1 and L_2), both reinforced with the same type of reinforcer. As noted in Chapter 2, there were two groups of animals which had two different reinforcers: one group had pellets as a reinforcer and another group had sucrose. Previous work in my lab suggested pellets sustain more responding than sucrose. After initial training, the contingencies on the levers were then switched. The group of animals that had pellets as a reinforcer received sucrose and, the group that had sucrose received pellets. When the contingencies were switched, both ST and GT, changed their responding for food-well entries, increasing their responding from sucrose to pellets and decreasing their responding from pellets to sucrose. Lever-press responding remained stable following the switch in contingencies (Figure 5).

In both Experiment 1 and Experiment 2, lever-press behaviour was less sensitive to changes in reinforcement contingencies than was food-well behaviour. This difference in sensitivity was equally apparent in rats that were classified as STs and GTs. To summarize, these results indicate that differential sensitivity to change is a property of the behaviour that is being measured not an overall property of the animal being governed by one system or another.

As discussed in Chapter 3 (section 3.6.1. Excitatory conditioning and extinction), HeiDI predicts that V_{CS-US} and V_{COMB} decline, but V_{US-CS} does not, which leads to differential sensitivity to US changes: CS-orientated behaviour is extinguished slower than US-oriented behaviour. The results obtained in Experiment 1 and Experiment 2 are predicted by HeiDI.

Experiments 3, 4, 5 and 6 aimed to test new prediction derived from HeiDI.

HeiDI assumes that α_{CS} and β_{US} are fixed for a given CS and US in a given animal, but propose that the perceived salience of the CS (relating to α_{CS}) and US (relating to β_{US}), and hence c*α_{CS} and V_{CS-US} (Equations 4 and 5), can vary between animals. Experiment 3 examined how performance is affected by a simple manipulation, the number of food pellets (one or two), which constituted the US on a given trial. HeiDI predicts higher levels of US-oriented behaviour (or goal-tracking) for USs of higher value. However, the results from this experiment, showed that a US assumed to have a higher value (2 pellets) produced more sign-tracking compared to a US with a lower value (1 pellet), and no differences were observed in goal-tracking. The sample of rats in Experiment 3 was relatively low, and Experiment 4 aimed to replicate the findings. However, the distribution of behaviour in this cohort, with very few animals that showed any sign-tracking behaviour, made the interpretation of the results in Experiment 4 difficult. This experiment could not be expected to replicate the findings of Experiment 3, as the animals behaved atypically.

Discrimination learning paradigms show that pigeons acquire more rapidly a feature positive discrimination (FP; e.g., AB+/B-) than a feature negative discrimination (FN; e.g., B+/AB-) (Jenkins & Sainsbury, 1970). The feature positive effect is a well-established characteristic of Pavlovian phenomena generally, but relatively little research has investigated how it translates into conditioned behaviour. HeiDI predicts a feature positive effect: learning should be better in an FP design than an FN design. When the bias is arranged towards the CS (when $\alpha_{CS} > \beta_{US}$, Figure 17), this is more evident in $\mathrm{R_{CS}}$ than $\mathrm{R_{US}}$. When the bias is arranged toward the US (when $\alpha_{CS} < \beta_{US}$, Figure 18), for a FP design this discrimination should be more evident in $\mathrm{R_{CS}}$ than $\mathrm{R_{US}}$, however for an FN design the discrimination is equal or better for $\mathrm{R_{US}}$ compared to $\mathrm{R_{CS}}$ (again, suggesting the feature positive effect should be more apparent in $\mathrm{R_{CS}}$ than $\mathrm{R_{US}}$).

Simulations predicted different levels of observed differences when the bias is arranged towards the CS or US and for this reason for half of the animals A is a lever (Experiment 5) and for the other half is a light (Experiment 6), which potentially imply different levels of α.

The results from Experiment 5 and Experiment 6, show overall that the animals learnt the feature positive discrimination, however for the feature negative discrimination they showed little evidence of learning. Interestingly, the feature positive effect was more obvious in Experiment 5, where responding to A (the unique feature) was measured. Directing behaviour towards the unique feature is consistent with the results found by Jenkins and Sainsbury (1970). It is unclear why the animals did not learn the feature negative discrimination but it is possible, even though unlikely, the amount of training was not enough. Another possibility lies in the use of simultaneous rather successive presentation of the AB compound. Other novel predictions derived from HeiDI, such as the feature positive effect being more evident in sign-tracking than goal tracking, are difficult to assess. These predictions rely on the fact that animals acquire a feature negative discrimination. These experiments highlight, yet again, the importance of how learning is assessed, as different measurements could lead to different conclusions. If the conclusion were made on the basis of Experiment 6 (Light condition), measuring behaviour towards B would lead to two wrong conclusions. First, I would be inclined to conclude that STs do not learn a feature positive discrimination or second, that learning is mainly expressed by interacting with the food-well. Experiment 5 (Lever condition) reveals the fact that learning is observed in both lever presses and food-well entries in both STs and GTs, however it is mainly driven by responding to the unique feature, which in Experiment 6 is not measured.

Overall the empirical work presented in this thesis elucidates one important question: is it the animal or the behaviour? Some of the previous results have been attributed to the animals' "personality" (e.g., Ahrens et al., 2016; Bissonette et al., 2015; Flagel et al., 2009; Lesaint et al., 2014; Lovic et al., 2011). However, the results presented here indicate that the apparent "inflexibility" or "resistance to change" is a property of the behaviour being measured. An open remaining question concerns whether a unified single mechanism theory (e.g., HeiDI) is able to explain how learning translates in behaviour, or a dual-mechanism is necessary to explain all the observed differences in conditioned behaviour. The empirical work presented here has not fully addressed this crucial question. The results obtained in Experiment 1 and Experiment 2 are consistent with both of these accounts.

Experiment 3 and Experiment 4 could have provided evidence for the origin of individual differences and if they are related to the US value, however the results were inconclusive.

Other novel predictions about the distribution of conditioned behaviour between the CS and the US, were addressed in Experiment 5 and Experiment 6, however they relied on the animals learning a feature negative discrimination.

In the following section I will discuss both how HeiDI is able to accommodate known phenomena, which have posed a challenge in the past to other learning theories (including those that rely on dual-process mechanisms to explain the patters of sign- and goal-tracking behaviour), and describe other novel predictions of how learning is translated into performance.

5.2 Theoretical implications: HeiDI

In dispelling out-dated (academic textbook) descriptions of Pavlovian conditioning, Rescorla (1988, p. 151) referred to three primary issues to be addressed in the study of any learning process: “What are the circumstances that produce learning? What is the content of the learning? How does that learning affect the organism’s behaviour?”. It is perhaps especially surprising that in the context of Pavlovian learning the final issue – concerning conditioned behaviour itself - has become secondary to theorizing directed toward addressing the first two questions. Indeed formal theories of Pavlovian learning have often followed the simplifying stance expressed by Rescorla and Wagner (1972) that it is “sufficient simply to assume that the mapping of Vs into magnitude or probability of conditioned responding preserves their ordering.”. The fact that the form of conditioned behaviour depends on the nature of both the CS and US (e.g., Holland, 1977, 1984) and that there are marked individual differences in how learning is exhibited (e.g., Patitucci et al., 2016) represent a significant impetus for developing theories that recognize this variety. HeiDI does this.

5.2.1 Conditions, content and performance

I started by simplifying the Rescorla-Wagner learning rule for forward, CS-US associations, and supplementing it with a formally equivalent rule for reciprocal, US-CS associations (see Equation 1 and 2). The values returned by these equations were then combined (to form V_{COMB}) using a rule that weights the associative value of the stimulus that is present (e.g., V_{CS-US}) more than an association involving associatively activated nodes (e.g., V_{US-CS}; see Equation 3). Finally, when the CS is presented, V_{COMB} is distributed in two forms of behaviour (CS-oriented, R_{CS}, and US-oriented, R_{US}) according to the ratio of α_{CS} and V_{CS-US} (see Equation 4 and 6). The resulting model, HeiDI, provides the following answers to the three questions posed by Rescorla (1984): (1) On a given trial, learning occurs to the extent that there is a difference between the perceived salience of an event (reflected in β_{US}) and the perceived salience of the retrieved representation of that event based on the combined associative strengths of the stimuli presented on that trial ($\sum \text{V}_{TOTAL-US}$; or a difference between $\text{c}^{*}\alpha_{CS}$ and $\sum \text{V}_{TOTAL-CS}$). (2) Learning is represented in the reciprocal associations between the nodes activated by different stimuli (e.g., CS and US). (3) Performance reflects the relative intensity of the CS (as reflected in α_{CS}) and the associative strength of the CS (V_{CS-US}; which reflects β_{US} through $\text{c}^{*}\beta_{US}$) multiplied by the combined associative strengths involving the CS and US (i.e., V_{COMB}). In this way, HeiDI provides a way to capture two classes of conditioned behaviour, and individual differences therein, together with the effect of group-level manipulations.

I have highlighted the application of HeiDI to sign-tracking and goal-tracking, which are examples of the general distinction between CS-oriented and US-oriented behaviours. The spatial separation of the two classes of response and the ease with which they are automatically recorded certainly means that they have some methodological advantages over other responses (e.g., those elicited by aversive USs). Nevertheless, I assume that many Pavlovian conditioning procedures result in greater variety in conditioned responses than is routinely measured and used to guide theorizing. I have already illustrated how this practice might complicate interpretation of patterns of results in the case of blocking. However, the two classes of responses that I have considered might themselves be further divided, with the individual elements of the CS and US giving rise to the different responses defined (r1-6; see Jenkins & Moore, 1973). Expanding HeiDI to accommodate this complexity would not present specific theoretical challenge: with each individual element having its own α or β values and affiliated (unconditioned) responses. However, there are issues that do require further discussion. These involve how associations between the components of a compound stimulus might affect performance, and the nature of the representations of the CS and US.

5.2.2 Associations between the components of a compound

Conditioned responding to a CS is not only determined by whether it has a direct association with a US. For example, after exposure to a stimulus compound (AB), conditioned responding that is established to B will also be evident when A is presented (e.g., Rescorla & Cunningham, 1978). This effect is known as sensory preconditioning and it is often attributed to the formation of an associative chain that allows A to activate the US through A-B and B-US associations (but see, Lin & Honey, 2016). I have already provided an analysis of how A-B links might form (see Equation 7 and 8), and have appealed to such links in providing an analysis of downshift unblocking (cf. Rescorla & Colwill, 1983). The way in which the links in the chain can be combined to determine the level of performance generated by A can be derived from an extension of Equation 3: $V_{CHAIN} = 1/c^{*}V_{A-B}\ V_{COMB-B}$, where $V_{COMB-B}= V_{B-US} + (1/c^{*}V_{B-US} * V_{US-B})$. This formulation means that $V_{CHAIN} < V_{COMB-B}$ if $V_{A-B} < 1$. The way in which V_{CHAIN} is distributed into R_{CS} and R_{US} can be determined using Equations 4 and 5: $c^{*}\alpha_{A}$ is substituted for $c^{*}\alpha_{CS}$, $|1/c^{*}V_{A-B} * V_{B-US}|$ is substituted for $|1/c^{*}V_{CS-US}|$, and V_{CHAIN} replaces V_{COMB}. In terms of the nature of the behaviour elicited by A, the most obvious prediction is that it will mirror that evoked by B through direct conditioning (Holland, 1984). However, according to HeiDI the distribution of CS-oriented and US-oriented behaviour will differ between A and B: with CS-oriented responding being more evident (and US-oriented behaviour less evident) during A than during B: To the extent that while $c^*\alpha_A$ and $c^*\alpha_B$ will be the same, $|1/c^*V_{A-B} * V_{B-US}| < |V_{B-US}|$ (see Dwyer, 2012). This analysis of sensory preconditioning, and of the potential impact of within-compound associations in conditioning procedures more broadly, is relatively straightforward. However, there is another approach to conditioned performance that has also been applied to sensory preconditioning and cue competition effects (e.g., overshadowing and blocking). It deserves consideration because it addresses some of the same issues and phenomena as HeiDI.

The comparator model proposed by Stout and Miller (2007) focuses on how performance to a test stimulus, A, is affected by the stimuli with which it was trained (e.g., B after conditioning with an AB compound). This model builds on the idea that performance to A at test is determined by a comparison between (i) the representation of the US directly retrieved by A, and (ii) the representation of the same US indirectly retrieved by the associative chain: A-B and B-US (see Miller & Matute, 1998). In this case, B is called the comparator stimulus for A, and following pairings of AB with a US, the tendency for A to generate performance at test is held to be restricted by the fact that its comparator, B, has retrieved a memory of the US. The analysis thereby explains overshadowing and blocking, but also other findings that are problematic for an unreconstructed Rescorla-Wagner model. However, in the case of sensory preconditioning, where AB is first non-reinforced, the model is forced to assume that the fact that B has acquired excitatory associative properties during a second stage increases the potential for A to generate performance. These differing effects of the comparator term (B; termed subtractive and additive) are held to be determined by experience with comparing the US representation retrieved by A with the US representation indirectly retrieved by B. The additive effect occurs when there has been little or no opportunity to experience the two types of retrieved representations (e.g., during simple exposure to AB in sensory preconditioning), and the subtractive effect increases with experience that affords such a comparison (e.g., during multi-trial compound conditioning; Stout & Miller, 2007, p. 765). In any case, like the Rescorla-Wagner model, the more sophisticated analysis of performance developed by Stout and Miller (2007) provides no ready explanation for the fact that different behavioural measures can provide support for opposing conclusions about how associative strength is translated into performance, which is the focus of interest here. That being said, the fact that within HeiDI the distribution of CS-oriented and US-oriented components of performance reflects the relative values of c*α and V_{CS-US} involves a comparison process of sorts. Certainly, changing the associative strength of stimuli before testing will not only affect V_{COMB}, but will also affect R_{CS} and R_{US} through changing V_{CS-US}. As I have already noted, in the context of my previous discussion of blocking, a secure interpretation of the impact of such changes on performance requires behavioural assays that are sensitive to both R_{CS} and R_{US}.

5.2.3 Elemental and configural processes

A final issue, which I mentioned in the section on summation (section 3.6.4), concerns how models that do not have configural processes address the fact that animals can learn discriminations that are not linearly separable. For example, animals can learn that a tone signals food and a clicker signals no food in one experimental context and the tone signals no food and a clicker signals food in a second context (see Allman et al., 2004). This type of discrimination is interesting because an 'elemental' animal – one only capable of representing individual events – should be incapable of learning them: The tone and clicker have the same reinforcement history, as do the spotted and checked chambers, and therefore each of the four combinations or compounds should be equally capable of generating performance. There is an ongoing debate about how different combinations of the same stimuli might be represented in ways that would permit these discriminations to be acquired (e.g., Brandon et al., 2000; Honey et al., 2010; Pearce, 1994). For example, different stimulus elements of a given auditory stimulus might become active depending on the context in which they are encountered (e.g., Brandon et al., 2000), or the elements activated by a given pattern of stimulation might come to activate a shared configural representation (e.g., Honey et al., 2010; Pearce, 1994). In either case, the elements or configurations thereof (or both; see Honey, Iordanova, & Good, 2014) could be subject to the same learning and performance rules described in Equations 1 to 6 (see also Delamater, 2012). However, I should also note that the response units (r1-r6) within the proposed associative architecture for HeiDI (see Figure 7 provide another locus in which combinations of CSs and indeed USs might be represented: The strength of the connections from combinations of CSs and USs to these response units could be modified during conditioning (for a related discussion, see Honey et al., 2010).

5.2.4 Limitations and further development

I have already noted that Equation 6 provides a simplistic analysis of how changes in R_{CS} and R_{US} might affect activity in a set of response-generating units (r1-r6). However, taking a step back, what is needed in order to provide a detailed assessment of the accuracy of the predictions that I have derived from HeiDI, is estimates of the perceived salience of both the CS and US on an individual-by-individual basis. Armed with these estimates, I could then provide a quantitative analysis of the fit between predictions of the model and the behaviour of animals on an individual basis. It has been argued that palatability might provide an estimate of perceived US salience (cf. Patitucci et al., 2016), and one potential estimate of the perceived salience of a CS is the unconditioned orienting behaviour that its presentation provokes before conditioning has taken place (cf. Kaye & Pearce, 1984).

5.2.5 Concluding comments

Pavlovian conditioning has provided a fertile testbed in which to investigate issues concerning when associative learning occurs, its content, and how it is translated into performance. Of these three issues, formal models have paid least attention to how learning is translated into performance: consideration of performance has been secondary to analyses of the conditions and content of learning. HeiDI begins to redress this imbalance by providing an integrated analysis of all three issues. This analysis could be developed in order to provide a more quantitative analysis, modelling performance at an individual-by-individual level, with the characteristics of the schematic network fully specified. As already noted, it could also be extended to explicitly distinguish between different features of both the CS and the US, which could be tied to different types of response (see also Delamater, 2012). In the process of developing this relatively simple model, it has become clear that it is difficult to address one of Rescorla's three issues without a detailed consideration of the others: developing a more complete understanding of associative learning through the study of Pavlovian conditioning involves multiple constraint satisfaction (Marr, 1982). HeiDI provides general insights into learning, its content and performance that are – at least in part – born out of a more detailed analysis of the variety and individual differences in conditioned behaviour. This evidence has been too often neglected, given its theoretical importance and potential translational significance.

It is true that the empirical results presented here are not fully consistent with the novel predictions of HeiDI: increasing the number of reinforcers did not selectively influence goal-tracking over sign tracking in Experiment 3 and Experiment 4; and the feature positive effect was not generally larger for sign- than goal-tracking responses (compare Experiment 5 with Experiment 6). However, there are reasons to think that these may have been due to issues other rather than flawed predictions: changing reinforcer number may not be a direct influence on β_{US}; and the absence of feature negative discrimination makes assessing the feature positive effect ambiguous. So, both dual process (e.g., S-S vs

S-R; or model free vs model based) accounts of the differences in sign- and goal-tracking remain as viable as the single-process account offered by HeiDI. That said, these dual process accounts do not offer the same range of novel explanations for prior results that have proved problematic for existing associative theory. Thus, there is a clear argument to be made for the superiority of the type of single-process account offered by HeiDI.

Chapter 6

Appendix

6.1 Reproducibility and Open Science Framework (OSF)

This document is Sweave R project. It combines Sweave files (`.Rnw` extension) when R code needs to be integrated within the text or simple LaTeXfiles (`.tex` extension) when only text is needed. The data and the code are freely available on OSF (https://osf.io/h4fyg) website. The README file (OSF/Thesis/README/README.pdf) details the structure of the Sweave R project.

For reproducibility purposes the code for data cleaning and manipulation of the data from all the experiments reported here have been uploaded on the Open Science Framework/Experiments (osf.io/h4fyg) website. The codes for the experiments are the following:

- Experiment 1: DDR066 (Experiments/2017)
- Experiment 2: RH001 (Experiments/2017)
- Experiment 3: DDR074 (Experiments/2017)
- Experiment 4: DDR084 (Experiments/2018)
- Experiment 5 and 6: DDR087 (Experiments/2018)

6.2 HeiDI app

For conducting the simulations, I built a Shiny dashboard app (HeiDI) using R programming language. The code for the simulations together with a README file which details the necessary steps to reproduce the simulations is available on the OSF website (https://osf.io/h4fyg/).

App structure and files

The main file is ‘app.R’. This file is responsible for the main body of the app and sources the files needed to conduct the simulations: ‘formulas.html’, ‘load.packages’, ‘model.R’, ‘my_theme.R’, ‘style.css’ as well as all the files in the ‘phenomenon’ folder.

app.R: Whenever possible, to create a cleaner appearance the app sources other files. The `ui` function builds a user interface and the `server` defines the server logic. More information about shiny apps can be found at https://shiny.rstudio.com/tutorial/written-tutorial/lesson1/.

formulas.html: writes the formulas displayed in the ’Formulas’ tab in html language.

install_packages.R: this file installs all the packages used for this app.

load_packages.R: sources from the app to load all the packages.

model.R: contains functions for modelling, converting from wide to long format and labelling dataframes. Each ’model’ function has a complementary ‘to.long’ (transforms generated data to long format) and ‘label’ (re-labels variables to use for plots) function. Every ‘model’ function has commented the phenomenon for which it is being used. There are three functions and each of them is used for a set number of variables (e.g. two α values and one β value).

my_theme.R: contains functions for controlling plot formatting (e.g. text size, spacing, axis ticks) in ‘ggplot’ plots.

phenomenon folder contains files sourced from the app to create simulations for each of the phenomenon.

README folder: contains the */README.tex* file that was used to generate this document.

rsconnect folder: generated by the app.

www folder:

/style.css: controls appearance (e.g. text format, box sizes, colour) of the app.

log folder: contains log files to track changes in the app. Each numbered `log` file tracks changes using a time stamp.

The app is not final and will include simulations of several more phenomena.

6.3 Additional notes

For reference generation the “natbib” package (Daly & Ogawa, 2013) has been used.

A special thank you for the Stackoverflow community for all the support in all R, Sweave and LaTeXrelated questions.

Chapter 4, Experiment 5 and Experiment 6: In the app conditioned inhibition simulates only FN design AB-/B+ (FP not simulated yet in the app). This has no implication for the conclusions drawn, however it is worth mentioning for reproducibility purposes.

References

Abramson, C. I., Cakmak, I., Duell, M. E., Bates-Albers, L. M., Zuniga, E. M., Pendegraft, L., … Wells, H. (2013). Feature-positive and feature-negative learning in honey bees. *The Journal of experimental biology*, *216*(Pt 2), 224–9.

Ahrens, A. M., Singer, B. F., Fitzpatrick, C. J., Morrow, J. D., & Robinson, T. E. (2016). Rats that sign-track are resistant to Pavlovian but not instrumental extinction. *Behavioural brain research*, *296*, 418–430.

Aitken, M. R. F., Bennett, C. H., McLaren, I. P. L., & Mackintosh, N. J. (1996). Perceptual differentiation during categorization learning by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *22*(1), 43–50.

Allman, M. J., & Honey, R. C. (2005). Associative change in connectionist networks: an addendum. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*(3), 363–7.

Allman, M. J., Ward-Robinson, J., & Honey, R. C. (2004). Associative change in the representations acquired during conditional discriminations: further analysis of the nature of conditional learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*(2), 118–28.

Anselme, P., Robinson, M. J. F., & Berridge, K. C. (2013). Reward uncertainty enhances incentive salience attribution as sign-tracking. *Behavioural Brain Research*, *238*, 53–61.

Arcediano, F., Escobar, M., & Miller, R. R. (2005). Bidirectional Associations in Humans and Rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*(3), 301–318.

Asch, S. E., & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society*, *106*, 135–163.

Atkinson, R. C., & Estes, W. K. (1963). Stimulus sampling theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol2 ed.). New York: Wiley.

Auguie, B., & Anotonov, A. (2017). *package 'gridExtra': Miscellaneous Functions for "Grid" Graphics.* Retrieved from https://cran.r-project.org/web/packages/gridExtra/index.html

Baker, A. G., & Mercier, P. (1982). Prior experience with the conditioning events: Evidence for a rich cognitive representation. In M. L. Commons, R. J. Hernstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior* (pp. 117–143). Cambridge, MA: Ballinger.

Barnet, R. C., & Miller, R. R. (1996). Second-order excitation mediated by a backward conditioned inhibitor. *Journal of experimental psychology. Animal behavior processes*, *22*(3), 279–96.

Beierholm, U. R., Anen, C., Quartz, S., & Bossaerts, P. (2011). Separate encoding of model-based and model-free valuations in the human brain. *NeuroImage*, *58*(3), 955–962.

Bissonette, G. B., Orsini, C. A., Simon, N. W., Calu, D. J., Nasser, H. M., Chen, Y., & Fiscella, K. (2015). Individual variability in behavioral flexibility predicts sign-tracking tendency. *Front Behav Neurosci*, *9*, Article 289.

Blanchard, R. J., Fukunaga, K. K., & Blanchard, C. (1976). *Environmental control of defensive reactions to footshock* (Vol. 8; Tech. Rep. No. 2).

Boughner, R. L., & Papini, M. R. (2003). Appetitive latent inhibition in rats: Now you see it (sign tracking), now you don't (goal tracking). *Animal Learning & Behavior*, *31*(4), 387–392.

Bouton, M. E. (1984). Differential control by context in the inflation and reinstatement paradigms. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*(1), 56–74.

Bouton, M. E. (2004). Context and Behavioral Processes in Extinction. *Learning & Memory*, *11*(5), 485–494.

Brandon, S. E., Vogel, E. H., & Wagner, A. R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behavioural*

Brain Research, *110*(1-2), 67–72.

Brimer, C. J. (1970). Inhibition and disinhibition of an operant response as a function of the amount and type of prior training. *Psychonomic Science*, *21*(3), 191–192.

Brimer, C. J., & Kamin, L. J. (1963). Disinhibiton, habituation, sensitization, and the conditioned emotional response. *Journal of Comparative and Physiological Psychology*, *56*(3), 508–516.

Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, *11*(1), 1–8.

Browne, M. P. (1976). *The role of primary reinforcement and overt movements in autoshaping in the pigeon* (Vol. 4; Tech. Rep. No. 3).

Bush, R. R., & Mosteller, F. (1951). A model for stimulus generalization and discrimination. *Psychological Review*, *58*(6), 413–423.

Cinotti, F., Marchand, A. R., Roesch, M. R., Girard, B., & Khamassi, M. (2019). Impacts of inter-trial interval duration on a computational model of sign-tracking vs. goal-tracking behaviour. *Psychopharmacology*, 2373–2388.

Cohen-Hatton, S. R., Haddon, J. E., George, D. N., & Honey, R. C. (2013). Pavlovian-to-instrumental transfer: Paradoxical effects of the Pavlovian relationship explained. *Journal of Experimental Psychology: Animal Behavior Processes*, *39*(1), 14–23.

Cole, R. P., & Miller, R. R. (1999). Conditioned Excitation and Conditioned Inhibition Acquired through Backward Conditioning. *Learning and Motivation*, *30*(2), 129–156.

Costa, D. S., & Boakes, R. A. (2009). Context blocking in rat autoshaping: Sign-tracking versus goal-tracking. *Learning and Motivation*, *40*(2), 178–185.

Cotton, M. M., Goodall, G., & Mackintosh, N. J. (1982). Inhibitory conditioning resulting from a reduction in the magnitude of reinforcement. *The Quarterly journal of experimental psychology. B, Comparative and physiological psychology*, *34 (Pt 3)*, 163–80.

Crowell, C. R., & Bernhardt, T. P. (1979). The feature-positive effect and sign-tracking behavior during discrimination learning in the rat. *Animal Learning & Behavior*, *7*(3), 313–317.

Daly, P. W., & Ogawa, A. (2013). *package 'natbib': Flexible bibliography support.* Retrieved from `https://ctan.org/pkg/natbib?lang=en`

Davey, G. C. L., & Cleland, G. G. (1982). *The effect of partial reinforcement on the acquisition and extinction of sign-tracking and goal-tracking in the rat* (Vol. 19; Tech. Rep. No. 2).

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–15.

Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. *Cognitive, affective & behavioral neuroscience*, *14*(2), 473–92.

Delamater, A. R. (2012). On the nature of CS and US representations in Pavlovian learning. *Learning & Behavior*, *40*(1), 1–23.

Derman, R. C., Schneider, K., Juarez, S., & Delamater, A. R. (2018). Sign-tracking is an expectancy-mediated behavior that relies on prediction error mechanisms. *Learning & Memory*, *25*(10), 550–563.

Dickinson, A., & Balleine, B. W. (2002a). The role of learning in motivation. In C. Gallistel (Ed.), *Stevens' handbook of experimental psychology. vol. 3.* (pp. 497–533). Wiley.

Dickinson, A., & Balleine, B. W. (2002b). The Role of Learning in the Operation of Motivational Systems. In *Stevens' handbook of experimental psychology.* Hoboken, NJ, USA: John Wiley & Sons, Inc.

Dickinson, A., Hall, G., & Mackintosh, N. J. (1976). Surprise and the attenuation of blocking. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*(4), 313–322.

Duvarci, S., Nader, K., & LeDoux, J. E. (2008). De novo mRNA synthesis is required for both consolidation and reconsolidation of fear memories in the amygdala. *Learning & Memory*, *15*(10), 747–755.

Dwyer, D. M. (2012). Licking and Liking: The Assessment of Hedonic Responses in Rodents. *Quarterly Journal of Experimental Psychology*, *65*(3), 371–394.

Dwyer, D. M., Burgess, K. V., & Honey, R. C. (2012). Avoidance but not aversion following sensory preconditioning with flavors: A challenge to stimulus substitution. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*(4), 359–368.

Dwyer, D. M., Figueroa, J., Gasalla, P., & López, M. (2018). Reward Adaptation and the Mechanisms of Learning: Contrast Changes Reward Value in Rats and Drives Learning. *Psychological Science*, *29*(2), 219–227.

Dwyer, D. M., & Mackintosh, N. J. (2002). Alternating exposure to two compound flavors creates inhibitory associations between their unique features. *Animal learning & behavior*, *30*(3), 201–7.

Engberg, L. A., Hansen, G., Welker, R. L., & Thomas, D. R. (1972). Acquisition of Key-Pecking via Autoshaping as a Function of Prior Experience: "Learned Laziness"? *Science*, *178*(4064), 1002–1004.

Fitzpatrick, C. J., Gopalakrishnan, S., Cogan, E. S., Yager, L. M., Meyer, P. J., Lovic, V., … Morrow, J. D. (2013). Variation in the Form of Pavlovian Conditioned Approach Behavior among Outbred Male Sprague-Dawley Rats from Different Vendors and Colonies: Sign-Tracking vs. Goal-Tracking. *PLoS ONE*, *8*(10), e75042.

Flagel, S. B., Akil, H., & Robinson, T. E. (2009). Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology*, *56*(SUPPL. 1), 139–148.

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., … Akil, H. (2011). A selective role for dopamine in stimulus–reward learning. *Nature*, *469*(7328), 53–57.

Flanagan, B., & Webb, W. B. (1964). Disinhibition and external inhibition in fixed interval operant conditioning. *Psychonomic Science*, *1*(1-12), 123–124.

Gallistel, C. R. (1990). *The organization of learning.* Cambridge, MA: MIT Press.

Gallistel, C. R., & Gibbon, J. (2000). Time, Rate and Conditioning. *Psychological Review*, *107*(2), 289–344.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*(3), 227–247.

Gokey, D. S., & Collins, R. L. (1980). Conditioned inhibition in feature negative discrimination learning with pigeons. *Animal Learning & Behavior*, *8*(2), 231–236.

Gottlieb, D. A. (2004). Acquisition with partial and continuous reinforcement in pigeon autoshaping. *Learning & behavior*, *32*(3), 321–34.

Gottlieb, D. A. (2006). Effects of partial reinforcement and time between reinforced trials on terminal response rate in pigeon autoshaping. *Behavioural Processes*, *72*(1), 6–13.

Gromer, D. (2019). *package 'apa': Format Outputs of Statistical Tests According to APA Guidelines.* Retrieved from https://cran.r-project.org/web/packages/apa/index.html

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review*, *87*, 1–51.

Hall, G. (1991). *Perceptual and Associative Learning.* Oxford, UK: Oxford University Press.

Hall, G., & Honey, R. C. (1989). Contextual effects in conditioning, latent inhibition, and habituation: Associative and retrieval functions of contextual cues. *Journal of Experimental Psychology: Animal Behavior Processes*, *15*(3), 232–241.

Hall, G., & Minor, H. (1984). A Search for Context-Stimulus Associations in Latent Inhibition. *The Quarterly Journal of Experimental Psychology Section B*, *36*(2b), 145–169.

Hearst, E., & Franklin, S. R. (1977). Positive and negative relations between a signal and food: Approach-withdrawal behavior to the signal. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*(1), 37–52.

Hearst, E., & Jenkins, H. M. (1974). *Sign tracking: The stimulus-reinforcer relation and directed action.* Austin: TX: Psychonomic Society.

Hinrichs, J. V. (1968). Disinhibition of delay in fixed-interval instrumental conditioning. *Psychonomic Science*, *12*(7), 313–314.

Holland, P. C. (1977). Conditioned stimulus as a determinant of the form of the Pavlovian conditioned response. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*(1), 77–104.

Holland, P. C. (1984). Origins of Behavior in Pavlovian Conditioning. *Psychology of Learning and Motivation*, *18*, 129–174.

Holland, P. C. (1988). Excitation and inhibition in unblocking. *Journal of experimental psychology. Animal behavior processes*, *14*(3), 261–79.

Holmes, N. M., Chan, Y. Y., & Westbrook, R. F. (2019). A combination of common and individual error terms is not needed to explain associative changes when cues with different training histories are conditioned in compound: A review of Rescorla's compound test procedure. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(2), 242–256.

Honey, R. C., & Bolhuis, J. J. (1997). Imprinting, Conditioning, and Within-event Learning:. *Quantitative analyses of behavior*, *50*(2), 97–110.

Honey, R. C., & Good, M. (1993). Selective hippocampal lesions abolish the contextual specificity of latent inhibition and conditioning. *Behavioral neuroscience*, *107*(1), 23–33.

Honey, R. C., Iordanova, M. D., & Good, M. (2010). Latent inhibition and habituation: evaluation of an associative analysis. In R. E. Lubow & I. Weiner (Eds.), *Latent inhibition* (pp. 163–182). Cambridge: Cambridge University Press.

Honey, R. C., Iordanova, M. D., & Good, M. (2014). Associative structures in animal learning: Dissociating elemental and configural processes. *Neurobiology of*

Learning and Memory, *108*, 96–103.

Honey, R. C., & Ward-Robinson, J. (2002). Acquired equivalence and distinctivenes of cues: I. Exploring a neural network approach. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(4), 378–387.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton-Century-Crofts.

Hull, C. L. (1949). Stimulus intensity dynamism (V) and stimulus generalization. *Psychological Review*, *56*(2), 67–76.

JASP Team. (2018). *JASP [Computer software].*

Jenkins, H. M., & Moore, B. R. (1973). The form of the auto-shaped response with food or water reinforcers. *Journal of the experimental analysis of behavior*, *20*(2), 163–81.

Jenkins, H. M., & Sainsbury, R. S. (1970). Discrimination learning with the distinctive feature on positive and negative trials. In D. Mostofsky (Ed.), *Attention: Contemporary theory and analysis* (pp. 239–273). New York: Appleton-Century-Crofts.

Kamin, L. J. (1969). Selective association and conditioning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 42–89). Halifax: Dalhousie University Press.

Kaye, H., & Pearce, J. M. (1984). The strength of the orienting response during Pavlovian conditioning. *Journal of experimental psychology. Animal behavior processes*, *10*(1), 90–109.

Konorski, J. (1948). *Conditioned reflexes and neuron organization.* Cambridge: Cambridge University Press.

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

Kruschke, J. K. (1992). ALCOVE: an exemplar-based connectionist model of category learning. *Psychological review*, *99*(1), 22–44.

Lawrence, M. A. (2016). *package 'ez': Easy Analysis and Visualization of Factorial Experiments.* Retrieved from https://cran.r-project.org/web/packages/ez/

index.html

Lee, B., Gentry, R. N., Bissonette, G. B., Herman, R. J., Mallon, J. J., Bryden, D. W., … Roesch, M. R. (2018). Manipulating the revision of reward value during the intertrial interval increases sign tracking and dopamine release. *PLOS Biology*, *16*(9), e2004015.

Lesaint, F., Sigaud, O., Flagel, S. B., Robinson, T. E., & Khamassi, M. (2014). Modelling Individual Differences in the Form of Pavlovian Conditioned Approach Responses: A Dual Learning Systems Approach with Factored Representations. *PLoS Comput Biol*, *10*(2).

Leslie, J. C., & Millenson, J. R. (1996). *Principles of behavioral analysis.* Amsterdam, The Netherlands: Harwood Academic Publishers.

Lin, T. E., & Honey, R. C. (2016). Learning About Stimuli That Are Present and Those That Are Not. In *The wiley handbook on the cognitive neuroscience of learning* (pp. 69–85). Chichester, UK: John Wiley & Sons, Ltd.

Lonsdorf, T. B., Weike, A. I., Nikamo, P., Schalling, M., Hamm, A. O., & Öhman, A. (2009). Genetic Gating of Human Fear Learning and Extinction. *Psychological Science*, *20*(2), 198–206.

Lotz, A., Uengoer, M., Koenig, S., Pearce, J. M., & Lachnit, H. (2012). An exploration of the feature-positive effect in adult humans. *Learning & Behavior*, *40*(2), 222–230.

Lovic, V., Saunders, B. T., Yager, L. M., & Robinson, T. E. (2011). Rats prone to attribute incentive salience to reward cues are also prone to impulsive action. *Behavioural Brain Research*, *223*(2), 255–261.

Lubow, R. E. (1989). *Latent inhibition and conditioned attention theory.* Cambridge: Cambridge University Press.

Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: The effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of Comparative and Physiological Psychology*, *52*(4), 415–419.

Mackintosh, N. J. (1974). *The psychology of animal learning.* London: Academic Press.

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

Mackintosh, N. J. (1976). Overshadowing and stimulus intensity. *Animal Learning & Behavior*, *4*(2), 186–192.

Maes, E., Boddez, Y., Alfei, J. M., Krypotos, A. M., D'Hooge, R., De Houwer, J., & Beckers, T. (2016). The elusive nature of the blocking effect: 15 failures to replicate. *Journal of Experimental Psychology: General*, *145*(9), e49–e71.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* Cambridge, MA: MIT Press.

Matzel, L. D., Han, Y. R., Grossman, H., Karnik, M. S., Patel, D., Scott, N., ... Gandhi, C. C. (2003). Individual differences in the expression of a "general" learning ability in mice. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *23*(16), 6423–33.

Matzel, L. D., & Sauce, B. (2017). Individual differences: Case studies of rodent and primate intelligence. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*(4), 325–340.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375–407.

McLaren, I. P. L., Kaye, H., & Mackintosh, N. J. (1989). An associative theory of the representation of stimuli: applications to perceptual learning and latent inhibition. In R. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neurobiology* (pp. 102–130). Oxford: Clarendon Press.

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

Meyer, P. J., Lovic, V., Saunders, B. T., Yager, L. M., Flagel, S. B., Morrow, J. D., & Robinson, T. E. (2012). Quantifying individual variation in the propensity to attribute incentive salience to reward cues. *PLoS ONE*, *7*(6), e38987.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological bulletin*, *117*(3), 363–86.

Miller, R. R., & Matute, H. (1998). Competition Between Outcomes. *Psychological Science*, *9*(2), 146–149.

Miller, R. R., & Matzel, L. D. (1988). The Comparator Hypothesis: A Response Rule for The Expression of Associations. *Psychology of Learning and Motivation*, *22*, 51–92.

Morrison, S. E., Bamkole, M. A., & Nicola, S. M. (2015). Sign tracking, but not goal tracking, is resistant to outcome devaluation. *Frontiers in Neuroscience*, *9*(DEC), 1–12.

Mundy, M. E., Dwyer, D. M., & Honey, R. C. (2006). Inhibitory associations contribute to perceptual learning in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*(2), 178–184.

Murphy, R. A., & Honey, R. C. (2016). *The Wiley Handbook on the Cognitive Neuroscience of Learning* (R. A. Murphy & R. C. Honey, Eds.). Chichester, UK: John Wiley & Sons, Ltd.

Ostlund, S. B., & Balleine, B. W. (2008). On habits and addiction: An associative analysis of compulsive drug seeking. *Drug discovery today. Disease models*, *5*(4), 235–245.

Patitucci, E., Nelson, A. J. D., Dwyer, D. M., & Honey, R. C. (2016). The Origins of Individual Differences in How Learning Is Expressed in Rats: A General-Process Perspective. *J Exp Psychol Anim Learn Cogn*, *42*(4), 313–324.

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

Pavlov, I. P. (1930). A brief outline of the higher nervous activity. In C. Murchison (Ed.), *Psychologies of the 1930* (pp. 207–220). Worcester, MA: Clark Universities Press.

Pavlov, I. P. (1941). *Conditioned reflexes and psychiatry.* New York, NY: International

Publishers.

Pearce, J. M. (1994). Similarity and discrimination: a selective review and a connectionist model. *Psychological review*, *101*(4), 587–607.

Pearce, J. M., Aydin, A., & Redhead, E. S. (1997). Configural analysis of summation in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*(1), 84–94.

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

Pearce, J. M., & Mackintosh, N. J. (2010). Two theories of attention: A review and possible integration. In C. Mitchel & M. E. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 11–39). Oxford: Oxford University Press.

Pearce, J. M., Redhead, E. S., & George, D. N. (2002). Summation in autoshaping is affected by the similarity of the visual stimuli to the stimulation they replace. *Journal of experimental psychology. Animal behavior processes*, *28*(2), 175–89.

Pena-Oliver, Y., Giuliano, C., Economidou, D., Goodlett, C. R., Robbins, T. W., Dalley, J. W., & Everitt, B. J. (2015). Alcohol-Preferring Rats Show Goal Oriented Behaviour to Food Incentives but Are Neither Sign-Trackers Nor Impulsive. *PLoS ONE*, *10*(5), e0131016.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of comparative and physiological psychology*, *66*(1), 1–5.

Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, *72*(2), 77–94.

Rescorla, R. A. (1971). Variation in the effectiveness of reinforcement and nonreinforcement following prior inhibitory conditioning. *Learning and Motivation*, *2*(2), 113–123.

Rescorla, R. A. (1973). Effect of US habituation following conditioning. *Journal of* *comparative and physiological psychology*, *82*(1), 137–43.

Rescorla, R. A. (1988). Pavlovian conditioning. It's not what you think it is. *The American psychologist*, *43*(3), 151–60.

Rescorla, R. A. (2000). Associative changes in excitors and inhibitors differ when they are conditioned in compound. *Journal of experimental psychology. Animal behavior processes*, *26*(4), 428–38.

Rescorla, R. A. (2001). Unequal associative changes when excitors and neutral stimuli are conditioned in compound. *The Quarterly journal of experimental psychology. B, Comparative and physiological psychology*, *54*(1), 53–68.

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

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

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In B. A.H. & W. Prokasy (Eds.), *Clasical conditioning ii:current research and theory* (pp. 64–69). New York: Appleton-Century-Crofts.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374.

Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E. (2017). Bayesian analysis of factorial designs. *Psychological Methods*, *22*(2), 304–321.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237.

RStudio, T. (2015). *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, Inc. Retrieved from `https://rstudio.com`

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

Schultz, W. (2006). Behavioral Theories and the Neurophysiology of Reward. *Annual Review of Psychology*, *57*(1), 87–115.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*, *275*(5306), 1593–9.

Shettleworth, S. J. (1972). Constraints on Learning. In D. S. Lehrman, R. A. Hinde, & E. Shaw (Eds.), *Advances in the study of behavior* (pp. 1–68). Academic Press.

Skinner, B. F. (1948). "Superstition" in the pigeon. *Journal of Experimental Psychology*, *38*(2), 168–172.

Spence, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review*, *43*(5), 427–449.

Spence, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, *44*(5), 430–444.

Stout, S. C., & Miller, R. R. (2007). Sometimes-competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review*, *114*(3), 759–783.

Terrace, H. S., Gibbon, J., Farrell, L., & Baldock, M. D. (1975). Temporal factors influencing the acquisition and maintenance of an autoshaped keypeck. *Animal Learning & Behavior*, *3*(1), 53–62.

Thorndike, E. L. (1927). The Law of Effect. *The American Journal of Psychology*, *39*(1/4), 212.

Timberlake, W. (1994). Behavior systems, associationism, and Pavlovian conditioning. *Psychonomic Bulletin & Review*, *1*(4), 405–420.

Timberlake, W., Wahl, G., & King, D. A. (1982). Stimulus and response contingencies in the misbehavior of rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *8*(1), 62–85.

Urcelay, G. P. (2017). Competition and facilitation in compound conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*(4), 303–314.

Volkow, N. D., Fowler, J. S., Wang, G., & Swanson, J. M. (2004). Feature review: Dopamine in drug abuse and addiction: results from imaging studies and treatment implications. *Molecular Psychiatry*, *9*, 557–569.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–148). Hillsdale, N.J.: Erlbaum.

Wagner, A. R., & Brandon, S. E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (ÆSOP). In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theories* (pp. 149–189). Hillsdale, N.J.: Erlbaum.

Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology*, *64*, 191–199.

Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: Applications of a theory. In R. A. Boakes & M. S. Haliday (Eds.), *Inhibition and learning* (pp. 301–336). New York: Academic Press.

Wasserman, E. A. (1974). Stimulus-reinforcer predictiveness and selective discrimination learning in pigeons. *Journal of Experimental Psychology*, *103*(2), 284–297.

Wasserman, E. A., Franklin, S. R., & Hearst, E. (1974). Pavlovian appetitive contingencies and approach versus withdrawal to conditioned stimuli in pigeons. *Journal of comparative and physiological psychology*, *86*(4), 616–27.

Watson, J. B. (1924). *Behaviorism.* New Yoek: People's Institute Publishing Company.

Watt, A., & Honey, R. C. (1997). Combining CSs associated with the same or different USs. *The Quarterly journal of experimental psychology. B, Comparative and physiological psychology*, *50*(4), 350–67.

Wickham, H., Chang, W., Henry, L., Lin Pederse, T., Takashashi, K., Wike, C., & Woo, K. (2018). *package 'ggplot2': Create Elegant Data Visualisations Using the Grammar of Graphics.* Retrieved from https://ggplot2.tidyverse.org

Wickham, H., & Henry, L. (2018). *package 'tidyr': Easily Tidy Data with 'spread()' and 'gather()' Functions.* Retrieved from https://cran.r-project.org/web/packages/tidyr/index.html

Wickham, H., Henry, L., & Muller, K. (2019). *package 'dplyr': A grammar of data manipulation.* Retrieved from https://cran.r-project.org/web/packages/dplyr/index.html

Williams, D. R., & Williams, H. (1969). Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *Journal of the experimental analysis of behavior*, *12*(4), 511–20.

Wilson, P. N., Boumphrey, P., & Pearce, J. M. (1992). Restoration of the Orienting Response to a Light by a Change in its Predictive Accuracy:. *Quarterly Journal of Experimental Psychology*, *44*(B), 17–36.

Zimmer-Hart, C. L., & Rescorla, R. A. (1974). Extinction of Pavlovian conditioned inhibition. *Journal of comparative and physiological psychology*, *86*(5), 837–45.