THE DISCRIMINATION OF MAGNITUDE

This thesis is presented for the degree of Doctor of Philosophy

2015

Richard A. Inman

DECLARATION

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

Signed …… (candidate)

Date ……

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of ……(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed …… (candidate)

Date ……

STATEMENT 2

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

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

Signed …… (candidate)

Date ……

STATEMENT 3

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

Signed …… (candidate)

Date ……

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

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

Signed …… (candidate)

Date ……

Large parts of Chapter 3 and Chapter 4 appear in:

Inman, R.A., Honey, R.C., Eccles, G.L., & Pearce, J.M. (2015). Asymmetry in the discrimination of quantity by rats: The role of the intertrial interval. *Learning & Behavior*. Advance online publication. DOI: 10.3758/s13420-015-0191-0

Inman, R. A., Honey, R.C., & Pearce, J. M. (2015). Asymmetry in the discrimination of magnitude: The role of stimulus generalisation. *Journal of Experimental Psychology: Animal Learning and Cognition.* Advance online publication. DOI: 10.1037/xan0000073

Acknowledgements

First I would like to express my gratitude to Professor John Pearce whose invaluable guidance and continuous enthusiasm made this thesis possible. I am incredibly lucky to have had you as a supervisor. I would also like to say thank you to Professor Rob Honey who has been incredibly supportive as a secondary supervisor.

I would like to thank everybody who has helped to make my PhD experience immensely enjoyable. Thank you to Becky Wright, Natasha Dumigan, Luke Montuori, Brice Dassy, and many other friends for tea breaks, dinners, interesting debate and constant comradeship throughout the last four years. Cheers to Clive Mann for being incredibly helpful down in the BNL. Thank you also to Bill Macken, Tom Freeman, Mark Good, and all other members of M56, past and present, with whom I have enjoyed performing and practicing with tremendously.

Mum and Dad, you have been incredibly patient and understanding when I needed it and your continuous support has helped me to strive to be the best I can be. Thank you so much for having been there for me always.

Finally, I want to say a huge heartfelt thank you to Sara whose smiles and laughter continue to brighten up my life. Words simply cannot express… Obrigado por seres simplesmente tu.

Summary

A number of theories of conditioning are based on the assumption that generalisation gradients either side of a stimulus are symmetrical. From this assumption, the prediction follows that the rate at which a discrimination is acquired between two stimuli will be unaffected by which of them signals the reinforcer (S+). In contrast to this prediction, when S+ and a non-reinforced stimulus (S-) are different in terms of their relative position on a magnitude continuum, this symmetry does not always hold. In many cases, discriminations with a high-magnitude S+ and low-magnitude S- are acquired more readily than when S+ is low-magnitude and S- is high-magnitude. The primary aim of this thesis was therefore to offer an account for this asymmetry. Chapter 2 presents evidence for such an asymmetry with clickers differing in intensity. Chapters 3 and 4 present an asymmetry with arrays differing in the number of black squares on a white background, although it emerged that the mechanisms responsible for this effect were different than for auditory intensity. One possibility, argued in Chapter 5, is that these modalities present contrasting results due to differences in how the stimuli are represented at the receptor level. With this considered, in an attempt to account for both sets of data in terms of theories of conditioning, an adaptation to Pearce´s (1987) configural theory is proposed whereby similarity between stimuli is based on the proportion of common elements, rather than the number of common elements. It is further argued that this amendment provides a more satisfactory account of the reported results than the original theory of Pearce. In addition, the amended theory is shown to account for a wider range of results than the Rescorla-Wagner (1972) model.

Contents

CHAPTER 1

General Introduction

After an animal has been trained to respond to a stimulus in a particular way, perhaps by moving towards a food magazine during a tone that signals the delivery of food, it will also make the same response, but less frequently, when the characteristics of the tone are altered. Few would deny that the similarity between the two tones exerts a fundamental influence on the transfer of responding from one to the other (see Shepard, 1987). Moreover, this relationship is generally regarded to be symmetrical, so that the similarity of tone A to tone B is considered as being the same as the similarity of tone B to tone A. On this basis it follows that the symmetry in stimulus similarity will also result in a symmetry in stimulus generalisation. Thus the strength of response elicited by tone B after training with tone A should be the same as the strength of response elicited by tone A after training with tone B. These principles are encapsulated in a variety of different theories of associative learning that have been, or are currently influential. In view of this theoretical consensus, it is surprising to note that animals do not always behave in the way that would be expected if the similarity between two stimuli is always symmetrical. The purpose of the present thesis is to explore the implications of these anomalous results for our theoretical understanding of associative learning.

By way of providing a background to the research that will be described in later chapters, the present chapter will first review different theoretical accounts of the mechanisms that permit similarity to exert its influence on behaviour. The next section will then review evidence that supports the claims above concerning the importance of similarity on transfer of responding from one stimulus to another, and that supports the claim that this influence is symmetrical. The final section of the chapter will then review the evidence showing that similarity relationship is not

always symmetrical, and review the explanations that have been offered for this outcome.

Theories of Discrimination Learning

A wide range of theoretical accounts have been proposed in order to account for the mechanisms behind stimulus generalisation and discrimination learning. For the sake of brevity the following section will review just three. These theories were chosen because of their relevance to discussion in later chapters. The first, Spence's (1936, 1937) gradient interaction theory, represents one of the earlier considerations of discrimination learning. The remaining two theories, the Rescorla-Wagner (1972) model and Pearce's (1987, 1994) configural theory, represent contemporary and highly influential alternatives. All three will be considered in terms of how they either predict or assume symmetrical stimulus generalisation and the impact this has on discrimination learning.

Gradient Interaction (Spence)

Stimulus Generalisation

According to Spence (1936, 1937) when subjects are given training in which a stimulus, e.g. tone A, is consistently paired with a reward such as food, they will acquire an excitatory tendency to approach and to respond to the stimulus. Importantly, during conditioning subjects learn about the absolute physical properties of the stimulus. Thus any excitatory tendencies to approach a stimulus will generalise to other similar stimuli that differ in these physical properties. Moreover, the amount of generalisation that occurs, and the subsequent level of responding to a new stimulus presented after training, will depend on the degree of

similarity to the trained stimulus. The solid curve in Figure 1 presents an example of a hypothetical excitatory generalisation gradient that has formed around a reinforced stimulus, S+, where the height of the curve corresponds to the level of responding to stimuli that differ to S+ on the same dimension. In order to account for the Weber-Fechner relationship[1] between sensory and stimulus dimensions, Spence further assumed that similarity is a logarithmic function of the dimension under scrutiny (see Figure 1). As a result, the amount of generalisation peaks at the training stimulus and then decreases symmetrically either side. An increase or decrease of equal distance away from the S+ will thus result in the same decrease in responding regardless of the direction of change. As an aside, it should be noted that the functions displayed in Figure 1 are parabolic. Spence later modified his theory to assume that generalisation gradients were instead bell-shaped Gaussian functions (1942). The predictions made by the theory, however, remain the same.

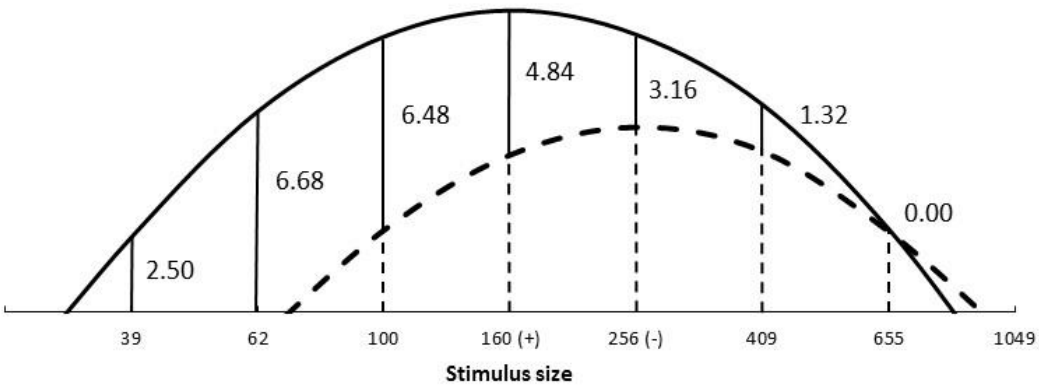

Figure 1. Hypothetical symmetrical generalisation gradients that form around S+ and S- during a simple discrimination.

[1] The Weber-Fechner law describes how the perceivable difference between stimuli decreases as stimulus magnitude increases. Thus at higher magnitudes a greater absolute difference between stimuli is required to equate to the same degree of perceived similarity than at lower magnitudes.

Discrimination Learning

In a discrimination between two stimuli, subjects are required to differentiate between a stimulus that signals an outcome (S+) and a second stimulus that does not signal the outcome (S-). Over successive training trials animals come to respond in the presence of S+ and inhibit their responses in the presence of the S-. The dashed line in Figure 1 demonstrates that the underlying gradients of inhibition that form around an S- are also symmetrical. This gradient reflects the degree to which animals will inhibit responding in the presence of similar stimuli. What is also clear from Figure 1 is that the total level of responding to the S+ when trained alongside S- will be determined by the algebraic sum of the excitatory and inhibitory gradients. Thus in a discrimination where S+ and S- are highly similar, the large amount of gradient overlap will result in a weaker response to S+ than if S+ and S- were highly dissimilar. Furthermore, because the generalisation gradients are assumed to be symmetrical around stimuli, in a discrimination task such as A+ B- the interaction between the two gradients and resulting level of responding to S+ will be the same as for the discrimination B+A-.

Spence's (1936, 1937) gradient interaction theory has been used to great effect to explain phenomena such as peak-shift (Hanson, 1959) which involves training with stimuli from a single dimension. The theory struggles when it comes to explaining phenomena that result from training with stimuli from two or more dimensions. A classic example is blocking (Kamin, 1969). A blocking design begins with the pairing of one stimulus with an outcome, A+, followed by further training in which a second stimulus is paired with the first, AB+. The second stimulus, B, is then tested for a conditioned response (CR). Kamin found that after this blocking

treatment stimulus B elicited a weaker CR than for a second non-blocking group of rats not given the initial A+ training. Spence's theory is unable to account for this finding because of its non-selective nature. Thus exposure to B should elicit the same strength of CR, regardless of whether or not the rats initially experienced training with A.

The Rescorla-Wagner Model

A highly influential theory of conditioning that can account for blocking is the Rescorla-Wagner (1972) model. According to this theory, the strength of a CR depends on the strength of the association between a conditioned stimulus (CS) and the outcome, or unconditioned stimulus (US). The change in the strength of this association that occurs on each conditioning trial can be described by Equation 1.

$$\Delta V_A = \alpha \times \beta \times (\lambda - V_T) \quad (1)$$

On any given trial in which a CS (in this instance stimulus A) is paired with a US, the change in associative strength, ΔV_A, is directly linked to the difference between an asymptotic level of associative strength, which is set by the magnitude of the US, λ, and the combined associative strength of all the stimuli presented during the trial, V_T. This value is then modified by two learning rate parameters; α, which is dictated by the salience of the CS, and β, which is dictated by the salience of the US. The values of these constants lie between 0 and 1. The fact that this model takes account of the associative strength of all the stimuli presented in a trial has made this theory successful in explaining a number of phenomena that could not be accounted for by Spence. Referring back to blocking, the Rescorla-Wagner (1972) model correctly predicts that the strength of CR to B should be less for the blocking group

than the non-blocking group. During initial training with A+, stimulus A enters into an excitatory association with the outcome and with extended training this level reaches asymptote. Consequently, when element B is added for AB+ training, B cannot acquire any excitatory association of its own because according to Equation 1, when the combined associative strength of the stimuli, V_T, is already at asymptote, λ, because of previous training with A, then the change in associative strength must be zero. In contrast, for the non-blocking group, the absence of previous A+ training allows both A and B to acquire associative strength during AB+ training.

Stimulus Generalisation

The most detailed account of how the Rescorla-Wagner (1972) model can be used to explain stimulus generalisation has been provided by Blough (1975). According to Blough a stimulus can be represented by the activations it elicits in a number of constituent elements. Each of these elements is able to acquire associative strength and thus the total amount of responding to a S+ after conditioning is determined by the sum of the associative strengths of its constituent elements. Moreover, if a novel stimulus is presented after training it will also activate some of the constituent elements of the S+ and the more dissimilar the novel stimulus is to the S+, the fewer common elements it activates. Importantly, the relationship between similarity and the number of activated common elements is Gaussian and therefore the strength of a CR to a novel stimulus as a function of its similarity can be described by a normal distribution around the S+.

Discrimination Learning

The Rescorla-Wagner (1972) model also makes the prediction that an A+B- discrimination should be acquired as readily as a B+A- discrimination. According to Equation 1, for the A+B- discrimination stimulus A will gradually acquire associative strength until it reaches asymptote, whilst stimulus B, which is never paired with the US, will not gain any associative strength. Exactly the same prediction is made for the B+A- discrimination assuming that the constants α and β remain the same. Of course, it is unlikely that two stimuli share nothing in common, and thus a more accurate representation of stimuli might include one or more common elements. For example, we might represent a high and a low frequency tone as comprising unique elements A and B, and a common element, X (e.g. AX+BX- and BX+AX-). Even so, if we represent the discriminations in this way the Rescorla-Wagner model continues to predict that the two stimuli should be differentiated at the same rate regardless of which stimulus is assigned as S+.

Stimulus Similarity (Pearce)

The Rescorla -Wagner (1972) model can be classified as an elemental theory because of its assumption that the individual stimuli in a compound and also the constituent elements of each stimulus (Blough, 1975), can separately enter into associations with the reinforcer during conditioning. A contrasting class of configural theories postulate that patterns of stimulation as a whole enter into association with the reinforcer (e.g. Friedman & Gelfrand, 1964; Gulliksen & Wolfle, 1938; Pearce 1987, 1994). Equation 2 shows the way, according to Pearce, in which conditioning to a pattern (in this instance AX) will progress on any given conditioning trial during an AX+BX- discrimination.

$$\Delta V_{AX} = \beta \times (\lambda - (V_{AX} + {}_{AX}S_{BX} \times V_{BX})) \quad (2)$$

The equation is similar to the Rescorla-Wagner model (Equation 1) in the sense that the change in associative strength to AX, ΔV_{AX}, is determined by the difference between an asymptote in learning, λ, and what has already been learned, multiplied by a learning rate parameter, β. What differentiates the two theories is that rather than simply being determined by the difference between asymptote and the sum of the associative strength of all stimuli present on the trial, Pearce's (1987, 1994) configural theory also considers the associative strength of other stimuli based on their similarity to CS+, ${}_{AX}S_{BX}$. Thus the change in associative strength for a stimulus on any given trial is the discrepancy between asymptote and the associative strength of the stimulus plus the associative strength that generalises to this stimulus from similar stimuli. This generalisation coefficient can be derived from Equation 3 where N_C refers to the number of elements common to both AX and BX, whilst N_{AX} and N_{BX} refer to the number of elements in AX and BX respectively.

$${}_{AX}S_{BX} = N_C^2/(N_{AX} \times N_{BX}) \quad (3)$$

Stimulus Generalisation and Discrimination Learning

Pearce's (1987, 1994) model is based on the assumption that the amount of generalisation between two patterns of stimulation is dependent on their similarity. Moreover, it follows from Equation 3 that generalisation between two compounds will be symmetrical because the similarity of, say, AX to BX is the same as BX to AX. As a result, Equation 2 readily predicts that AX+BX- discriminations will be acquired at the same rate as BX+AX- discriminations. Pearce, however, offers no account of stimulus generalisation based on a single stimulus. Nonetheless, if

Pearce's model adopts a logic similar to that of Blough's (1975) proposal that there is a Gaussian relationship between stimulus similarity and the number of activated common elements, then it is able to predict symmetrical generalisation gradients. Consider a case where conditioning with A+ is followed by a test in which stimulus B is presented. During this test the pattern of stimulation activated by B will also include some elements that are common to A. These elements can be represented as X. The level of CR elicited by B will then, according to Equation 3, depend on the number of common elements it shares with A – a greater number will result in a stronger CR. As B becomes more dissimilar to A the level of shared activation decreases in a Gaussian fashion and thus so too does the strength of CR. In this manner Pearce's theory predicts a symmetrical decrease in responding either side of the CS (in terms of position on a stimulus dimension). As the similarity of test stimuli increases or decreases away from the CS, the number of common elements shared by the two stimuli also decreases.

Evidence of symmetry

Given that all three of the aforementioned theories of conditioning assume that the decrease in CR as a function of similarity is symmetrical either side of the S+, one might expect there to be a large amount of evidence in support of this shared assumption. Such evidence could be identified from at least two sources. First, initial training with an S+ followed by generalisation tests with novel stimuli should result in symmetrical generalisation gradients. Second, when required to discriminate between two stimuli the task should be acquired at the same rate regardless of which stimulus signals the reinforcer. However, as it will become clear, there is

surprisingly little evidence to support the second of these predictions, and some that contradicts the first.

Symmetrical Generalisation Gradients

A number of studies appear to suggest that responding to a test stimulus decreases as it becomes more dissimilar from an initial training stimulus. Importantly, there is evidence that this decrease in responding as a function of similarity is symmetrical. This point is illustrated well by Guttman and Kalish (1956). In this experiment four groups of pigeons were trained to peck at an illuminated response key in order to obtain food. These groups were differentiated by the wavelength of light used to illuminate the key (see Figure 2). Once a steady rate of pecking had been established, all pigeons were given test trials in which they were given multiple presentations of 11 different wavelengths of light, all trials of which were conducted in extinction. In support of the claim that generalisation gradients are symmetrical, the number of responses declined symmetrically as the wavelength became more dissimilar from the trained stimulus. In other words, the slope of the gradient was the same when the wavelength of the test stimulus decreased and increased away from the trained stimulus. Similar symmetrical gradients have been found with a variety of stimuli and species, for example Moore (1972), who gave rabbits eye-blink conditioning with a 1200-Hz tone followed by a generalisation test with tones ranging between 400 Hz and 2000 Hz.

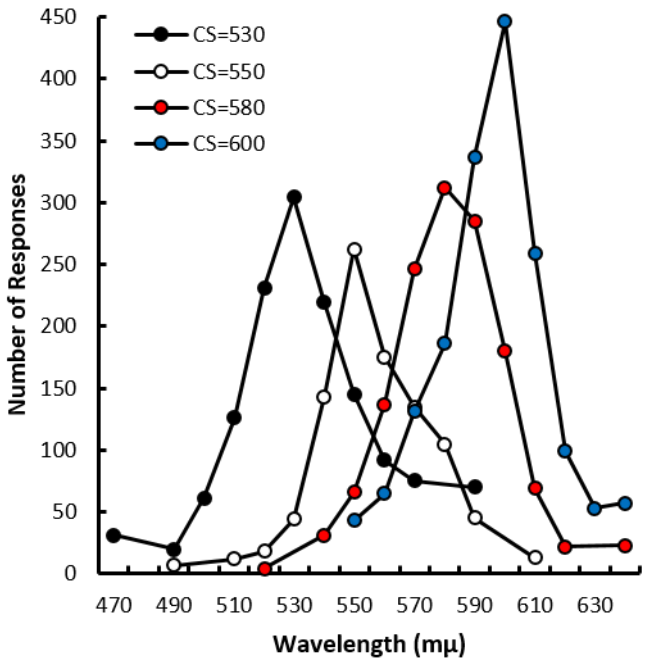

Figure 2. Example mean generalisation gradients from four groups of pigeons. Data from Guttman and Kalish (1956). Each group was trained to peck a key light of a specific wavelength and then tested for stimulus generalisation.

Discrimination Learning

If generalisation gradients are indeed symmetrical then it follows that the rate at which a discrimination is acquired should be unaffected by which of two discriminanda signal the reinforcer. One study that supports this assumption was conducted by Sutherland and Williams (1969). Two groups of rats were trained to discriminate between a regular and an irregular checkerboard. The animals were placed on a platform and then required to jump to one of two other platforms, one below each of the stimuli. A jump to the platform below the stimulus designated as the reinforced conditioned stimulus (S+) resulted in the rat being given food, whilst a jump to the platform below the stimulus designated as the nonreinforced conditioned stimulus (S-) did not result in any food. For half the animals the regular checkerboard signalled food and the irregular checkerboard signalled the absence of

food (A+B-) whilst for the remaining animals this relationship was reversed so that the irregular checkerboard signalled food and the regular checkerboard signalled the absence of food (B+A-). Both groups reached a discrimination criterion over the two days of training, that is, they successfully came to jump more often to the platform beneath S+ than S-. Moreover, consistent with the prediction that the two discriminations should be acquired at the same rate, the mean number of trials taken to reach this criterion did not significantly differ between groups.

A large number of other experiments have also utilised discrimination training. Often, like Sutherland and Williams (1969), these experiments are fully counterbalanced so that when one group is given an A+B- discrimination there is also a second group presented with a B+A- discrimination. However, almost always little or no mention is made of a symmetry in acquisition. It is therefore dangerous to assume that the acquisition between two counterbalanced groups is symmetrical in all cases.

Stimuli differing in magnitude: Asymmetrical generalisation gradients and asymmetrical discrimination acquisition

All of the theories reviewed in the previous section either predict or assume that; (1) generalisation gradients around stimuli should be symmetrical (an effect observed by Guttman and Kalish, 1956), and (2) that the acquisition of a discrimination between two stimuli should be the same regardless of which signals the reinforcer. Unfortunately for these theories, generalisation gradients do not always appear to be symmetrical. These demonstrations of asymmetry have been

revealed with stimuli that differ in magnitude[2]. At first sight it would seem that the similarity of a high-magnitude stimulus (e.g. a loud tone, *T*) to a low-magnitude stimulus (e.g. a soft tone, *t*) should be the same as a low-magnitude stimulus to a high intensity stimulus. If this were the case then the strength of the response to *T* after conditioning with *t* will be the same for *t* after conditioning with *T*. The results from the experiments described next challenge this prediction.

Asymmetrical generalisation gradients

The asymmetrical nature of intensity generalisation gradients can be seen with reference back to the classic experiments conducted by Pavlov (1927). After initial training in which a CS was consistently paired with the delivery of food – the subjects here were hungry dogs - Pavlov tested the strength of the CR, the amount of salivation, to novel stimuli that differed from the original in terms of intensity. The data from 67 such studies from Pavlov's laboratory have been summarised by Razran (1949) and are presented in Figure 3. In contrast to the symmetrical decrements in CR either side of the CS as demonstrated by Guttman and Kalish (1956; see also Moore, 1972), the level of CR in this case was found to increase linearly as a function of stimulus intensity. In other words, the higher the intensity of the test stimulus, the stronger was the response on the test trial. As a result, the

[2] Magnitude will be defined as the position of a stimulus along a dimension (e.g. brightness), where greater distance from 'zero' on that dimension corresponds to greater magnitude. Intensity, which I shall also refer to, especially in relation to auditory stimuli, relates to the perception of magnitude. With many dimensions such as brightness or loudness there is a clear relationship between a stimulus' magnitude and its intensity. However, with other dimensions there is a less obvious relationship. For example, it is less clear if a long rectangle is perceived to be more intense than a short rectangle. Whether the distinction between these two types of stimuli is an important one will be discussed in the general discussion.

strongest test CR was not observed to the training CS, but to a stimulus that was more intense than the training CS.

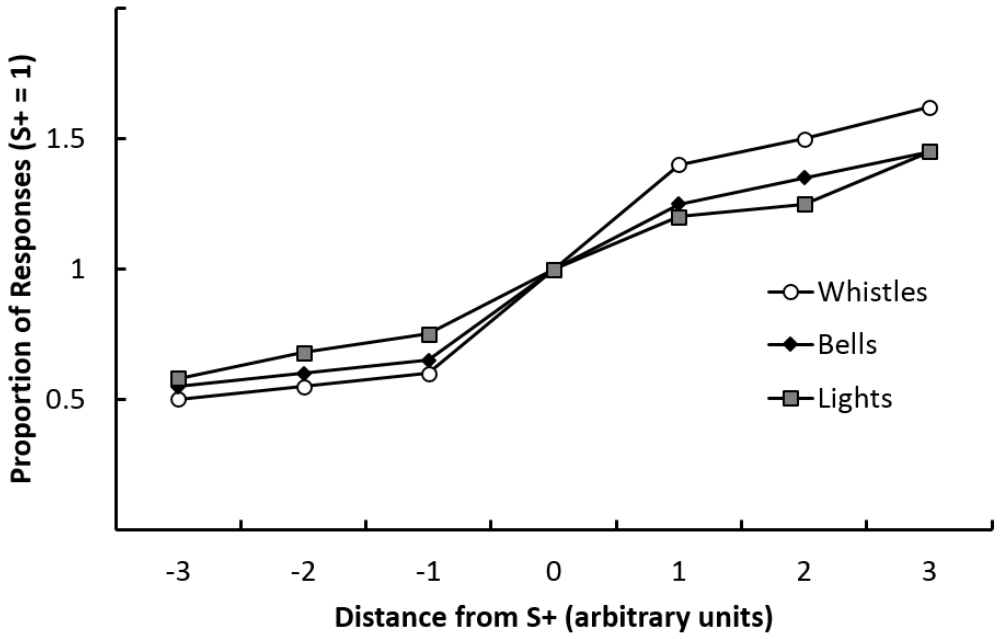

Figure 3. Summary of the monotonic generalisation gradients from over 250 of Pavlov experiments. Data from Razran (1949).

Demonstrations of this pattern of stimulus generalisation are not restricted to Pavlov's laboratory. In an experiment by Scavio and Gormenzano (1974) two groups of rabbits were given multiple sessions of eye blink conditioning in which a tone was consistently paired with a small shock. For the loud group the tone was 86 dB whilst for the soft group the tone was 65 dB. Across a number of sessions of acquisition training the rabbits came to make an eye blink at the onset of the tone. Following this training the rabbits were given presentations of an 86-dB, 79-dB, 72-dB and 65-dB tone in extinction. Results indicated that the level of responding was a direct function of the CS intensity. Both groups demonstrated the lowest level of responding to the 65-dB tone and the highest level of responding to the 86-dB tone.

Not all experiments have demonstrated the same monotonic gradients observed by Razran (1949) and Scavio and Gormenzano (1974). In an experiment described by Grice and Saltz (1950) rats were trained to approach and pass through a door onto which was attached a white circle. For half the rats this circle had an area of 20 cm^2 whilst for the remaining rats the circle was larger, with an area of 79 cm^2. After training, subgroups of rats from the 20-cm^2 group were given extinction trials with circles of 20, 32, 50, and 79 cm^2. Subgroups of rats from the 79-cm^2 group were given extinction trials with circles of 79, 63, 50, 32, and 20 cm^2. Consistent with the pattern of results described by Razran (1949) rats trained with the large circle showed a declining level of responding to smaller circles. In contrast, rats trained with the small circle made a slight increase in responding to the 32-cm^2 circles, although the level of responding to stimuli larger than this did decline. Thus gradients moving from high to low intensity and from low to high intensity seem to be asymmetrical, but not always in the same manner.

A related pattern of results has been found more recently by Kosaki and Pearce (2015). Two groups of rats were trained to swim to one of two submerged platforms that were located in front of two opposing walls of a square white swimming pool. Black panels were attached to the centre of the walls adjacent to the platforms. For one group these panels were 25 cm long whilst for a second group these panels were 100 cm. Over a number of sessions, rats in both groups readily learned to locate the hidden platforms. After this treatment both groups were tested in an arena in which one panel was 100 cm, the other was 25 cm and the platform was removed. Rats in the 100-cm group spent more time searching for the platform in front of the 100-cm panel than the 25-cm panel, suggesting a conventional gradient where test stimuli

differing from the training stimulus elicit a reduced CR. Conversely, rats in the 25-cm group spent an approximately equal amount of time searching in front of both panels. This suggests that the gradient moving from small to large panels was unconventional in the sense that there had been very little reduction in CR. Thus, in contrast to the assumption that the amount of generalisation between two stimuli should be equal, i.e. that the influence of a 100-cm panel should generalise as much to a 25-cm panel as the influence of a 25-cm panel will generalise to a 100-cm panel, these data indicate that there may be a different relationship between high- and low-intensity stimuli. In this case generalisation from a short to a long panel was greater than from a long to a short panel.

Asymmetrical generalisation gradients have also been identified when animals were initially given training with two stimuli in a discrimination. Pierrel and Sherman (1960) trained rats to discriminate between a 70-dB and 90-dB tone. For the 70+/90- group presses on a lever during the 70-dB tone, but not the 90-dB tone, were reinforced according to a variable interval 60-s schedule (VI-60 s). For the 90+/70- group lever presses were rewarded during the 90-dB tone but not the 70-dB tone. This training was then followed by generalisation tests in which the rats were presented with the original S+ and S-, and seven novel different stimuli with intensities between 60 dB and 100 dB increasing in increments of 5 dB. Lever presses during the presentations of these novel stimuli were not rewarded with food. Initially, the generalisation gradients were found to be monotonic; rats pressed the lever more often in the presence of stimuli at the extreme end of the dimension away from the S-. However, after the first two sessions the monotonic gradients became peaked, that is to say as intensity moved from S- to S+, and then beyond, so did the

level of responding up to a point after which the level of responding began to decrease. This suggests that the shape of intensity gradients may change with extended testing (Ghirlanda & Enquist, 2003; although this is perhaps not surprising considering the effects of nonreinforcement). Nonetheless, regardless of the precise shape of the gradients, they remained asymmetrical demonstrating a higher rate of responding to stimuli beyond the S+ in the direction away from the S- than to the S+ itself. Importantly, this was the case when S+ was the high intensity and the low intensity stimulus. A simple explanation for this effect, which shall be discussed later, is peak shift.

A similar pattern of results to Pierrel and Sherman (1960) with a different stimulus dimension has been identified by Ernst, Engberg and Thomas (1971). In this experiment pigeons were initially trained to peck at a response key in order to obtain grain. For half the birds key pecks in the presence of a high intensity light (12.45 mL) but not a low intensity light (3.94 mL) were rewarded according to a VI-60 s schedule. For the remaining birds this reward contingency was reversed. This training was then followed by a test session in which the birds were presented with nine different intensities of the light (between 0.70 and 70.00 mL) in extinction. All birds presented asymmetrical response gradients with a maximum rate of responding beyond S+ in the direction away from the S- (see Figure 4). For two of the six pigeons the gradients monotonically increased beyond S+ whilst for the remaining four pigeons the gradients displayed a peak which was displaced beyond the S+ away from the S-.

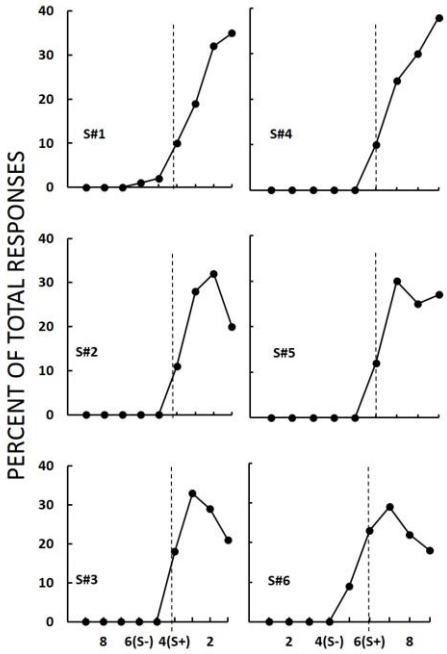

Figure 4. Examples of generalisation gradients from subjects in the light+/dim- groups (left-hand panels) and dim+/light- groups (right-hand panels). The numbers on the abscissas correspond to intensity values (in mL): 1 = 70.00; 2 = 39.36; 3 = 22.14; 4 = 12.45; 5 = 7.00; 6 = 3.94; 7 = 2.21; 8 = 1.24; and 9 = 0.70. Data from Ernst, Engberg & Thomas (1971).

Taken together, data from the experiments just described suggest that the gradients obtained by Razran (1949) and Scavio and Gormezano (1974) do not accurately describe the relationship between stimulus intensity and the strength of CR. Nonetheless, it does appear that these gradients are not symmetrical. After analysing 38 intensity gradients from a range of studies, including those described above (16 monotonic and 22 peaked), Ghirlanda and Enquist (2003) found that there was a large number of instances where the strength of responding to stimuli beyond S+ in the direction away from S- was never weaker than to S+ itself. This outcome was the case for 30 out of the 38 examined gradients and a Fisher's exact probability test revealed this to be a significant effect, $p < 10^{-8}$. This clear departure from the symmetrical gradients assumed by theories of conditioning has obvious implications

for the predictions they make concerning discriminations between two stimuli differing in magnitude. In particular, if generalisation gradients are asymmetrical then the amount of generalisation between two stimuli that differ in intensity should be different depending on whether the high or the low intensity stimulus signals the reinforcer. Thus it seems likely that intensity discriminations are acquired at different rates based on which stimulus is assigned as the CS+. In the following section, I present some examples where this seems to be the case.

Discriminations based on stimulus magnitude

A small and varied literature exists concerning the ability of animals to discriminate between stimuli based on magnitude. In direct contrast to the predictions made by the theories of discrimination outlined above, changing which of two stimuli signals the reinforcer alters the ease at which animals can acquire the task. In these instances, when a high magnitude stimulus signals the reinforcer and the low magnitude stimulus signals the absence of the reinforcer, the discriminations appear to be acquired much more readily than when the low magnitude stimulus signals the reinforcer and the high magnitude stimulus signals the absence of a reinforcer. Interestingly, although this pattern of results is not predicted by the theories of conditioning described earlier, it was in fact directly predicted by Hull's (1952) theory of behaviour (although Hull specifically referred to intensity as opposed to magnitude):

"*When the simple discrimination of two stimulus intensities occurs, the difference between the intensities remaining constant, the process is more effective in terms of the net reaction potential yield when reinforcement is given to the more intense rather than to the less intense of the two discriminanda*" (Theorem 17B, Hull, 1952).

This effect has indeed been demonstrated with a range of stimuli including the amplitude of auditory stimuli (Jakubowsak & Zielinski, 1976; Pierrel, Sherman, Blue & Hedge, 1970), odour (Pelz, Gerber, & Menzel, 1997), quantity (Vonk & Beran, 2012; Watanabe, 1998), temporal duration (e.g. Bouton & Garcia-Gutierrez, 2006; Bouton & Hendrix, 2011), and length (Kosaki, Jones, & Pearce, 2013). Data from each of these different dimensions will be considered now in turn.

Auditory Intensity

Jakubowska and Zielinski (1976) provide an early example of the asymmetry in magnitude discriminations when two groups of rats were trained to differentiate between two different intensities of a white noise in a conditioned emotional response (CER) procedure. Rats were initially trained to press a lever in order to receive a food reward in the presence of a 70-dB and a 50-dB white noise. For the 70+/50- group, the presentation of the high-intensity white noise ended with the delivery of a foot shock. The low-intensity white noise was never paired with shock. For a second 50+/70- group, this contingency was reversed so that the low-intensity white noise was paired with a shock and the high-intensity white noise was not. Over a number of training sessions rats in both groups came to suppress the rate of lever pressing during the presentation of the stimulus associated with shock whilst maintaining a high level of lever pressing during the stimulus not associated with shock. However, rats in the 70+/50- group were quicker to suppress their responding during the presentation of the CS+ than rats in the 50+/70- group. This pattern of results suggests that the two stimuli were more easily differentiated when the high-magnitude stimulus signalled a shock than when the low-magnitude stimulus signalled a shock. This result presents an obvious challenge to Stimulus Intensity

Dynamism (Hull, 1949, 1952) because the CR here was the suppression of a lever press, whereas Hull proposed that the effect of intensity was to energise subjects to respond more.

Pierrel, Sherman, Blue and Hegge (1970) presented one group of rats with an instrumental appetitive discrimination between a high-intensity S+ (100-dB tone) and a low-intensity S- (80-dB tone), and a second group of rats with a discrimination in which this contingency was reversed. Over the course of training the rats were given two, eight-hour sessions in which the presentations of S+ and S- were alternated. During the inter-session intervals, which lasted for four hours, rats remained in the test chambers. Rats in the 100+/80- group came to make more lever presses in the presence of the S+, and suppressed lever pressing during S- more quickly than those in the 80+/100- group indicating that the former group acquired the discrimination more readily than the latter.

Quantity

A similar asymmetry in the discrimination of magnitude was obtained by Vonk and Beran (2012) when investigating the ability of three American black bears to discriminate based on the relative number of dots in two arrays. These two arrays were presented simultaneously, on either side of a touch-screen monitor. The two arrays never contained the same quantity of dots and varied in the number they contained between a value of one and ten. One of the bears was trained to select the stimulus containing the larger number of dots in order to receive a food reward whilst the remaining two bears were rewarded for choosing the stimulus with the fewer number of dots. The bear trained with the more+/fewer- discrimination required fewer sessions of training (22) to reach the 80% discrimination criterion

than both bears in the fewer+/more- discrimination (at least 36 sessions) suggesting that the more+/fewer- task was easier to acquire. Of course, it is possible that the bears had been using a dimension other than number in order to solve these discriminations. To assess this possibility, Vonk and Beran conducted further tests in which the number of dots varied independently with the area of the dots. In support of the idea that bears had been solving the discriminations based on the relative number of dots, the bears reached above-chance levels of responding to their respective more+/fewer- and fewer+/more- discriminations despite this manipulation. Similarly to Vonk and Beran, Watanabe (1998) found that two pigeons were readily able to discriminate between stimuli when four red-balls signalled the delivery of food and when two red-balls signalled the absence of food. Again, further tests in which the nature of the stimuli were changed, for example using novel objects, suggested that they had been using number rather than any other dimension in order to solve the discriminations. Importantly, a second pair of birds was unable to differentiate between the stimuli when the two red-balls signalled the delivery of food and the four red balls signalled the absence of food. Of course, the limited sample size of this study and that of Vonk and Beran (2012) makes it difficult to trust the reliability of these asymmetries.

This limitation of the above studies is highlighted when it is appreciated that other studies have failed to demonstrate a difference between the acquisition of a many+/few- discrimination and few+/many- discriminations. For example, Agrillo, Dadda, Serena and Bisazza (2009) trained mosquitofish to approach and swim through one of two opposing doors, each of which was located beneath a set of geometric shapes. One of these sets contained two shapes whilst the other contained

three. Movement through the correct door allowed the fish to swim to an outer tank and re-join conspecifics. Even when the non-numerical data such as cumulative area and luminance were controlled for, there were no significant differences in the proportion of correct first choices made by fish trained to approach the three- and two-object stimuli, suggesting that there was no asymmetry. However, prior to training the fish were given a number of pre-training sessions in a different tank with the same S+ and S-. In this tank movement through the door under the stimulus designated as S+ resulted in the fish being able to shoal with conspecifics. Thus it is possible that any asymmetry that might have developed in this task may have disappeared before training in the second tank. The results from the first stage were, however, not reported.

Temporal duration

Kyd, Pearce, Haselgrove, Amin and Aggleton (2007), examined if rats could discriminate between two different durations of a tone; one long and one short. For two short-duration groups these durations were 1.5 s and 0.5 s, whilst for two long-duration groups the durations were 12 s and 3 s. Each CS was followed by a 10-s trace interval which, after S+ but not S-, was followed by the delivery of food. For both pairs of groups, when the long CS was followed by a food reward and the short CS was not, the discrimination was more easily solved than when the short CS was followed by food and the long CS was not. In both instances this asymmetry emerged due to the failure of rats in the short+/long- groups to inhibit responding in the trace interval following the long stimulus.

Another temporal asymmetry has been identified in a series of experiments in which the duration of the inter-trial interval (ITI) rather than the CS was manipulated

(Bouton & García-Gutiérrez, 2006; Bouton & Hendrix, 2011). Rats were given multiple presentations of a 10-s tone which was preceded by an ITI which was either 16 min or 4 min in duration. For one group of rats the tone was followed by a delivery of food after the 16-min ITI but not the 4-min ITI (16+/4-). A second group was presented with the same task with the exception that the tone was followed by food after a 4-min, but not a 16-min ITI (4+/16-). Rats given the 16+/4- discrimination came to respond significantly more during the CS after the 16-min ITI than after the 4-min ITI. In contrast, rats in the 4+/16- group failed to demonstrate differential responding and in particular this effect appeared to be driven by rats' inability to inhibit responding after the 16-min ITI. A similar experiment in which humans were required to discriminate between ITIs of 4 s and 1 s found a comparable asymmetry. In this case, participants anticipated a positive target (an alien) and responded to it (a key press to 'shoot' the alien) readily when it followed a long ITI than a short ITI. Furthermore, in the same manner as rats, when the positive target followed the short but not the long ITI, there was little evidence of the discrimination being solved (Astley, Aird & Bouton, 2015).

Using a different design in which the duration of a stimulus rather than the ITI signalled the presence or absence of reward, Todd, Winterbauer and Bouton (2010) presented rats with food after a 10-s target tone when a preceding feature stimulus, a white noise, was 4 min but not 1 min (4+/1-), or 1 min but not 4 min (1+/4-). Trials were separated by an ITI (tone offset to noise onset) in which no auditory stimuli were presented (See Figure 5). Rats displayed better discriminatory behaviour in the 4+/1- group than the 1+/4- group. Indeed, there was little evidence of the 1+/4- group showing the discrimination, again a result of a failure to inhibit

responding following a 4 min noise by the 1+/4- group. A second experiment used shorter durations for the noise (60 s vs. 15 s). When conditioning to this cue was reduced by the addition of extra non-reinforced trials the same long+/short- asymmetry was observed. It is of interest to note than when the extra non-reinforced trials were omitted, then rats showed better discriminatory behaviour during the feature stimulus for a 15+/60- task than a 60+/15- task. To account for the reversed asymmetry Todd, Winterbauer and Bouton noted the asymmetry had not been observed during the target CS (the tone), but instead in a 10-s period preceding the tone. Thus it appears in this instance the white noise had been the primary stimulus that elicited conditioned responding. Considering that shorter duration stimuli have been identified to result in better conditioning than long duration stimuli (Bouton & Sunsay, 2003), it follows that 15+/60- discriminations should be better acquired than 60+/15- discriminations due to the better excitatory conditioning of the 15-s white noise than the 60-s white noise.

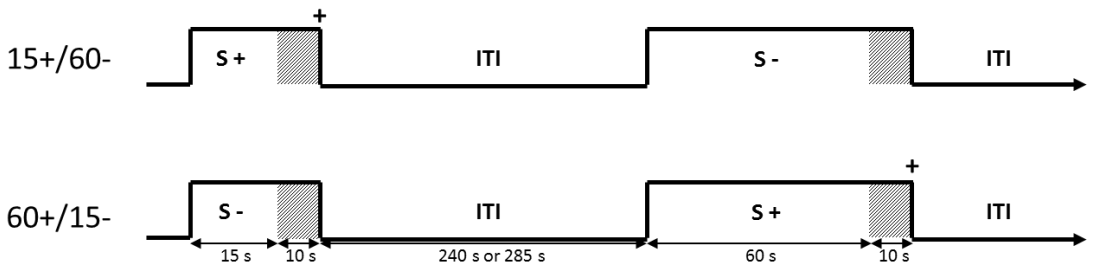

Figure 5. Summary of the experimental design used by Todd, Winterbauer, and Bouton (2010). The stimuli labelled as S+ and S- are a white noise feature-stimulus of different durations. The grey areas show the 10-s target CS (a tone). The symbol + represents the delivery of two food pellets. S+ and S- were presented in a pseudo-random order.

The experiments so far described concerning temporal duration have demonstrated what Bouton and colleagues referred to as the long+ effect. This refers

to the asymmetry whereby discriminations in which a long but not a short duration signals the delivery of a reward are acquired more readily than discriminations in which a short, but not a long duration signals the delivery of a reward. Interestingly, there is at least one other report, in addition to Todd, Winterbauer and Bouton (2010), of an asymmetry in temporal duration discriminations where the short+/long- task was acquired more readily than long+/short- task. Kehoe and Bosenberg (2002) presented rabbits with trials in which a 66-s feature stimulus (a tone) was followed after a feature-target interval by a 400-ms target stimulus (a flashing light). The feature-target interval was either 5 s or 45 s. For one group, the presentation of the target stimulus terminated with a small electric shock near the rabbit's eye after the 5-s, but not the 45-s interval whilst, for a second group, the shock followed the 45-s but not the 5-s interval. After 11 days of training rabbits in both groups were more likely to elicit a CR (an eye blink) following the feature-target interval paired with shock than after the feature-target interval not paired with shock. However, rabbits in the 45+/5- group required many more sessions to inhibit the CR during the target stimulus following CS- than the 5+/45- group.

One account for the pattern of results observed by Kehoe and Bosenberg (2002) is that the rabbits were not using the duration of the feature-target interval to solve the discriminations. Instead, it is possible that the tasks were solved based on the strength, or intensity, of the memory trace of the feature stimulus. Assuming that the memory of the tone decays over time and that this memory trace will be weaker after a 45-s interval than a 5-s interval, it is possible to describe the 45+/5- discriminations as weak+/strong-, and the 5+/45- discrimination as strong+/weak-. In this manner, the pattern of results obtained by Kehoe and Bosenberg is completely

consistent with those from other demonstrations of an asymmetry in the discrimination of intensity. Whether explaining the results in this manner is justified remains to be determined.

Odour

Using an unusual stimulus dimension and species Pelz, Gerber and Menzel (1997) identified an asymmetry when honeybees were trained to discriminate between two intensities of an odour. Bees were given trials in which a high or low concentration of linalool was presented in a steady air flow. For half the bees the high concentration was followed by a drop of sucrose whilst the low concentration was not followed by sucrose. For the remaining animals the low concentration was followed by sucrose and the high concentration was not. The measure of conditioned responding was the percentage of proboscis extension during the presentation of the odour. Bees in the high+/low- group were able to acquire the discrimination over trials as indicated by a greater percentage of proboscis extension in the presence of the CS+ than CS-. Bees in the low+/high- group however were unable to acquire the discrimination, in fact displaying greater proboscis extension across trials to the non-reinforced high intensity odour than to the reinforced low intensity. A control group in which bees were required to discriminate between the low intensity odour and an odourless solvent demonstrated that bees were able to detect the low concentration.

Length

Kosaki, Jones and Pearce (2013) trained rats to find one of two hidden platforms in a square swimming pool, the walls of which were grey (see Figure 6). Attached to the centre of each wall was a black panel. For two opposing walls the

panels were long (100 cm), while for the other two opposing walls the panels were short (50 cm). Landmarks above the centre of each panel helped rats to find the platforms. For half the animals the hidden platforms were in the middle of the long, but not the short panels (long+/short-), while for the remaining animals the platforms were in the middle of the short but not the long panels (short+/long-). Following this training, they were then given test trials in which the hidden platforms were removed. Rats in the long+/short- group spent significantly more time searching for the platform in front of the correct length panel than rats in the short+/long- group. This effect was identified when the long panel was twice and four times the size of the short panel and when the arena was a rectangle rather than a square.

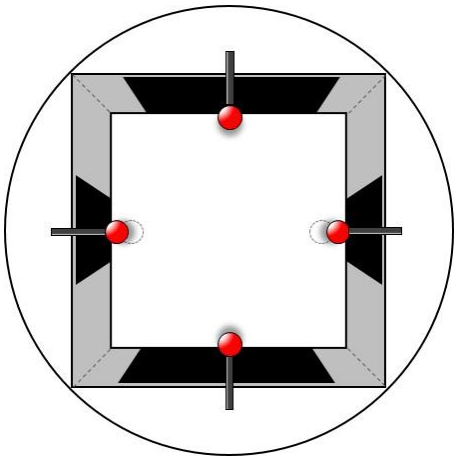

Figure 6. Illustration of the water-maze used by Kosaki et al. (2013), showing the grey walls, landmarks above the centre of each panel (filled circles), black panels, and submerged platforms (dashed circles).

Explanations for asymmetrical generalisation gradients and discriminations

The review of evidence above suggests that there might be something unique about magnitude dimensions. First, gradients that form around stimuli on a magnitude dimension appear to be asymmetrical. For example, after conditioning

with a training stimulus, presentations of a higher-magnitude test stimulus often results in a stronger CR than to presentations of a lower-magnitude test stimulus (e.g. Razran, 1949). Second, discriminations in which the high-magnitude stimulus signals a reward and a low-magnitude stimulus signals the absence of reward are often acquired more readily than the reverse treatment (e.g. Pierrel et al. 1970). These data pose a challenge to the theories considered earlier, which were shown to predict symmetrical generalisation gradients and symmetrical discriminations. The following section of the thesis will explore how these theories might account for the aforementioned findings. First, I will return to the theory of Hull (1949), which proposes a direct relationship between stimulus intensity and conditioned responding. Following this, some adaptations of a stimulus generalisation account of discrimination learning (Spence, 1936, 1937) will be considered. Finally, I will describe an alternative account offered by Bouton and colleagues (Bouton & García-Gutíerrez, 2006; Bouton & Hendrix, 2011) in which high intensity stimuli might be considered to be made up of more elements than low intensity stimuli.

Stimulus Intensity Dynamism

Hull (1949) proposed that the magnitude of a conditioned response shares a linear relationship with the intensity of a stimulus. This positive correlation is referred to as stimulus intensity dynamism. According to this ‘law of strength’ (Gray, 1965, p.180), asymmetrical response gradients either side of a S+ are a result of an energising effect which serves to strengthen the CR to stimuli more intense than S+, and also to energise responding to S+ itself above less intense stimuli. A generalisation decrement will occur with stimuli that are different to S+, but the energising effect of stimulus intensity dynamism will, at least to some extent,

counter this effect with stimuli of greater magnitude than S+, and enhance this effect with stimuli that are lower magnitude than S+. Hull's theory thus readily predicts the increasing monotonic gradients described by Razran (1949) and demonstrated by Scavio and Gormezano (1974).

The energising effect can also explain some of the asymmetries seen in intensity discriminations (see Theorem 17B, Hull, 1952, as quoted earlier). Recall that Pierrel et al. (1970) trained rats to discriminate between a 100-dB tone and 80-dB tone. For one group lever presses during the presentation of the loud, but not the soft tone resulted in the delivery of food while for the second group lever presses during the presentation of the soft, but not the loud tone resulted in food. According to stimulus intensity dynamism, for the 100+/80- group rats will be energised to make more lever presses in the presence of the 100-dB tone than the 80-dB tone and as a result will come to solve this task rather quickly. In contrast, for the 80+/100- group the rats will continue to be more energised by the 100-dB tone, and will therefore make a large number of incorrect responses during non-reinforced trials. Thus the 80+/100- discrimination will be acquired more slowly than the 100+/80- discrimination.

There are, however, reasons for believing that the above account does not apply to all the results considered in the previous section. For example, stimulus intensity dynamism (Hull, 1949) cannot account for monotonic gradients that extend beyond the S+ in the direction away from S- when S+ is the less intense stimulus (e.g. Ernst et al., 1971; Pierrel & Sherman, 1960). This is because as stimuli become less intense than S+ the energising effect from stimulus intensity dynamism on the

strength of responding should also be reduced. Stimuli less intense than S+ should thus elicit a weaker CR than stimuli more intense than S+.

Stimulus intensity dynamism (Hull, 1949) also struggles to account for the asymmetry demonstrated by Kyd et al., (2007). In one experiment, two groups of rats were presented with tones of 1.5 s and 0.5 s in duration, followed by a 10-s trace interval in which no tone was presented. For one of these groups, Group Long, the 1.5-s tone signalled the delivery of food after the trace interval whilst for the other group, Group Short, the 0.5-s tone signalled food. The results demonstrated an asymmetry in the acquisition of the discrimination (Group Long acquired the discrimination more readily than Group Short) but this effect was observed during the trace interval and not the tone. Considering that the trace intervals were always the same duration and the same auditory intensity it is difficult to find a reason why stimulus intensity dynamism would be responsible for the observed asymmetry during this period.

A final issue with stimulus intensity dynamism is that Hull (1949) specifically referred to the effect of stimulus intensity whilst also not commenting on stimulus magnitude more generally. There are then some dimensions described above that display an asymmetry but where it is less clear how they are perceived in terms of intensity, such as panels in a swimming pool (Kosaki et al., 2013). There seems to be no obvious reason why a long black panel with short grey panels either side should be regarded as more intense than a short panel with longer grey panels either side. Indeed, at a concrete level the walls with short black panels can be described as being brighter than walls with long black panels. Thus whenever a subject looks at a wall with a short panel the rate of neuronal firing by photoreceptors should be

greater than when the subject looks at a wall with a long panel. In this manner, walls with short panels can be described as being more intense than walls with long panels. As a result, stimulus intensity dynamism predicts that in the experiments conducted by Kosaki et al., Group Short+ should have acquired the discrimination more readily than Group Long+, the opposite to what was observed. Without a clear method for relating the magnitude of a stimulus to its intensity, it is difficult to apply stimulus intensity dynamism to many of the demonstrations of the asymmetry in magnitude discriminations including panel length (Kosaki et al.), quantity (e.g. Vonk & Beran, 2012) and temporal duration (e.g. Bouton & Hendrix, 2011).

Gradient Interaction

Gradient interaction theory (Spence, 1936, 1937) predicts that a discrimination between two stimuli should be acquired at the same rate regardless of which stimulus signals the reinforcer. This prediction follows from the assumed symmetrical nature of the gradients that form around CS+ and CS-. However, the fact that high-intensity+/low-intensity- discriminations are acquired more readily than low-intensity+/high-intensity- discriminations, and that the generalisation gradients around intensity stimuli are asymmetrical, indicates that this primary assumption might be incorrect. In fact, the experiments just reviewed may pose less of a challenge to the theory of Spence than might at first sight appear to be the case, once account is taken of the role played by background cues.

Consider the experiment conducted by Kosaki and Pearce (2015). In this experiment rats were trained to find one of two hidden platforms in an arena which contained either two short panels or two long panels. The platforms were always located in front of the panels. The rats were then tested in an arena which contained

one long panel and one short panel, and in which the platform had been removed. Rats trained with short panels spent an equal amount of time swimming in front of the long and short panels, indicating a large amount of generalisation from short to long, whilst rats in trained with long panels spent significantly more time swimming in front of the long panel indicating relatively little generalisation from long to short. An explanation for why this pattern of results might occur has been offered by Mackintosh (1974) whose account follows proposals by Perkins (1953), and Logan (1954). According to Mackintosh:

"*The greater the intensity of a CS, the less the generalisation of inhibition from non-reinforced background stimuli and therefore the faster the rate of learning*." (Mackintosh, 1974, pp.532-533)

For rats given a long+/short- discrimination the long panel will acquire an excitatory association with the platform while the short panel will acquire an inhibitory association. In the manner described by Spence, these associations will generalise in a symmetrical fashion along the dimension of length, to other similar stimuli. In addition, because the platform is always located in front of a panel, the grey areas of wall surrounding the panels, the background stimuli, will enter into an inhibitory association with the platform. According to Perkins (1953) and Logan (1954), background stimuli can be considered as 'situation-minus-CS' or zero-intensity stimuli. If this is the case then it follows that excitation acquired by the long black panels will be relatively unaffected by the generalisation of inhibition from the grey panels and the discrimination will be acquired rapidly. In the case of the long+/short- discrimination the long panels gain associative strength quickly and the discrimination is rapidly solved. In the case of a short+/long- task, the excitation

gained by the short panel will be masked by the generalisation of inhibition from the grey walls and as a result the capacity for the reinforced short panels to elicit a response will be less than the reinforced long panels, thus the asymmetry emerges between long+/short- and short+/long- discriminations.

The same argument can be made for other demonstrations. For example, Jakubowska and Zielinski (1976) trained rats in a conditioned emotional response (CER) experiment in which the presence of one, but not another intensity of white noise was followed by an electric shock. Rats soon came to suppress a lever-press response in the presence of the stimulus paired with shock, but not the stimulus not paired with shock. These, trials were separated by an inter-trial interval (ITI) in which no white noise was presented. During these periods no shocks were delivered and thus any cues present will have acquired an inhibitory association with shock. Importantly, these cues will include the absence of the white noise which might be considered as an auditory stimulus of less intensity than either of the two stimuli. Thus an inhibitory association between the absence of a white noise and the shock will form over training. As a result the generalisation of inhibition from the ITI will reduce the capacity of the S+ to elicit a response in the 50+/70- discrimination, thus making the discrimination harder to acquire than when the S+ is the high intensity stimulus.

Cues from a nonreinforced ITI may be a source of generalisation for a number of other studies with different dimensions. Watanabe (1998) for example gave pigeons multiple presentations of visual stimuli containing four or two red-balls. Importantly, trials were separated by a 5-s blackout period in which no stimuli could be seen. Similarly, Vonk and Beran (2012), who trained bears to discriminate

between ‘many’ vs. ‘few’ dots, also presented subjects with a blank screen after incorrect trials. If these ITIs are represented by the subjects as zero on a quantity dimension then it follows that they will be considered to be more similar to the low quantity stimuli than the high quantity stimuli. Greater generalisation of inhibition would then be expected from the ITI to these stimuli and hence, for the reasons explained above, discriminations in which the larger quantity signals the reinforcer will be more readily acquired than when the smaller quantity signals the reinforcer.

Unlike Stimulus Intensity Dynamism (Hull, 1949), gradient interaction theory can also account for reversed generalisation gradients such as those demonstrated by Ernst et al. (1971). In this experiment pigeons were trained to discriminate between a two intensities of a light. They were then given generalisation tests with a variety of light intensities including intensities beyond the S+ in the direction away from the S-. Responding to test stimuli beyond S+ elicited greater responding to S+ even when S+ was a low intensity light. This can readily be explained with reference back to Figure 2. Recall that the strength of response elicited by a stimulus is determined by the difference between the hypothetical excitatory and inhibitory gradients. Because the net excitation is greater for S+ than S- it correctly follows that subjects will respond more in the presence of S+ than S-. Furthermore, it is evident from Figure 2 that a stimulus to the left of the S+, a less intense stimulus, will have an even greater net excitation than S+ itself. Thus, regardless of whether S+ is of a higher or lower intensity than S- it follows that tests with stimuli further along the dimension away from S- will initially elicit stronger responding than S+. In other words, the results can be viewed as a demonstration of the peak shift.

The results from experiments by Bouton and Hendrix (2011) and Bouton and García- Gutíerrez (2006), pose a challenge to gradient interaction theory (Spence, 1936, 1937). Rats received food after a tone when successive presentations of this stimulus were separated by a long, but not a short inter-trial interval (long+/short-), or they received the opposite of this treatment (short+/long-). Despite the fact that cues indicating the trial outcome were present throughout the experimental session, and there were thus no periods of nonreinforced exposure to cues in the absence of S+ and S-, an asymmetry was still observed: long+/short- discriminations were acquired more readily than the short+/long- discriminations. Nevertheless, the fact that the principles of stimulus generalisation explain many of the data described earlier in the chapter suggests that this account, and the later modifications made by Mackintosh (1974), Perkins (1953), and Logan (1954), should be taken seriously as a possible explanation for the asymmetries in magnitude discriminations.

The Feature-Positive Account

A contemporary account has been offered by Bouton and colleagues in order to explain the asymmetry seen in the discrimination of temporal duration (Bouton & García-Gutiérrez, 2006; Bouton & Hendrix, 2011). They proposed that the different durations of a stimulus, including the ITI might be represented as a series of temporal elements. In this fashion a short duration can be represented by the hypothetical element A. Furthermore, a long duration can be considered as the short duration, A, plus additional duration, B, thus AB. A long+ short- discrimination can consequently be described as an AB+A- discrimination whilst a short+ long- discrimination is an A+AB- discrimination. This logic can also be applied to any magnitude dimension beyond temporal durations and to any magnitude dimension

where a high-magnitude stimulus can be considered as including a lower-magnitude stimulus. For example, it is possible to describe a visual array containing a small number of dots as A, and an array containing a larger number of dots as comprising all the dots in A plus additional dots, B. Thus a few+/many- discrimination can be described as A+AB-, whilst a many+/few- discrimination can be represented as AB+A-. AB+A- is an example of a *feature-positive* discrimination where the unique element signals the reward, whilst A+AB- is an example of a *feature-negative* discrimination where the unique cue signals the absence of reward. It has been frequently demonstrated that feature-positive discriminations are more readily acquired than feature-negative discriminations (e.g. Hearst, 1978; Jenkins & Sainsbury, 1970). Why this is the case can be readily explained with reference to the Rescorla-Wagner model (1972). During training, AB+ in an AB+A- discrimination will gain excitatory associative strength more rapidly than A+ in an A+AB- discrimination due to the fact that both elements A and B accrue excitation independently. In other words, AB+ is more salient than A+ and thus gains excitatory strength more rapidly. Furthermore, in the *feature-negative* A+AB- discrimination, B has to acquire inhibitory properties, which only occurs once A has entered into an association with the reward. Thus, AB+A- or high-intensity+/low-intensity- discriminations are acquired more readily than A+AB-, or low-intensity+/high-intensity- discriminations. In order to confirm these predictions do indeed follow from the Rescorla-Wagner model two computer simulations were conducted. Figure 7 presents these simulations, which were based on the equation proposed by Rescorla and Wagner (Equation 1), for the acquisition of associative strength to A and AB with an A+AB- and AB+A- discrimination. The learning parameter α was set at 0.2 and, in keeping with arguments put forward by Wagner

and Rescorla (1972) and Jones and Pearce (2014), the value for the learning rate parameter, β, was greater for reinforced than nonreinforced trials; β+ = 0.5 and β- = 0.25. The value set for asymptote, λ, was 100. It is quite evident that the feature-positive discrimination is predicted to be acquired more readily than the feature-negative discrimination.

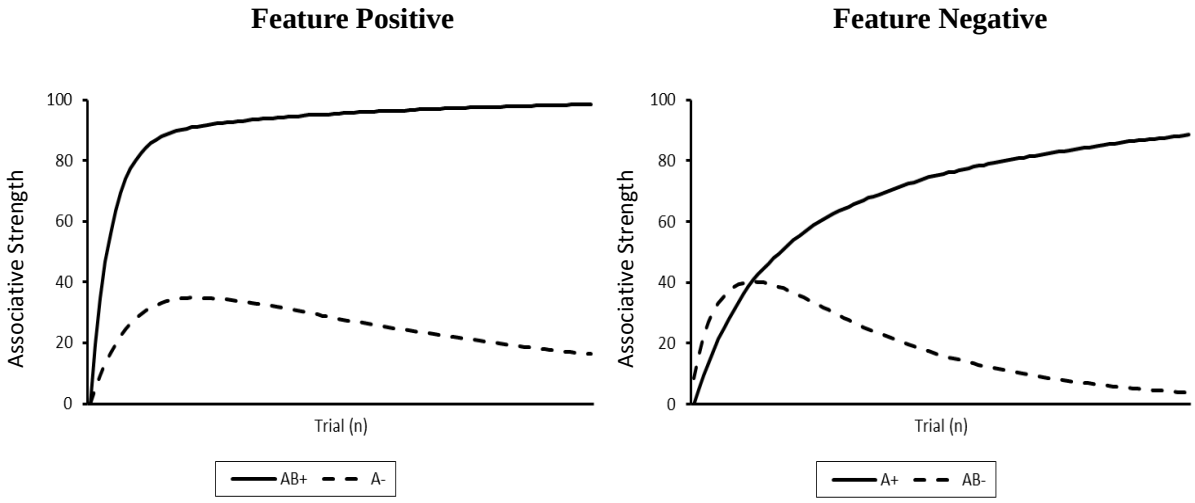

Figure 7. Computer simulations of the changes in associative strength for an AB+A- (left-hand panel) and A+AB- (right-hand panel) discrimination as predicted by the Rescorla-Wagner (1972) model.

Despite its success in accounting for the asymmetry across a range of different stimulus dimensions, the account based on the Rescorla-Wagner (1972) theory is unable to explain all empirical data. First, experiments conducted by Pierrel and Sherman (1960), Ernst et al. (1971) and Razran (1949) have all demonstrated that when a high intensity test stimulus, is presented after conditioning with a low intensity stimulus, the test stimulus elicits a higher level of responding than the training stimulus. If the training stimulus is construed as A, and the test stimulus as

AB, the Rescorla-Wagner (1972) model predicts that responding to both stimuli will be determined by the associative strength of element A on the first test trial. Additional test trials are then predicted in extinction because element B will accrue an inhibitory association.

The Rescorla-Wagner (1972) account of the asymmetry in magnitude discriminations is also challenged by the results obtained in an experiment by Kosaki et al. (2013) in which two groups of rats were trained to find a submerged platform in a square water-maze. Two opposing walls had a short panel, while the other two opposing walls had a long panel. The platforms for both groups were submerged in front of the short panels. For one group, the short panels were 15 cm and the long panels were 45 cm. For the second group the short panels were 70 cm and the long panels were 100 cm. Thus the absolute difference in cm between the panels was the same but the ratio between them was different. Rats were able to solve the discrimination and came to swim in front of the short panels significantly more in the 15+/45- group than the 70+/100- group. This is consistent with Weber's law (from which it follows that the ease of a discrimination will depend on the ratio between the two values rather than the absolute difference) but inconsistent with the Rescorla-Wagner model. In this instance, the Rescorla-Wagner model predicts the wrong outcome; the 70+/100- discrimination should be easier than the 15+/45- discrimination. To explain this prediction, it is useful to describe the 15+/45- discrimination as A+AC-, where A is 15 cm and C is 30 cm, and the 70+/100- discrimination as B+BC-, where B is 70 cm. Because B represents a larger area than A it can be considered to be more salient. As a result, the Rescorla-Wagner model makes the prediction that B will acquire more associative strength on each trial than

A, and consequently that C will acquire inhibitory associative strength more rapidly on BC- trials than AC- trials. Thus, some degree of caution should be taken when attempting to apply this model, at least in the way described above, to the discrimination of intensity.

Summary

Many theories of associative learning are based on the assumption that the similarity between two stimuli is symmetrical so that the similarity of stimulus A to stimulus B, is the same as the similarity of stimulus B to stimulus A. Furthermore, because stimulus generalisation is accepted to be related to the similarity between stimuli, it follows that the strength of a CR to stimulus B after conditioning with stimulus A should be the same as to stimulus A after conditioning with stimulus B. As a result, these theories make a common prediction that the rate at which a discrimination between two stimuli is acquired should be unaffected by which comes to signal the US.

However, I have presented evidence that suggests the relationship between similarity and generalised responding is not always symmetrical. Moreover, discriminations between two stimuli differing in intensity are also not symmetrical. In these instances, discriminations where the CS+ is a high intensity stimulus and the CS- is a low intensity stimulus are solved more readily than when the reverse contingency is true.

After presenting this evidence the challenge was to try and interpret it theoretically. One possible account of the findings is to consider intensity stimuli as sets of hypothetical elements. In this manner high intensity stimuli can be considered

as comprising more elements than low intensity stimuli. Thus the asymmetry in intensity discriminations can be described as an example of the feature-positive effect (Bouton & García-Gutiérrez, 2006; Bouton & Hendrix, 2011) which in turn can be readily explained by the Rescorla-Wagner (1972) model. However, the Rescorla-Wagner model is challenged by at least two experimental findings that it is unable to account for. First the model is unable to correctly predict that a 15+/45- discrimination was easier to solve than a 70+/100- discrimination (Kosaki et al. 2013). Furthermore, it cannot predict the observed higher levels of responding to a higher intensity test stimulus after conditioning with a lower intensity stimulus (e.g. Ernst et al. 1970). This therefore suggests that a degree of caution should be held when applying this theory to the evidence above.

A successful interpretation of these challenging findings can be obtained by referring to the principles of stimulus generalisation and gradient interaction. According to Spence (1936, 1937), the hypothetical generalisation gradients that form around stimuli after conditioning are symmetrical because the stimuli are located on a logarithmic scale. Thus, because of this scale, generalisation from 45 cm to 15 cm will be less than between 70 cm and 100 cm. Moreover, assuming that conditioning with a single stimulus, say A, is conducted with inter-trial intervals in which no stimulus is presented, and further that stimuli during this period will enter into inhibitory associations (Mackintosh, 1974, Perkins, 1953) it follows that responding to a high-intensity stimulus will elicit higher responding due to peak shift. It therefore seems that gradient interaction theory may also be a good account for many of the asymmetries described above (although, as we have just seen, it runs into problems with result from certain experiments involving temporal

discriminations). However, as far as I am aware, no direct test has been conducted to assess whether the mechanisms of stimulus generalisation can indeed account for the asymmetries described above. Therefore, considering that this theory places a great deal of emphasis on generalisation from nonreinforced background cues other than S+ and S-, a starting point for the investigations outline in Chapters 2 – 4 is to assess the role of these nonreinforced background cues in causing the asymmetry.

It has already been mentioned that Kosaki et al. (2013) have applied gradient interaction principles in order to account for their findings in the discrimination of length. Considering that their proposals were based on data obtained in a swimming pool and related to spatial memory, the starting point for this thesis is will be to replicate the asymmetry with a different dimension of intensity. Rats are readily able to discriminate between auditory stimuli of different intensities (e.g. Pierrel, Sherman, Blue, & Hegge 1970), and thus the initial experiments presented in Chapter 2 use a appetitive Pavlovian design and different intensities of a clicker. The fact that the asymmetry has been demonstrated in many different dimensions however suggests that any stimulus might have been chosen. Successful replication of the asymmetry and assessment of the gradient interaction account will not only be a valuable addition to the limited number of demonstrations of this phenomenon, but will be an important step into understanding how intensity should be incorporated into theories of associative learning in general.

CHAPTER 2

Asymmetry in the discrimination of auditory intensity by rats

Chapter 1 reviewed a set of experiments that show discriminations between two stimuli differing in magnitude are asymmetrical. Typically, when a high-magnitude stimulus signals a reinforcer, and a low-magnitude stimulus signals the absence of a reinforcer, the discrimination is acquired more rapidly than when the reverse treatment is true.

To account for one of these asymmetries, Kosaki et al. (2013) proposed that S+ and S- were influenced to different degrees by nonreinforced background cues, S_0. In the experiments conducted by Kosaki et al., rats were trained to escape a swimming pool by locating a submerged platform that was placed in front of a long black panel, but not a short black panel, or in front of a short black panel, but not a long black panel. Either side of the panels, the S+ and S-, were areas of grey wall. Importantly, the submerged platform was never located in front of these grey areas and thus they can be considered to be S_0. Assuming that S+ and S- were differentiated by their position on a dimension, which in this case was the dimension of length, and that the background cues were represented as 'zero-magnitude' (Perkins, 1953), then the background cues can be considered to be more similar to low-magnitude stimuli than high-magnitude stimuli. A low-magnitude S+ will therefore receive more generalisation of inhibition from S_0 than a high-magnitude S+, thus making a low magnitude+/high magnitude- discrimination more difficult to acquire than a high magnitude+/low magnitude- discrimination.

A similar asymmetry has been demonstrated when S+ and S- were two different intensities of an auditory stimulus. When subjects were required to discriminate between a loud-intensity S+ and a soft-intensity S-, loud+/soft-, the

discrimination was acquired more readily than with soft+/loud-. This effect has been observed when the stimuli under scrutiny were different intensities of a white noise and the US was an electric shock (Jakubowska & Zielinski, 1976), and when the stimuli were tones which were paired with a food reward for instrumental responding (Pierrel et al., 1970). For both experiments there were periods in the experiment, either in an intertrial-interval (Jakubowska & Zielinski), or an intersession-interval (Pierrel et al.), in which no stimuli, and no US, were presented. In a similar manner to the experiments conducted by Kosaki et al. (2013), these experiments presented an opportunity for background cues, the absence of an auditory stimulus, to acquire inhibition. Thus according to Kosaki et al., the asymmetries observed by Jakubowska and Zielinksi, and Pierrel et al., are a result of the absence of an auditory stimulus being more similar to a low-intensity sound than a high-intensity sound.

Whether the proposals of Kosaki et al. (2013) can be applied to the discrimination of auditory intensity has yet, as far as I am aware, to be tested. The aim of the present chapter was therefore to assess a number of predictions that should follow from these proposals using appetitive Pavlovian conditioning in rats. Experiment 1 tested the ability of two groups of rats to discriminate between a loud and a soft clicker using an appetitive Pavlovian design. A clicker was used instead of a tone or a white noise for three reasons. First, an unreported pilot study indicated that rats found high intensity white noises aversive and were at risk from fitting. Further investigation also demonstrated that it was difficult to keep the intensity of a tone consistent within the experimental test chambers due to standing waves. Finally, there has yet, as far as I am aware, to be a demonstration of an asymmetry in the discrimination of auditory intensity using a clicker.

For one group the loud clicker, but not the soft clicker, was followed by the delivery of sucrose. For the other group this relationship was reversed so that the soft clicker, but not the loud clicker, signalled the delivery of sucrose. Consistent with what has previously been observed with discrimination based on auditory intensity (e.g. Jakubowska and Zielinski, 1976) and indeed discriminations of magnitude from other dimensions such as length (Kosaki et al., 2013), and temporal duration (e.g. Bouton & García-Gutiérrez, 2006), the loud+/soft- task was acquired more readily than the soft+/loud- task. The purpose of the remaining four experiments was then to assess the role of the cues presented during the ITI in causing the asymmetry observed in Experiment 1.

Experiment 1

Two groups of rats were given appetitive Pavlovian conditioning with two different intensities of a clicker. For one of these intensities, trials always terminated with a delivery of sucrose solution into a food well, S+. This area could be accessed by poking the snout through a hole in the chamber wall, the action of which will henceforth be referred to as a 'snout entry'. For the other intensity stimulus, trials were never followed by a delivery of sucrose, S-. Over the course of training it was anticipated that rats would make more snout entries during S+ than S-. For the loud+/soft- group the presentation of a loud clicker served as S+ whilst the presentation of a soft clicker served as S-. For the soft+/loud- group, which used the same stimuli as the first group, the presentation of the soft clicker was S+ while the loud clicker was S-. If the asymmetry identified in other discriminations of auditory intensity (e.g. Jakubowska and Zielinski, 1976, Pierrel et al., 1970) is replicable, it follows that rats in the loud+/soft- group will increase the number of snout entries

made during S+ trials and decrease the number of snout entries made during S- trials more rapidly than rats in the soft+/loud- group.

Method

Subjects. The subjects were 20 male hooded Lister rats supplied by Harlan Olac (Bicester, Oxon, UK). All rats were housed in pairs in a temperature-controlled colony room (approximately 20°C) that was continuously illuminated for 12 hours per day, with lights on at 07:00. Rats had access to water *ad libitum* but were food deprived to between 80 – 85 % of their free feeding weights prior to the start of behavioural training (M = 237g), and maintained at this weight by being fed a restricted diet after each experimental session. They were randomly assigned in equal numbers (n = 10) to the two groups at the start of the experiment.

Apparatus. Eight conditioning chambers were used. The walls (28 cm × 30 cm; H × W) and ceiling of each chamber were constructed from clear Perspex. The floor was a metal grid floor positioned 5 cm above the base of the chamber that was lined with an absorbent, odour-removing paper. In the centre of the back wall, there was a circular hole, diameter 3 cm, the centre of which was 3 cm above the grid floor. The circular hole allowed access to a well into which sucrose solution (8% sugar, 92% water) was delivered. This area is henceforth referred to as the magazine. A peristaltic pump was located beneath each conditioning chamber, which delivered the sucrose solution via a plastic tube into the well. Half of these chambers had a lever either side of the magazine, while the other half had only one lever located to the left of the magazine. These levers had no part in this experiment and lever presses were not recorded. Auditory stimuli were delivered simultaneously to all chambers from a 5-ohm, speaker located on the ceiling of each chamber. A PC with

Whisker software, and programmed in Visual Basic 6.0, controlled the experimental events and recorded the duration and number of snout entries into the magazine from infrared sensors that were set into each chamber. Barriers were placed between the chambers to prevent the animals seeing each other. Throughout all phases of the experiment the testing room lights and main computer monitor were switched off.

Stimuli. Rats were presented with a clicker at two different intensities. The average intensity across the eight chambers for what shall be referred to for simplicity as the 'soft' intensity clicker was 78.3 dB (*SD* = 3.5 dB). The average intensity for the 'loud' clicker was 87.5 dB (*SD* = 3.1 dB). The average background noise across the chambers was 62.3 dB (*SD* = 3.7 dB).

Procedure. Rats were first given two hour-long sessions in which they were trained to obtain sucrose from the magazine. In each session, 1 ml of sucrose was delivered to the magazine once a minute, every minute, for the first 30 min, followed by a 30-min period in which no sucrose was delivered. No clickers were presented during these sessions. The rats then received 14 sessions of discrimination training. Each of these sessions included eight presentations of the reinforced stimulus and eight presentations of the non-reinforced stimulus in a random order, with the constraint that a stimulus could not be presented more than twice consecutively. On each trial the stimulus was presented for 15 s and successive trials were separated by an inter-trial interval (ITI) that was either four, six or eight minutes in duration (*M* = 6 min). No stimuli were presented during the ITI. For the loud+/soft- group, the presentation of the loud, but not the soft clicker signalled the delivery of 1 ml sucrose solution at the offset of the CS. For the soft+/loud- group, the presentation of

the soft clicker signalled the delivery of 1 ml sucrose, whilst the loud clicker signalled the absence of sucrose.

Data Analysis. Individual mean rates of responding, based on the number of times in which the beam of the infra-red sensor was interrupted during a trial, were recorded for all trials in each of the 14 sessions of discrimination training. The rates of responding during a 15-s interval prior to the onset of each trial were also recorded. For all experiments presented in this thesis, the analysis of the rates of responding to reinforced and non-reinforced trials over the sessions was conducted with analyses of variance (ANOVA) using a rejection criterion of $p < .05$. For the sake of clarity, and to account for within-group variation, the raw data for each session were also transformed into discrimination ratios. This further analysis was also conducted for all experiments. The ratios were of the form $A/(A + B)$, where A and B were the mean rates of responding on reinforced trials (S+) and nonreinforced trials (S-), respectively. A ratio greater than .50 indicates that more nose pokes were made during S+ than S-. The analysis of these ratios was also conducted with analyses of variance (ANOVA) using a rejection criterion of $p < .05$. The reported effect size for ANOVA with more than one factor is partial eta squared (η_p^2), while for comparisons between two means it is eta squared (η^2).

Results

The mean rates of responding to S+, S- and the 15-s pre-CS period of each of the 14 experimental sessions are presented in the upper panels of Figure 8. Rats in both groups came to make more snout entries during the presentation of S+ than S- although the magnitude of this difference was greater for the loud+/soft- group than the soft+/loud- group.

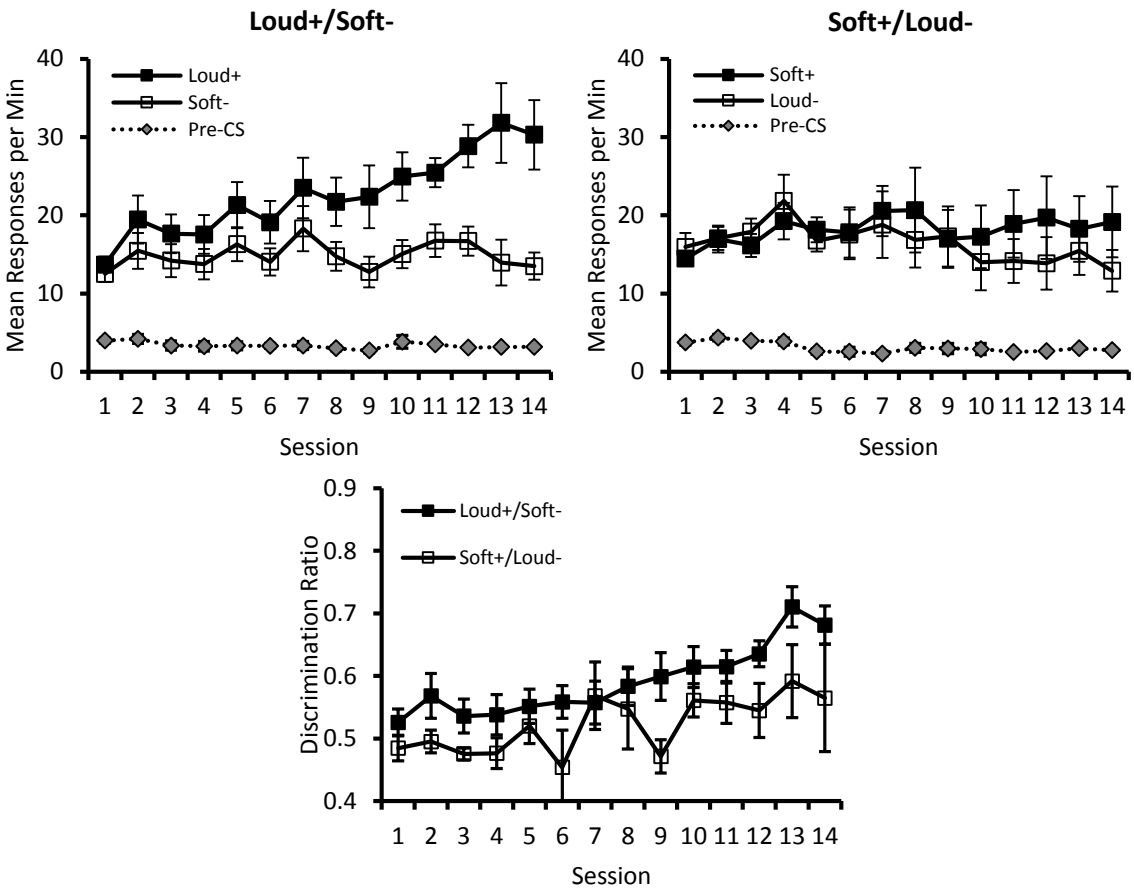

Figure 8. The upper panels present the mean rates of responding to S+ and S- for the 14 sessions of training for the loud+/soft- (left-hand panel) and soft+/loud- groups (right-hand panel) of Experiment 1. The lower panel presents the discrimination ratios for the two groups. Error bars represent ± 1 *SEM*.

Analysis of the mean rates of responding using a 3-way ANOVA with factors, trial-type (S+ vs. S-), group (loud+/soft- vs. soft+/loud-) and session did not reveal a significant 3-way interaction, $F(13, 234) = 1.35$, $p > .10$, but importantly did show a significant Trial-type × Group interaction, $F(1, 18) = 11.53$, $p = .003$, $\eta_p^2 = .39$, supporting the observation that there was an asymmetry in the acquisition of the discrimination. Further investigation of this significant interaction using tests of simple main effects revealed a significant effect of trial-type for the loud+/soft-

group, $F(1, 18) = 37.67$, $p < .001$, $\eta_p^2 = .68$, but not the soft+/loud- group, $F(1, 18) = 1.78$, $p > .10$. There was also no effect of group for the rate of responding during either the S+ or the S-, Fs$(1, 36) < 2.51$, ps $> .10$. The remaining results from the 3-way ANOVA were a significant main effect of trial-type, $F(1, 18) = 27.92$, $p < .001$, $\eta_p^2 = .61$, but the main effects of group, $F < 1$, and session, $F(13, 234) = 1.18$, $p > .10$, and the Session × Group interaction, were not significant, $F(13, 234) = 1.37$, $p > .10$.

Figure 8 also presents the mean rate of responding during the 15-s pre-CS period. A 2-way ANOVA revealed an overall main effect of session, $F(13, 234) = 2.07$, $p = .017$, $\eta_p^2 = .10$, but the effect of group and Group × Session interaction were not significant, Fs < 1.

The lower panel of Figure 8 presents the discrimination ratios for the two groups. Analysis of these ratios using a 2-way, Group × Session ANOVA revealed a significant main effect of group, $F(1, 18) = 10.13$, $p = .005$, $\eta_p^2 = .36$, confirming the observation that the loud+/soft- group acquired the discrimination more rapidly than the soft+/loud- group. This analysis also revealed a significant effect of session, $F(13, 234) = 3.30$, $p < .001$, $\eta_p^2 = .15$, but the interaction between these factors was not significant, $F < 1$.

Discussion

The results demonstrate a clear asymmetry in the acquisition of a discrimination between a loud and soft clicker. Rats in the loud+/soft- group acquired the discrimination more rapidly than rats in the soft+/loud- group, an observation consistent with other experiments in which rats were required to

discriminate between two auditory stimuli differing in intensity (Jakubowska and Zielinski, 1976, Pierrel et al., 1970). By using a clicker as the cue that varied in magnitude for the first time, these findings therefore extend the generality of results reported from other magnitude dimensions such as panel length (Kosaki, Jones, & Pearce, 2013), quantity of dots (Vonk & Beran, 2012; Watanabe, 1998), and temporal duration (Kyd et al., 2007, Bouton & García-Gutiérrez, 2006).

The purpose of Experiments 2 to 5 is to investigate the role of the ITI in causing this asymmetry. If the asymmetry is a result of generalisation from nonreinforced cues presented during the ITI then it follows that if the experimental stimuli are presented without an ITI then this source of generalisation will be removed and the asymmetry observed in Experiment 1 should be reduced, or indeed, completely abolished. Experiments 2 and 3 were conducted in order to directly test this prediction.

Experiment 2

The most direct way of testing the foregoing prediction would be to repeat Experiment 1, say, with the ITI removed and, therefore, with alternating 15-s presentations of S+ and S-. A preliminary investigation revealed that this method of training did not result in the successful acquisition of the discrimination. After further investigation, the following technique was adopted (see Figure 9). Four groups of rats were trained to discriminate between a loud and soft clicker of 73-s duration. For the two ITI groups, successive presentations of S+ and S- were separated by an ITI of 73 s in which no clickers were presented. For the loud+/soft-/ITI group the presentations of a loud clicker, S+, coincided with three randomly determined deliveries of food whilst no food was ever delivered during the

presentations of the soft clicker, S-. For the soft+/loud-/ITI group this reward contingency was reversed so that the soft clicker was S+ and the loud clicker was S-. For the two groups trained without an ITI, the loud+/soft-/no ITI group and soft+/loud/no ITI group, rats received alternating presentations of S+ and S- without an ITI. To equate session duration across groups the duration of the S- for no ITI groups was three times that of S+.

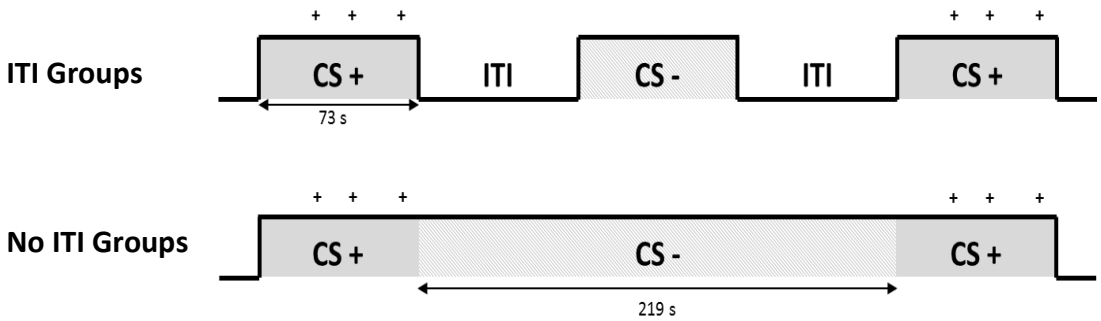

Figure 9. Design of Experiment 2. The symbol + represents the delivery of a single food pellet.

Subjects. The subjects were 32 male hooded Lister rats that were from the same supplier and housed in the same manner as those of Experiment 1. They were food deprived to between 80 – 85% of their free feeding weights (*M* = 264g) prior to the start of behavioural training and maintained at this weight for the duration of the experiment by being fed a restricted diet after each session. The rats were randomly assigned in equal numbers (*n* = 8) to each of the four groups.

Apparatus and Stimuli. Eight operant chambers different to those used in Experiment 1 were used for all stages of this experiment (23.0 × 24.5 × 21.0 cm, L ×

W × H). Each was housed in an individual sound- and light-attenuating chamber that was closed for the duration of each session. The chambers were constructed from three aluminium walls, an aluminium ceiling and a transparent Perspex door. A series of 16 stainless steel rods, 0.5 cm in diameter and 1.5 cm apart (centre – to – centre), served as the chamber floor for four of these chambers while a wire grid served as the floor for the remaining four chambers. Below these floors was a tray lined with absorbent, odour – removing paper. There was a 5.0 × 6.0 cm recessed food magazine in the front wall of the box into which 45-mg food pellets (traditional formula, P.J. Noyes, Lancaster NH) could be delivered. The base of the magazine was 0.5 cm above the floor. A clear Perspex flap which was hinged at the top covered the entrance to the magazine and could be pushed back to gain access. The number of snout entries into the food well was recorded by an infra-red sensor that was set into a rectangular frame that surrounded the entrance to the magazine. A PC with Whisker software, and programmed in Visual Basic 6.0 controlled experimental events and recorded the number of snout entries. Two loudspeakers in the ceiling of the chambers delivered two different intensity of clicker. The stimuli, which shall be referred to as the 'soft' and the 'loud' clicker were 57.5 dB (*SD* = 1.8 dB) and 82.2 dB (*SD* = 2.5 dB) respectively. The level of background noise in the chambers when no clicker was being presented was approximately 50 dB.

Procedure and Data Analysis. Rats were first given two 30-min sessions in which they were trained to obtain food pellets from the magazine. In each session a single food pellet was delivered to the magazine once a minute, every minute. In the first of these sessions, the flaps covering the magazine were taped open and two pellets were placed at the entrance of the magazine to encourage exploration. For the

second session the flaps were not taped open and no food additional pellets were placed in the magazine. No auditory stimuli were presented during this training. The rats were then given eight sessions of discrimination training in which ten loud and ten soft clickers were presented in an alternating order. Each clicker was presented continuously for a total of 73 s. For rats in the ITI groups, during the presentation of a S+ trial rats were rewarded at a randomly selected time point within 3 successive 20-s periods after an initial 10-s period in which no food pellets were delivered. The initial trial-type was chosen pseudo randomly but was identical for all groups at each session. Each 20-s period was separated by a 1-s interval in order to ensure a minimum separation period between successive food deliveries. Trials were separated by a 73-s ITI in which no clicker was presented. For rats in the no-ITI groups, the termination of one trial-type was followed immediately by the onset of the next. To control for subsequent differences in session duration and frequency of reinforcement, the duration of S- trials in the no-ITI groups was equivalent to two ITI periods and a single S- trial (219 s) in the ITI groups (See Figure 9).

The number of snout entries prior to the first delivery of a food pellet (pre-US) was recorded for every S+ trial. The timer that was in operation during S+ to determine when a food pellet was first delivered on a trial, was also in operation from the onset of each S-, but its purpose was solely to identify an interval during which responding would be recorded for the nonreinforced trials. The number of snout entries made throughout the ITI was also recorded for the two ITI groups.

Results

The results from the eight sessions of discrimination training can be seen in Figure 10. The top two panels indicate that when trials were separated by an ITI the loud+/soft- group acquired its discrimination more readily than the soft+/loud- group. A similar pattern of results can also be seen for the two groups in which trials were not separated by an ITI (lower panels).

A 4-way ANOVA with the factors of ITI (present or absent), group (loud+/soft- vs. soft+/loud-), trial-type, and session revealed a Trial-type × Group interaction, $F(1, 28) = 23.72$, $p < .001$, $\eta_p^2 = .46$, supporting the observation of an asymmetry in the acquisition of the discriminations. Test of simple effects revealed that there was a significant effect of group for the S+, $F(1, 56) = 12.56$, $p = .001$, $\eta_p^2 = .18$, but not the S-, $F < 1$. The effect of trial-type was significant for loud+/soft- and soft+/loud- groups, Fs$(1, 28) > 21.87$, ps $< .001$.

The remaining results from the 4-way interaction were a significant Session × ITI interaction, $F(7, 196) = 2.91$, $p = .006$, $\eta_p^2 = .09$, Trial-type × Session interaction, $F(7, 196) = 5.27$, $p < .001$, $\eta_p^2 = .16$, Session × ITI × Group interaction, $F(7, 196) = 3.00$, $p = .005$, $\eta_p^2 = .10$, Trial-type × Session × ITI interaction, $F(7, 196) = 2.26$, $p = .031$, $\eta_p^2 = .07$, Trial-type × ITI interaction, $F(1, 28) = 4.26$, $p = .048$, $\eta_p^2 = .13$, and main effects of trial-type, $F(1, 28) = 131.82$, $p < .001$, $\eta_p^2 = .82$, session, $F(7, 196) = 2.80$, $p = .009$, $\eta_p^2 = .09$, and ITI, $F(1, 28) = 8.34$, $p = .007$, $\eta_p^2 = .23$. The Trial-type × ITI × Group, $F < 1$, 4-way interaction, $F(7, 196) = 1.51$, $p > .10$, ITI × Group interaction, $F < 1$, Session × Group interaction, $F(7, 196) = 1.13$, $p > .10$, and Trial-type × Session × Group interaction, $F(7, 196) = 1.28$, $p > .10$, and the main effects of group, $F(1, 28) = 2.23$, $p > .10$, did not reach the accepted level of significance.

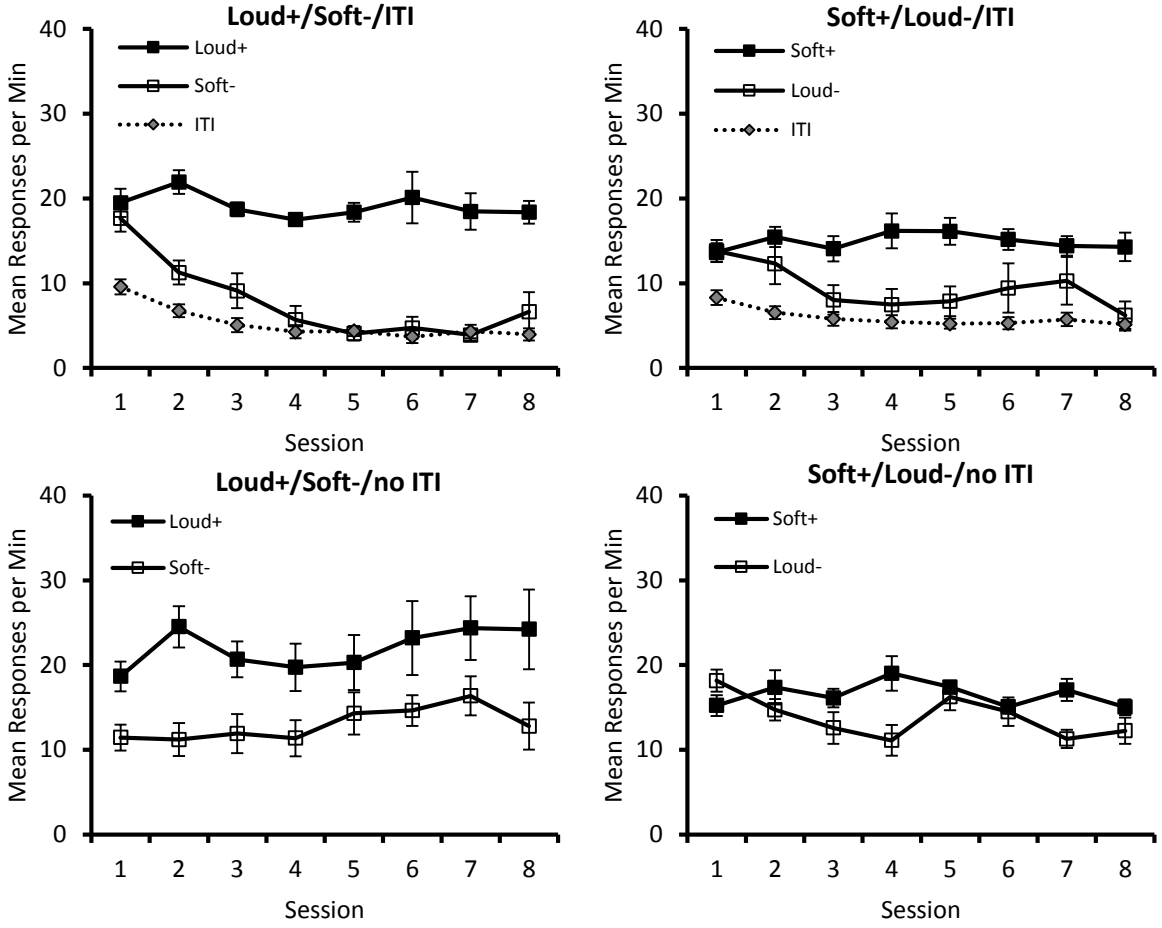

Figure 10. The mean rates of responding to S+ and S- for the eight sessions of training. The upper panels represent the two groups in which trials were separated by an ITI. Lower panels represent the two groups in which trials were presented successively without an ITI. Error bars indicate ± 1 *SEM*.

The upper panels of Figure 10 also present the mean rates of responding during the ITI. A 2-way ANOVA revealed a significant main effect session, $F(7, 98) = 15.99$, $p < .001$, $\eta_p^2 = .53$, but the main effect of group, $F < 1$, and Group × Session interaction, $F(7, 98) = 1.61$, $p > .10$, were not significant.

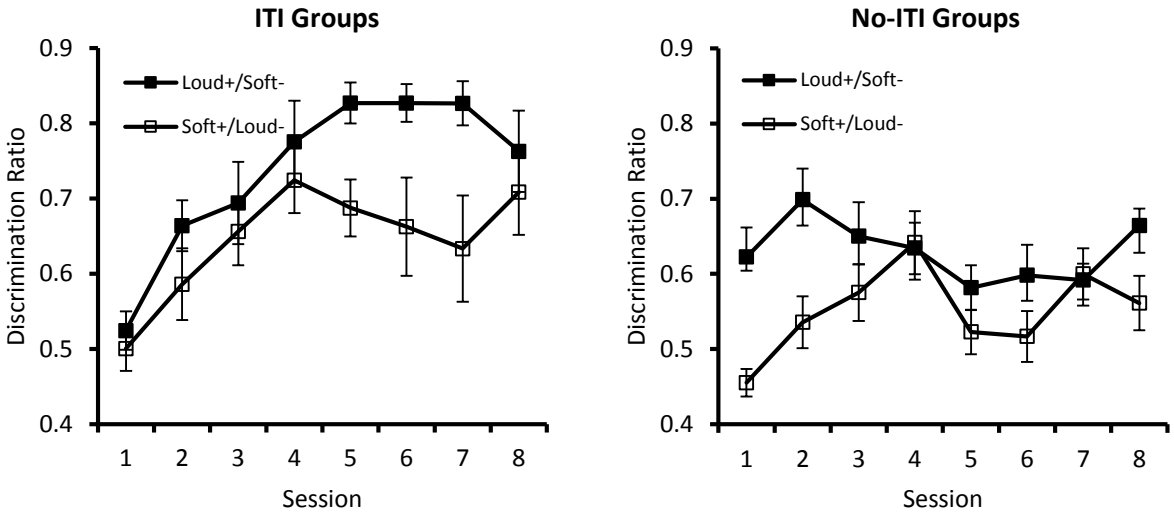

Figure 11. Discrimination ratios across the 8 sessions of Experiment 2 for groups with an ITI (left-hand panel) and for groups without an ITI (right-hand panel).

The discrimination ratios presented in Figure 11 are consistent with the results from the acquisition data. Both panels indicate that the loud+/soft- groups had higher discrimination ratios than the soft+/loud- groups. A 3-way ANOVA conducted for these ratios revealed a significant main effect of group, $F(1, 28) = 12.65$, $p = .001$, $\eta_p^2 = .31$, confirming this observation, and a significant 3-way interaction, $F(7, 196) = 2.54$, $p = .016$, $\eta_p^2 = .08$. Further analysis of simple effects for this significant interaction revealed a significant ITI × Session interaction for loud+/soft- groups, $F(7, 196) = 6.85$, $p < .001$, $\eta_p^2 = .20$, but not for soft+/loud- groups, $F(7, 196) = 1.03$, $p > .10$. The Group × Session interaction was not significant for the ITI groups, $F(7, 196) = 1.66$, $p > .10$, and no-ITI groups, $F(7, 196) = 1.80$, $p = .090$. An effect of group was observed for the ITI groups at Sessions 6 and 7, Fs$(1, 224) > 8.10$, ps $< .005$, $\eta_p^2 > .03$, and for the no-ITI groups at Sessions 1 and 2, Fs$(1, 224) > 7.73$, ps $< .006$, $\eta_p^2 > .03$. The remaining findings from 3-way ANOVA were a significant main effect of ITI, $F(1, 28) = 17.39$, $p < .001$, $\eta_p^2 = .38$, a significant main effect of session, $F(7, 196) = 8.53$, $p < .001$, $\eta_p^2 = .23$, and a

significant Session × ITI interaction, $F(7, 196) = 5.39$, $p < .001$, $\eta_p^2 = .17$. The Session × Group interaction and ITI × Group interactions were not significant, Fs < 1.

Discussion

The purpose of this experiment was to test the prediction that presenting stimuli successively without an ITI would remove the asymmetry effect. Consistent with the proposals of Kosaki et al., (2013), when trials were separated by an ITI, an asymmetry was observed in the acquisition of the discriminations. However, when stimuli were not separated by an ITI rats in the loud+/soft-/no ITI group acquired the discrimination more readily than the soft+/loud-/ no ITI group. This result therefore contradicts an account of the asymmetry in terms of gradient interaction as it suggests that the effect is not dependent on the generalisation of inhibition from cues presented during the ITI.

One possible explanation is that the asymmetry seen in the no-ITI groups is a result of the elongated exposure by this group to the S-. The groups trained in the no-ITI condition received three times as much exposure to the S- than groups in the ITI condition. It is therefore conceivable that the excessive exposure to S- by this group may have facilitated the discrimination by the loud+/soft-/no ITI group or hindered the acquisition of the soft+/loud-/no ITI group. The fact that there is no clear theoretical reason why this excessive exposure to S- should have this effect makes this explanation seem unlikely. Nevertheless, the purpose of Experiment 3 was to address this methodological shortcoming by equating S+ and S- durations. If the asymmetry is a result of the elongated S- then we might expect the asymmetry seen in the no ITI groups in Experiment 2 to disappear. However, if the asymmetry

remains it seems likely that some mechanism other than stimulus generalisation is responsible for this effect.

Experiment 3

The experiment contained two groups, a loud+/soft-/ITI group and soft+/loud-/ITI group, who received the same training as their namesakes in Experiment 2. Two further groups, a loud+/soft-/no-ITI group and soft+/loud-/no-ITI group, received training where alternating presentations of S+ and S- were not separated by an ITI. To control for the possible effects of the elongated S- period in Experiment 2, the duration of S+ and S- trials were equated (see Figure 12). According to the proposals put forward by Kosaki et al. (2013), and based on the outcome of Experiment 2, rats in the loud+/soft-/ITI group should acquire the discrimination more rapidly than those in the soft+/loud-/ITI group. In contrast, when trials are no longer separated by an ITI, it follows that both groups should acquire the discrimination at the same rate.

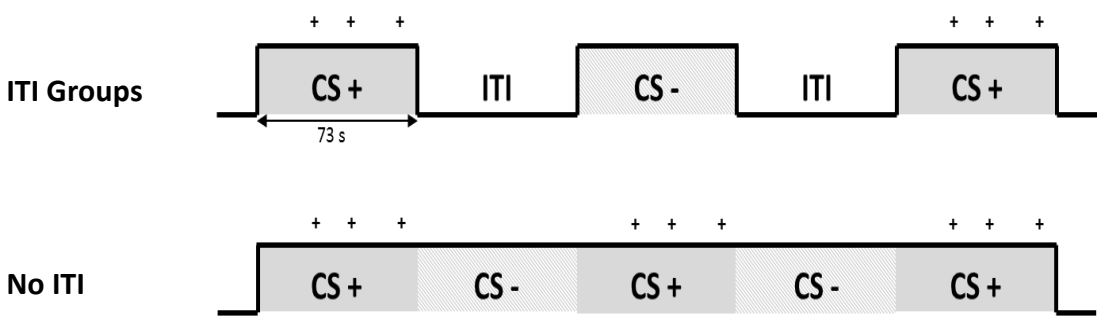

Figure 12. Design of Experiment 3. The symbol + represents the delivery of a single food pellet.

Method

Subjects, apparatus and procedure. The subjects were 32 naive male hooded Lister rats from the same supplier and housed in the same manner as Experiments 1 and 2. They were food deprived to, and maintained between 80 – 85% of their free feeding weights (*M* = 387 g). The rats were assigned randomly and in equal numbers to the groups (*n* = 8) prior to the start of the experiment. The chambers were the same as for Experiment 2. The mean intensity across the eight chambers for the soft and loud clickers was 58.1 dB (*SD* = 1.8 dB) and 82.0 dB (*SD* = 1.5 dB) respectively. The rats first received two sessions of pre-training, the method of which was identical to that described in Experiment 2. They then received six sessions of discrimination training.[3] For rats in the loud+/soft-/ITI and soft+/loud-/ITI groups, training was the same as described for their namesakes in Experiment 2. For the no ITI groups rats were given 20 trials of alternating presentations of S+ and S- without an ITI. The duration of both trial-types was 73 s. During S+, a single food pellet was delivered at a randomly selected time within three consecutive 20 s periods which followed an initial 10-s period in which no food was delivered. Each 20 s period was followed by 1 s in which no food was delivered. In total, each S+ was 73 s in duration and involved the delivery of three food pellets. No food was delivered during S-. Because the ITI was removed from these groups, the session duration for no ITI groups was half that of the ITI groups.

[3] Unlike Experiment 2, the present experiment presents six, rather than eight sessions of training as a result of a computer failure during Session 7.

Results

Across the six sessions of training the loud+/soft-/ITI group acquired the discrimination more readily than the soft+/loud-/ITI group (top panels; Figure 13). The lower panels of Figure 13 show that a similar pattern of results was obtained by the no-ITI groups.

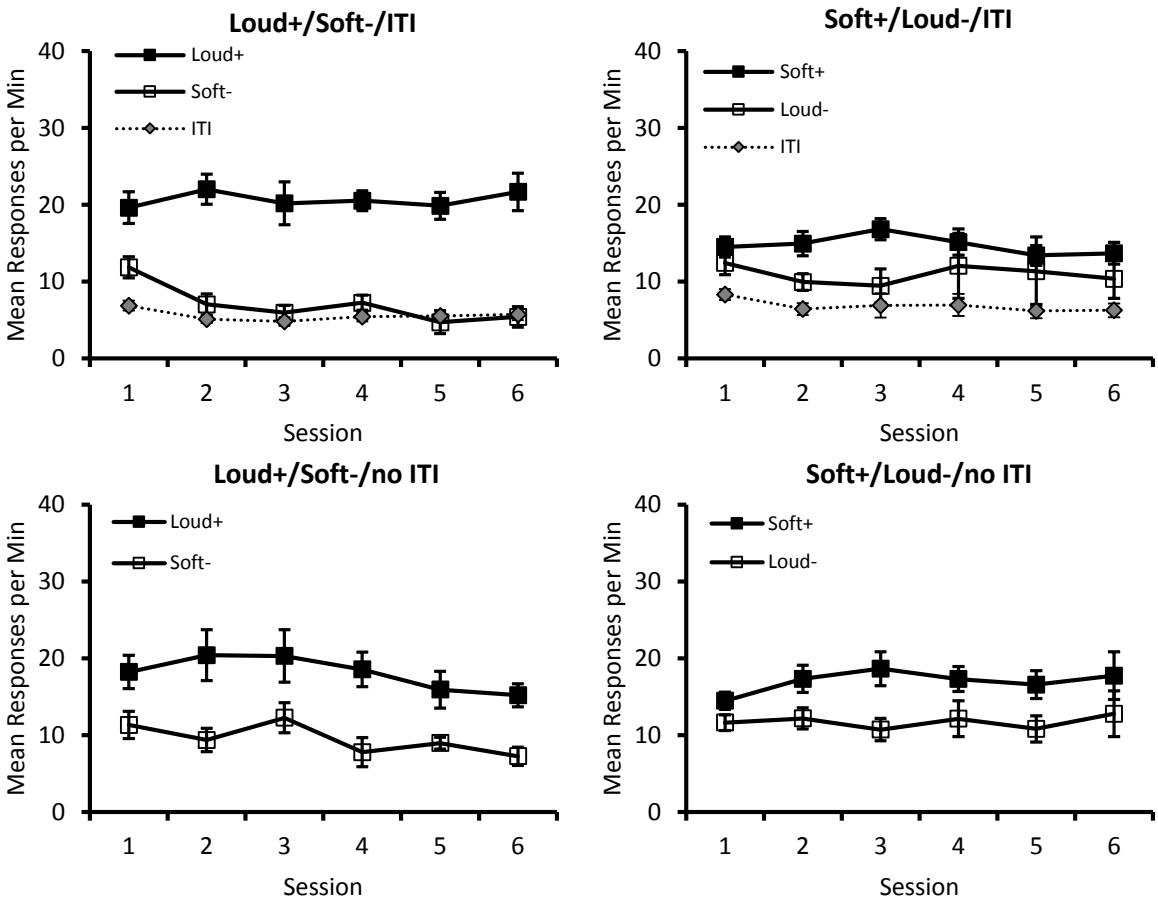

Figure 13. The mean rates of responding to S+ and S- for the six sessions of training. The upper panels represent the two groups in which trials were separated by an ITI. Lower panels represent the two groups in which trials were presented successively without an ITI. Error bars indicate ± 1 *SEM*.

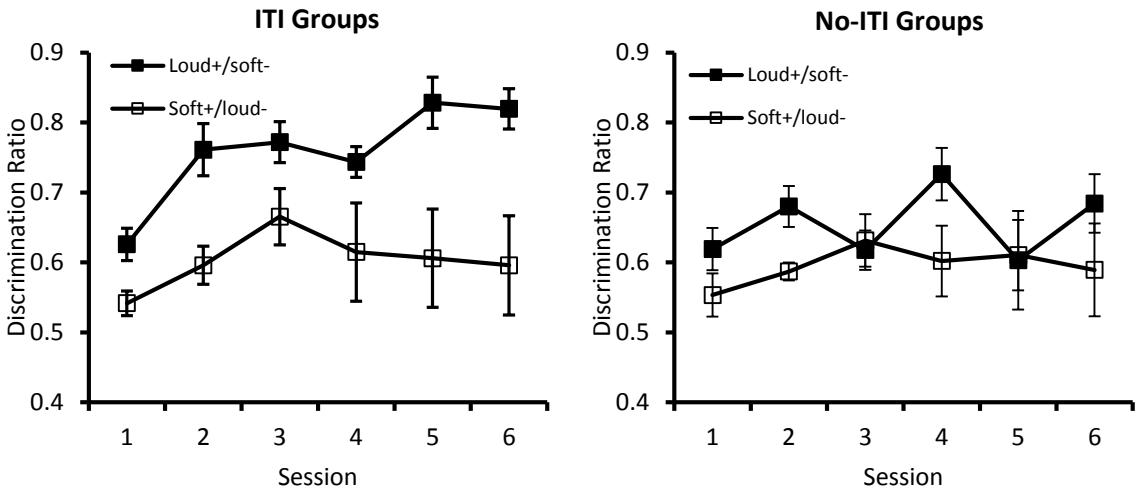

Figure 14. Discrimination ratios across the 6 sessions of Experiment 3 for groups with an ITI (left-hand panel) and for groups without an ITI (right-hand panel).

A 4-way ANOVA was conducted to test these observations. A significant Trial-type × Group interaction was revealed, supporting the observed asymmetry, $F(1, 28) = 13.55$, $p = .001$, $\eta_p^2 = .33$, but the Trial-type × ITI × Group interaction fell short of significance, $F(1, 28) = 3.34$, $p = .078$. Further investigation of the significant Trial-type × Group interaction revealed a significant effect of trial-type for loud+/soft- groups, $F(1, 28) = 77.91$, $p < .001$, $\eta_p^2 = .74$, and soft+/loud- groups, $F(1, 28) = 13.12$, $p = .001$, $\eta_p^2 = .32$. There was also an effect of group for S+, $F(1, 56) = 5.68$, $p = .021$, $\eta_p^2 = .09$, and for S-, $F(1, 56) = 4.36$, $p = .041$, $\eta_p^2 = .07$. The main effect of trial-type, $F(1, 28) = 77.50$, $p < .001$, $\eta_p^2 = .73$, and Trial-type × Session interaction, $F(5, 140) = 2.37$, $p = .042$, $\eta_p^2 = .08$, were also significant. The remaining effects of session, $F(5, 140) = 1.32$, $p > .10$, group , $F < 1$, and ITI, $F < 1$, and ITI × Group interaction, $F < 1$, Trial-type × ITI interaction, $F < 1$, Session × ITI interaction, $F < 1$, Session × Group interaction, $F(5, 140) = 1.11$, $p > .10$, Session × ITI × Group interaction, $F < 1$, Trial-type × Session × ITI interaction, $F < 1$, Trial-

type × Session × Group interaction, $F < 1$, and 4-way interaction, $F < 1$, were not significant.

An additional ANOVA was conducted on the mean rate of responding during the ITI. This analysis revealed a significant main effect of session, $F(5, 70) = 3.76$, $p = .004$, $\eta_p^2 = .21$, but the effect of group, $F(1, 14) = 1.33$, $p > .10$, and interaction were not significant, $F < 1$.

Figure 14 presents the discrimination ratios for the ITI and no-ITI groups. The left-hand panel indicates that the ratios were higher for the loud+/soft-/ITI group than the soft+/loud-/ITI group. A similar pattern is shown for the no-ITI groups, although the ratios for the loud+/soft- and soft+/loud- groups were similar at sessions 3 and 5. A 3-way ANOVA with factors ITI, group and session revealed a significant main effect of group, $F(1, 28) = 14.20$, $p = .001$, $\eta_p^2 = .34$, supporting the observation that the loud+/soft- groups acquired the discrimination more readily than the soft+/loud- groups. The main effect of session was also significant, $F(5, 140) = 3.53$, $p = .005$, $\eta_p^2 = .11$. The remaining findings of the analysis were that the 3-way interaction, $F(5, 140) = 1.32$, $p > .10$, Session × Group interaction, $F(5, 140) = 1.27$, $p > .10$, Session × ITI interaction, $F(5, 140) = 1.51$, $p > .10$, ITI × Group interaction, $F(1, 28) = 2.82$, $p > .10$, and main effect of ITI, $F(1, 28) = 3.80$, $p = .061$, were not significant.

Discussion

The results from the two ITI groups were as anticipated. Rats in the loud+/soft- group were observed to acquire the discrimination more readily than those in the soft+/loud- group. An analysis of the first six sessions revealed an

overall asymmetry across both the ITI, and no-ITI groups, a result consistent with those from Experiment 2.

The fact that removal of the ITI did not disrupt the asymmetry suggests that the asymmetry observed in Experiment 1, and in the ITI groups of Experiments 2 and 3, cannot be attributed to the generalisation of inhibition from background cues. These data therefore suggest that a degree of caution should be held when applying the proposals of Kosaki et al (2013) beyond discriminations based on length. Discussion about why this asymmetry might yet occur despite the absence of background cues will be withheld until Chapter 5.

The purpose of Experiments 4 and 5 was to examine a second prediction that follows from the principles of stimulus generalisation. In Experiment 1, and for the ITI groups in Experiments 2 and 3, the cues presented during the ITI were less intense than S+ and S-. According to the proposals of Kosaki et al (2013), if the intensity of nonreinforced background cues presented in the ITI is greater than the intensity of S+ and S- then the asymmetry should be reversed (see also Logan, 1954). This is because unlike Experiments 1 to 3 the ITI will be more similar to the high-intensity stimulus than the low-intensity stimulus and thus the high-intensity clicker will receive more generalisation of inhibition from the ITI than the low-intensity clicker.

Experiment 4

A simple way to test the aforementioned prediction would be to use the same stimuli as Experiments 1 to 3, but to change the role of the stimuli so that the loud clicker is presented during the ITI, and the absence of a clicker is used as a stimulus. In this instance it is the high-intensity stimulus, the soft clicker, which is more similar to the ITI than the low-intensity stimulus, the absence of a clicker. Two groups of rats were therefore trained to discriminate between a soft clicker and the absence of a clicker using, for the sake of consistency, the experimental design used for the ITI groups from Experiments 2 and 3 (see Figure 15). For both groups successive presentations of S+ and S- were separated by an ITI in which a loud clicker was presented throughout. For the soft+/no clicker- group the presentations of a soft clicker was assigned as S+ while an equivalent duration in which no clicker was presented was S-. For the no clicker+/soft- group this was reversed so that the absence of a clicker was S+ and the soft clicker was S-. If the proposals of Kosaki et al (2013; see also Logan, 1954) are correct, then it follows that the no clicker+/soft- group should acquire the discrimination more rapidly than the soft+/no clicker- group.

Figure 15. Design of Experiment 4. The symbol + represents the delivery of a single food pellet. For the soft+/no clicker- group S+ represents a soft clicker and S- represents the absence of a clicker For the no clicker+/soft- group, S+ represents the absence of a clicker and S- represents a soft clicker. The ITI was a loud intensity clicker.

Method

Subjects. The subjects were 32 male hooded Lister rats (*M* = 239 g) from the same supplier and housed in the same way as described in Experiment 1. They were reduced to between 80 – 85% of their free feeding weights and maintained at this level by being fed a restricted diet after every session. At the start of the experiment the rats were allocated randomly and in equal numbers to the two groups (*n* = 16).

Apparatus and Stimuli. The experiment was conducted in the same conditioning chambers used in Experiments 2 and 3. The stimulus, which shall be referred to as the 'soft' clicker, (*M* = 59.6 dB, *SD* = 2.1 dB), and the 'loud' clicker presented during the ITI (*M* = 81.4 dB, *SD* = 7.7 dB) were presented to the rats via speakers in the ceiling of the chambers. The level of auditory intensity provided by background noise during 'no clicker' trials was measured to be 49.0 dB (*SD* = 1.7 dB). A 2-kHz tone with a mean intensity of 70.0 dB was also presented throughout the initial two sessions of pre-training in order to prevent the absence of an auditory stimulus acquiring any association with food.

Procedure and Data Analysis. The rats were first given two sessions of pre-training with the same procedure as described in Experiment 1 with the exception that a 2-kHz tone was presented continuously throughout each session. They were then given six sessions of discrimination training in which ten loud and ten soft clickers were presented successively. During the presentation of a S+ trial rats were rewarded at a randomly selected time point within 3 successive 20-s periods after an initial 10-s period in which no food pellets were delivered. The initial trial-type, S+ or S-, was chosen pseudo randomly but was identical for all groups at each session. Each 20-s period was separated by a 1-s interval in order to ensure a minimum

separation period between successive food deliveries. For group soft+/no clicker- S+ was the soft clicker and S- was the absence of a clicker. For group no clicker+/soft- S+ was the absence of a clicker and S- was the soft clicker. Thus, S+ trials were either the soft clicker presented continuously for a total of 73 s, or the absence of a clicker for the same duration. Trials were separated by a 73-s ITI in which the loud clicker was presented. The duration of non-reinforced trials was also 73 s. Data were recorded and analysed in the manner described in Experiments 2 and 3.

Results

Inspection of the top panels of Figure 16 shows that rats in both groups came to respond more rapidly in the presence of the S+ than the S- over the six sessions of training. The soft+/no clicker- group acquired the discrimination more readily than the no clicker+/soft- group, but this observation was not supported by the statistical analysis

.

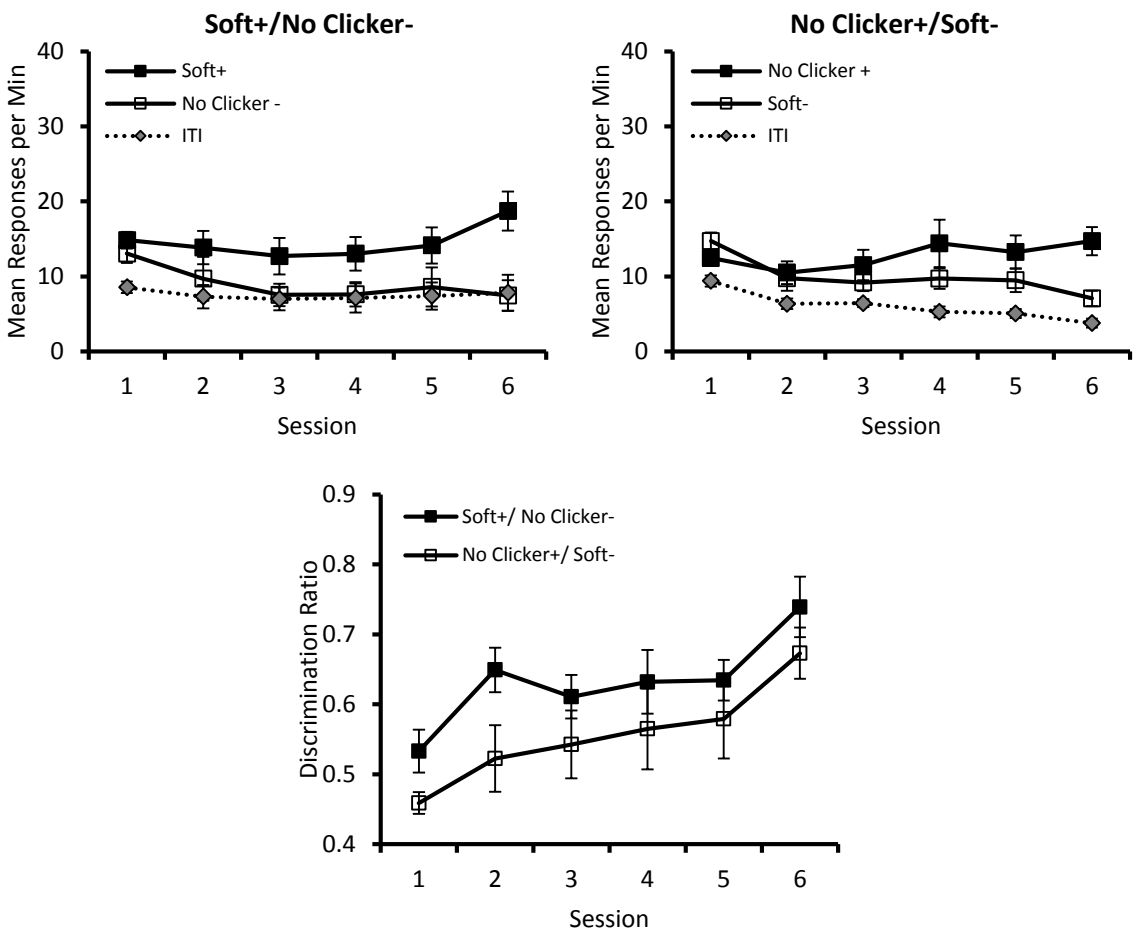

Figure 16. The mean rates of responding to S+ and S- for the six sessions of training (upper panels). The lower panel presents the discrimination ratios for the two groups. Error bars indicate ± 1 *SEM*.

A 3-way ANOVA with the factors group (soft+/no clicker- vs. no clicker+/soft-), trial-type and session revealed a significant main effect of trial-type, $F(1, 14) = 25.18$, $p < .001$, $\eta_p^2 = .64$, confirming the observation that the groups came to respond more to S+ than S-, but the Trial-type × Group interaction, $F(1, 14) = 2.71$, $p > .10$, and 3-way interaction, $F < 1$, were not significant. The ANOVA also revealed a significant Trial-type × Session interaction, $F(5, 70) = 8.86$, $p < .001$, $\eta_p^2 = .39$, but the main effect of session, $F(5, 70) = 2.12$, $p = .073$, and group, $F < 1$, and Session × Group interaction were not significant, $F < 1$.

The upper panels of Figure 16 also show that the mean rate of responding during the ITI was approximately the same as for S-. An ANOVA on this data revealed a significant effect of session, $F(5, 70) = 4.10$, $p = .003$, $\eta_p^2 = .23$, but the effect of group, $F < 1$, and Session × Group interaction, $F(5, 70) = 2.23$, $p = .06$ were not significant.

The lower panel of Figure 16 shows that discrimination ratios for the soft+/no clicker- group were larger for all sessions than the no clicker+/soft- group. A 2-way ANOVA of the mean discrimination ratios revealed a significant main effect of session, $F(5, 70) = 8.06$, $p < .001$, $\eta_p^2 = .37$. The main effect of group was marginally significant, $F(1, 14) = 3.81$, $p = .071$, and the interaction between these factors did not reach significance, $F < 1$.

Discussion

The results demonstrate that both groups of rats were able to successfully discriminate between the absence of a clicker and a soft clicker when trials were separated by presentations of a high intensity clicker. An inspection of the discrimination ratios presented in Figure 16 shows that the soft+/no clicker- discrimination may have been acquired more rapidly than the no clicker+/soft- discrimination. This result is surprising because it is the direct opposite of what is predicted from the proposals of Kosaki et al. (2013). According to these proposals, on a dimension of auditory intensity the loud clicker presented during the ITI can be considered as more similar to the soft clicker than to the absence of a clicker. Thus generalisation of inhibition from the ITI should mask the discriminatory performance of the soft+/no clicker- group more so than the no clicker+/soft- group. However the

difference in discrimination ratios seen in Figure 16 fell short of the accepted level of significance.

One problem with this experiment that might account for the pattern of results is that one of the conditioned stimuli was the absence of a clicker, which makes the design of this experiment conceptually different from other demonstrations of the asymmetry in which the discriminanda have always been two physical stimuli. Experiment 5 was conducted in an attempt to replicate Experiment 4 with the exception that rats were required to discriminate between a soft- and medium-intensity clicker rather than a soft-clicker and no-clicker stimuli.

Experiment 5

To address the shortcomings of Experiment 4, two groups of rats were trained to discriminate between a soft-, and medium-intensity clicker. A loud-intensity clicker was presented continuously throughout each ITI. For the soft+/medium- group the presentation of the soft-, but not the medium-intensity clicker signalled the delivery of a food reward. For the medium+/soft- group the presentation of the medium-intensity, but not the soft-intensity clicker signalled the reward. In all other respects the procedure for discrimination training was the same as for Experiment 4.

Method

Subjects, apparatus, and procedure. The subjects were 16 male Lister hooded rats that had previously been trained to discriminate between two different quantities of black dots (see Chapter 3, Experiment 7). The previous experiment was conducted in a different test room, with a different style of chamber, and with a

different reward (sucrose) to the experiment reported here. The rats were housed in the manner described in Experiment 1 and the experiment was conducted in the chambers described in Experiment 2. They were first given two sessions of pre-training in which they were trained to obtain food pellets from the magazine. Unlike Experiment 4 a tone was not presented during this training although in all other respects the procedure details were the same. They were then given six sessions of discrimination training in which they were trained to discriminate between a soft intensity clicker (M = 49.7 dB, SD = 0.8 dB) and medium intensity clicker (M = 64.2 dB, SD = 1.1 dB). Each session consisted of 24 trials. A loud intensity clicker (M = 82.8 dB, SD = 9.6 dB) was presented throughout each ITI. Any omitted procedural details were the same as those of Experiment 4.

Results

Rats in the medium+/soft- group acquired the discrimination more readily than those in the soft+/medium- group (see the upper panels of Figure 17). A 3-way ANOVA with the factors of trial-type, group and session did not however confirm this observation as the 3-way interaction, $F(5, 70) = 1.47$, $p > .10$, and Trial-type × Group interaction, $F(1, 14) = 3.65$, $p = .077$, were not significant. The analysis also revealed a significant main effect of trial-type, $F(1, 14) = 108.84$, $p < .001$, $\eta_p^2 = .89$, and session, $F(5, 70) = 3.53$, $p = .007$, $\eta_p^2 = .20$, and a significant Trial-type × Session interaction, $F(5, 70) = 9.68$, $p < .001$, $\eta_p^2 = .40$. The Session × Group interaction, $F(5, 70) = 1.44$, $p > .10$, and main effect of group, $F < 1$, did not reach significance.

Figure 17 also indicates that the mean rate of responding during the ITI was low and decreased over the six sessions. This observation was supported by a Group × Session ANOVA which revealed a significant main effect of session, $F(5, 70) = 8.44$, $p < .001$, $\eta_p^2 = .38$. The effect of group and Group × Session interaction were not found to be significant, Fs < 1.

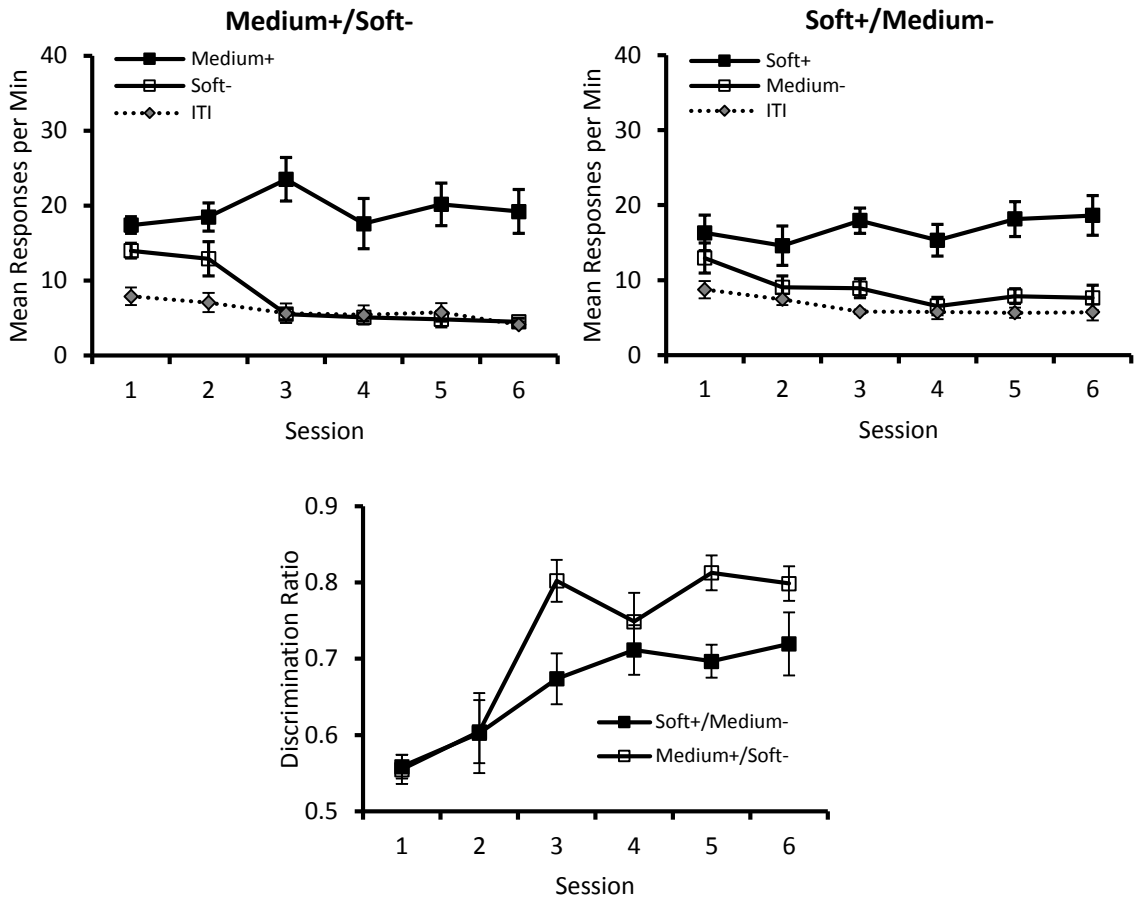

Figure 17. The mean rates of responding to S+ and S- for the six sessions of training for the two groups of Experiment 5 (top row). The lower panel presents the discrimination ratios for both groups. Error bars indicate ± 1 *SEM*.

Inspection of the lower panel of Figure 17 reveals a pattern of results that were consistent with the asymmetry in the acquisition data; from Session 3 onwards the medium+/soft- group showed superior discriminatory performance than the

soft+/medium- group. A Group × Session ANOVA for the discrimination ratios revealed a significant main effect of group, $F(1, 14) = 8.24$, $p = .012$, $\eta_p^2 = .37$. The remaining findings for this analysis were a significant main effect of session, $F(5, 70) = 15.26$, $p < .001$, $\eta_p^2 = .52$, but non-significant Session × Group interaction, $F(5, 70) = 1.62$, $p > .10$.

Discussion

The results indicate that the rats were readily able to discriminate between the soft- and medium-intensity clickers. Moreover, an asymmetry was observed whereby the medium+/soft- discrimination was acquired more readily than the soft+/medium- discrimination, a finding consistent with the pattern of results displayed in Experiment 4. These results therefore pose a problem for a stimulus generalisation account for the asymmetry in Experiment 1 as they are, in essence, the direct opposite of what should follow based on these principles. The fact that the results of Experiments 2 and 3 are also the opposite of what is predicted by this account suggests that the findings from this chapter present a serious challenge to the proposals of Kosaki et al (2013).

General Discussion

The experiments in the present chapter confirm that the asymmetry observed in discriminations of auditory intensity using tones (Pierrel et al. 1970), and white noises (Jakubowska & Zielinski, 1976), can also be found with discriminations between two clickers of different intensities. Experiment 1 demonstrated that the ease of a discrimination between a loud and soft clicker was dependent on which stimulus signalled the delivery of food. The acquisition of the task was benefitted

when the loud stimulus signalled food and the soft stimulus signalled the absence of food relative to when the opposite reward contingency was true.

According to Kosaki et al. (2013), whose proposals were initially applied to an asymmetry observed in the discrimination of length, the ease of a discrimination is determined, in part, by the similarity of the S+ to stimuli presented during the ITI. Thus in Experiment 1 the loud+/soft- group was assumed to acquire the discrimination more readily than the soft+/loud- group because the loud S+ is more dissimilar to the ITI, and therefore receives less generalisation of inhibition, than the soft S+. The purpose of Experiments 2 and 3 was to test the prediction that if stimuli are presented in a session without an ITI then, the asymmetry should be reduced or indeed eliminated. In stark contrast to this prediction, when the stimuli were presented consecutively without an ITI an asymmetry in favour of the loud+/soft- task was still observed.

Experiments 4 and 5 were conducted to further assess the proposals of Kosaki et al (2013). If the stimulation presented during the ITI is made more similar to the loud clicker than the soft clicker, then the asymmetry should be reversed; discriminations where the lowest intensity clicker signals food and the higher intensity clicker signals the absence of food should be acquired more readily when the reverse treatment is true. Once again the results from these experiments were in contrast to these proposals. Soft+/no clicker- discriminations were acquired more readily than no clicker+/soft- discriminations (Experiment 4; although a degree of caution should be held about forming conclusions based on this marginal result), and medium+/soft- discriminations were acquired more readily than soft+/medium- discriminations (Experiment 5).

It therefore seems that an alternative account to that of Kosaki et al. is required in order to explain the results from this chapter. Consideration of what this account might be will be withheld until Chapter 5.

Taken together, the results from Experiments 2 to 5 suggest that the proposals of Kosaki et al. (2013) to account for the asymmetry in length struggle to explain a similar asymmetry in the discrimination of auditory intensity. Whether or not this means the proposals are appropriate for discriminations of length, or indeed any other dimension of magnitude, remains beyond the scope of this chapter. The purpose of Chapter 3 was to test the proposals of Kosaki et al. but with a different magnitude dimension to auditory intensity.

CHAPTER 3

Asymmetry in the discrimination of quantity by rats

The experiments in Chapter 2 were conducted to demonstrate an asymmetry when rats are required to discriminate between two different intensities of a clicker, and then to address the role of stimuli present during the ITI in causing this asymmetry. They demonstrated, in direct contrast to what follows from an account of this asymmetry in terms of generalisation of inhibition from the ITI (Kosaki et al., 2013; Logan, 1954), that making the stimulus present during the ITI more similar to the high-intensity clicker, rather than the low-intensity clicker, did not result in soft+/loud- discriminations being acquired more readily than loud+/soft- discriminations. The experiments also revealed that removing the ITI altogether did not eliminate the asymmetry. These findings therefore present a significant challenge for the account of the asymmetry in magnitude discriminations proposed by Kosaki et al.

Of course it is possible that some unique feature of discriminations based on auditory intensity means that the mechanisms responsible for the asymmetry are different to those for length – the magnitude investigated by Kosaki et al (2013). As noted in the Chapter 1, there appears to be some distinction between dimensions where changes in magnitude can be understood as changes in perceived intensity; loudness, brightness, odour and so on, and others in which the relationship between magnitude and intensity are not immediately obvious; length, duration, number etc. Is it possible that the proposals of Kosaki et al. (1970) only apply to those in the latter category? One purpose of the present chapter is to test this possibility by examining whether the asymmetry in magnitude discriminations extends to tasks where the variations in the number of objects that are displayed serve as the signals for presence and absence of reward.

On the basis of the small amount of evidence that is available (refer to Chapter 1 for more detail), it appears that discriminations between different quantities or numbers of the same object are asymmetrical. In an experiment by Watanabe (1998), pigeons were presented with displays consisting of either two or four red balls with food made available for key pecking in the presence of one of the displays. The two birds who were required to peck for food in the presence of four balls, but not two, each solved the discrimination more readily than the remaining two birds who had to peck in the presence of two balls for food, but not four. For a similar finding from a study using three brown bears see Vonk & Beran (2012). Neither of these experiments was intended as an investigation of whether a discrimination between a large and a small number of the same objects is easier when reward is signalled by the larger rather than smaller number and the small sample sizes make it difficult to draw any clear theoretical conclusions from the experiments. Consequently, the initial purpose of the present experiments was to provide the first direct test of whether an asymmetry exists in discriminations based on different numbers of the same object. The experiments were conducted with rats because although an asymmetry has been found using this species, with discriminations based on the intensity of sound (Zeilinski & Jakubowska, 1977), the length of an object (Kosaki et al., 2013), and the duration of an auditory cue (Kyd et al. 2008) it remains to be determined if a similar asymmetry can be found with rats when the discrimination involves differences in quantity.

Experiment 6 was based on the design depicted in Figure 15. Two groups of rats received appetitive Pavlovian conditioning in which the conditioned stimuli (CS) were patterns containing either 5 identical black squares or 20 identical black

squares (see Figure 18). The squares were presented against a white background on a television screen that was illuminated white throughout the intertrial interval (ITI). Food was presented after patterns with 5 squares, S+, but not after patterns with 20 squares, S-, for the 5+/20- group. For the 20+/5- group there were 20 squares in S+ and 5 squares in S-. According to the proposals of Kosaki et al. (2013), the background cues that are present during the ITI, such as the white television screen, will enter into inhibitory associations by virtue of being present for prolonged periods in the absence of food. This inhibition will then generalise to the training stimuli, but it is likely that the extent of this generalisation will be greater to the pattern containing 5 rather than 20 squares, as the former will contain a larger proportion of background cues than the latter. As a consequence, the generalisation of inhibition from background cues is then predicted to facilitate the acquisition of the 20+/5- discrimination, and disrupt the acquisition of the 5+/20- discrimination. The results confirmed this prediction.

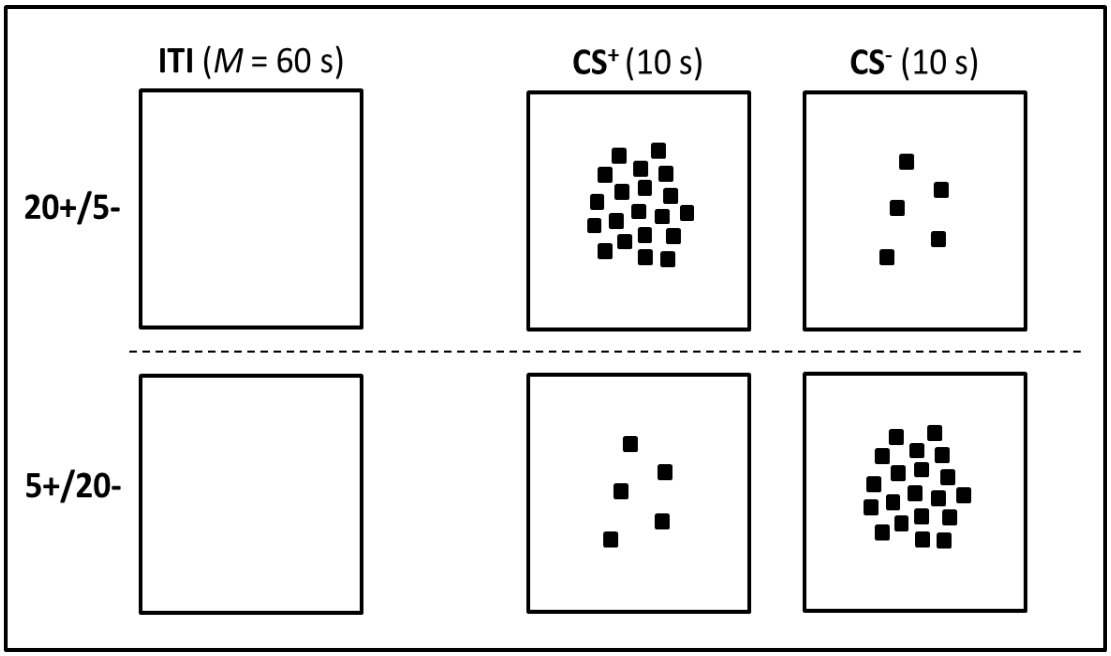

Figure 18. The stimuli used for Experiment 6. The figure is for illustrative purposes and does not depict accurately the images used in the experiments.

The remaining two experiments, which were, in essence, replications of Experiments in Chapter 2, were then conducted in order to assess the role played by cues present during the ITI on the asymmetry revealed in Experiment 6. In Experiment 7, the cues present during the ITI were made more similar to the large quantity, rather than the small quantity conditioned stimulus (CS), in order to determine if this manipulation would reverse the asymmetry. Experiment 8 was conducted to assess the effect of conducting the quantity discrimination without an ITI. If the stimulation during the ITI plays a crucial role in the asymmetry observed with the discriminations in Experiment 6, then conducting training without an ITI will abolish the asymmetry.

Experiment 6

Two groups of rats received appetitive Pavlovian conditioning in chambers from which they could view the stimuli depicted in Figure 18 on the screen of a computer monitor. Sucrose solution could be delivered to a dispenser on the outside of the wall nearest the monitor, and access to this dispenser was made possible by a small hole in the wall. Sucrose solution was available after patterns with 20 squares, but not 5 squares for the 20+/5- group, whereas the opposite arrangement was used for the 5+/20- group. As training progressed, it was expected that the frequency with which a rat's snout was inserted into the hole above the sucrose dispenser would be greater during patterns that signalled the imminent delivery of sucrose, rather than no delivery of sucrose. If this pattern of results should emerge, then it would indicate for the first time that this species is capable of solving magnitude discriminations based on different numbers of the same, visual object. More importantly, on the basis

of the proposals of Kosaki et al. (2013), the discrimination was expected to be acquired more readily by the 20+/5- group than the 5+/20- group.

Method

Subjects. The subjects were 32 male hooded Lister rats supplied by Harlan Olac (Bicester, Oxon, UK). Their mean free-feeding weight was 303 g. They were housed in pairs in a temperature-controlled colony room (approximately 20°C) that was continuously illuminated for 12 hours per day, with lights on at 07:00. They had access to water *ad libitum* but were food deprived to between 80 – 85 % of their free feeding weights prior to the start of the experiment. They were maintained at this weight by being fed a restricted diet after each experimental session. The rats had previously been used for an appetitive conditioning experiment for which they were divided into two groups that received different auditory discriminations. Prior to the present experiment they were randomly assigned to the two new groups (n = 16), with the constraint that each of the new groups contained 8 rats from each of the former groups.

Apparatus and stimuli. The experiment was conducted in the same eight conditioning chambers as Experiment 1 (Chapter 2), although the levers present in Experiment 1 had been removed. Beyond the magazine, at a distance of 8 cm, a 34 × 27 cm computer monitor was placed to present visual stimuli. The lower edge of the screen was in line with the chamber floor. Any other omitted details are the same as Experiment 1. Training stimuli consisted of either 5 or 20 filled black squares (1 × 1 cm) which were randomly arranged within a circular area (20 cm in diameter) in the centre of a white background. Ten different arrangements of both stimuli were used. Squares in the 5-and 20-square stimuli were separated by a mean distance of 4.5 cm

(*SD* = 2.4 cm) and 1.8 cm (*SD* = 0.6 cm) respectively. During the ITI the computer monitors were illuminated white.

Procedure. Rats were first given two sessions in which they were trained to approach the magazine in order to obtain sucrose solution in the manner described in Experiment 1. During these sessions the computer monitors were switched off. The rats then received 14, 60-min sessions of discrimination training. Within each session there were nine trials with patterns showing 5 squares, and nine trials with patterns showing 20 squares. The order of trials was determined randomly with the restriction that a trial-type could not be repeated more than twice consecutively. Each stimulus was presented for 15 s at a time and successive trials were separated by an ITI of two, three, or four minutes (*M* = 3 min). The duration of each ITI was determined randomly but with the constraint that each of the three possible intervals was selected six times in each session. During the ITI the computer monitors were illuminated entirely white. For the 20+/5- group the presentation of 20 squares was immediately followed by the delivery of 1 ml sucrose solution whilst the presentation of 5 squares was not. For the 5+/20- group, 5 squares signalled the delivery of sucrose and 20 squares signalled the absence of sucrose.

Data Analysis. Individual mean rates of responding, based on the number of times the beam of the infra-red sensor was interrupted during a trial, were recorded for all trials in each of the 14 sessions of discrimination training. The rates of responding during the ITI before every trial were also recorded. Any omitted details concerning data analysis are the same as for Chapter 1.

Results and Discussion

The mean rates of responding to S+, S- and throughout every ITI for each of the 14 experimental sessions are presented in Figure 19. Rats in both groups eventually responded more frequently during S+ than S-, but the magnitude of this discrimination was more pronounced in the 20+/5- than the 5+/20- group. Figure 19 also presents the results as discrimination ratios which, not surprisingly, show that the discrimination was acquired more readily by the 20+/5- than the 5+/20- group.

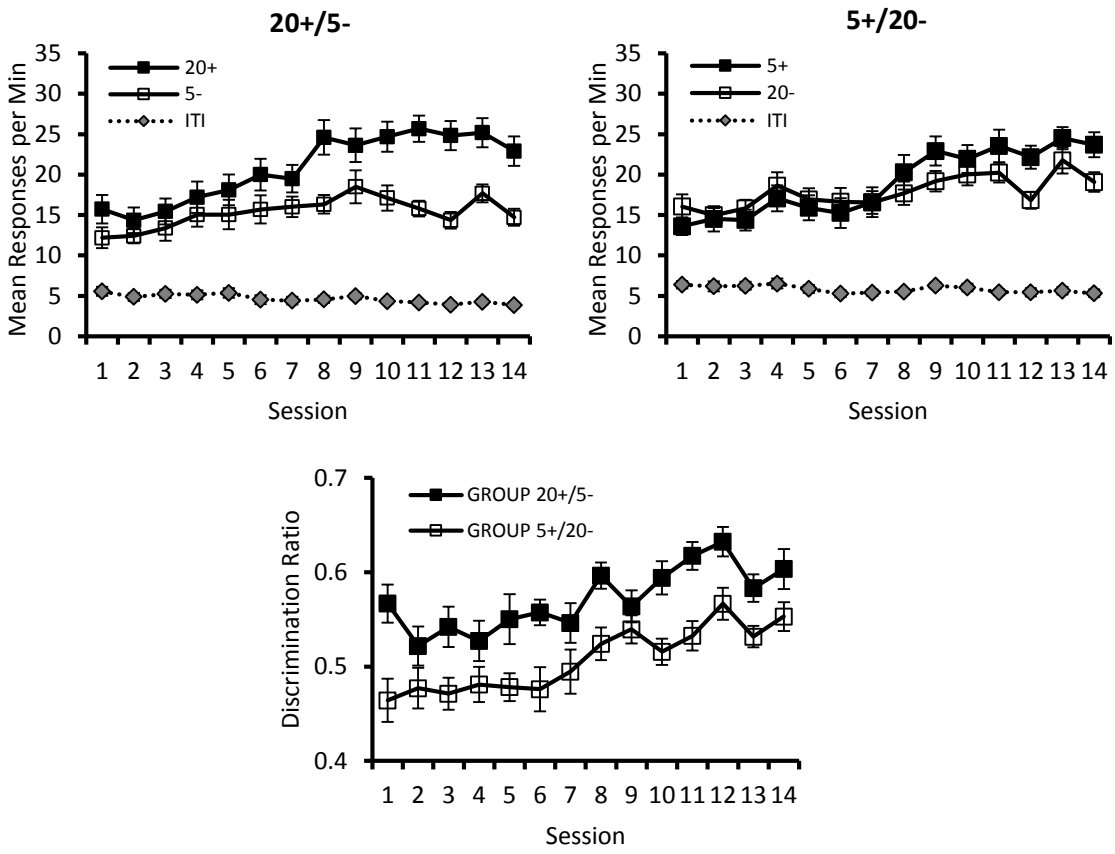

Figure 19. The mean rates of responding to S+ and S- for the 14 sessions of training for the 20+/5- (upper left-hand panel) and 5+/20- groups (upper right-hand panel) of Experiment 6. The lower panel shows the discrimination ratios derived from the mean rates of responding. Error bars represent ± 1 *SEM*.

The upper panels of Figure 19 show the mean rates of responding by the two groups during the reinforced, S+, and nonreinforced, S-, stimuli, and during the ITI, for the 14 sessions of training. The discrimination was acquired more readily by the 20+/5- group than the 5+/20- group. These observations were supported by a 3-way ANOVA of individual mean rates of responding during S+ and S- for each of the 14 sessions, which revealed a significant Stimulus × Group interaction, $F(1, 30) = 31.01$, $p < .001$, $\eta_p^2 = .51$. The remaining findings from the ANOVA were a significant effect of stimulus, $F(1, 30) = 71.30$, $p < .001$, $\eta_p^2 = .70$, and session, $F(13, 390) = 17.02$, $p < .001$, $\eta_p^2 = .36$, and Stimulus × Session interaction, $F(13, 390) = 9.19$, $p < .001$, $\eta_p^2 = .23$. The effect of group, and Session × Group and 3-way interactions, did not reach the accepted level for significance, *F*s < 1. Tests of simple main effects based on the significant Stimulus × Group interaction revealed a significant difference between responding to S+ and S- for the 20+/5- group, $F(1, 30) = 98.21$, $p < .001$, $\eta_p^2 = .77$, but no the 5+/20- group, $F(1, 30) = 4.13$, $p = .051$.

The lower panel of Figure 19 indicates that the discrimination ratios for the 20+/5- group were higher than for the 5+/20- group. Analysis of individual mean discrimination ratios with an ANOVA confirmed that the observed difference between the two groups was significant, $F(1, 30) = 49.80$, $p < .001$, $\eta_p^2 = .62$. The analysis also revealed a significant main effect of session, $F(13, 390) = 6.46$, $p < .001$, $\eta_p^2 = .18$, but the Session × Group interaction, $F < 1$, was not significant.

Inspection of the mean rates of ITI responding, which are displayed in the upper panels of Figure 19, shows that the rate of responding was generally more rapid in the 5+/20- group than the 20+/5- group. A two-way ANOVA of individual mean rates of responding during the ITIs for each of the 14 sessions revealed a

significant effect of group, $F(1, 30) = 4.99$, $p = .033$, $\eta_p^2 = .14$, and session, $F(13, 390) = 5.17$, $p < .000$, $\eta_p^2 = .15$. The Group × Session interaction was not significant, $F < 1$.

The results demonstrate for the first time with rats, an asymmetry in a magnitude discrimination based on different quantities of the same visual object. The results also support the proposal of Kosaki et al. (2013) that the acquisition of magnitude discriminations is influenced by the similarity between S+ and the cues present during the ITI. When the similarity between S+ and the ITI cues is close, as was the case for the 5+ 20- discrimination, then the discrimination will be harder to solve than when this difference is more pronounced, as was the case for the 20+ 5- discrimination.

Experiment 7

The two groups of rats in Experiment 7 received the same training as the two groups in Experiment 6, except that instead of being entirely white during the ITI, the television screen displayed 80 black squares clustered around the centre (see Figure 20). The purpose of this manipulation was to render the stimulation that was present during the ITI more similar to the patterns containing 20 black squares than those containing 5 black squares. Thus this experiment can be considered to be analogous to Experiment 5 in Chapter 2 where medium- and soft-intensity clicker trials were separated by an ITI in which a loud clicker was presented. According to the proposals of Kosaki et al. (2013), this manipulation will result in a reversal of the asymmetry seen in Experiment 10 with the 20+/5- discrimination being harder to solve than the 5+/20- discrimination.

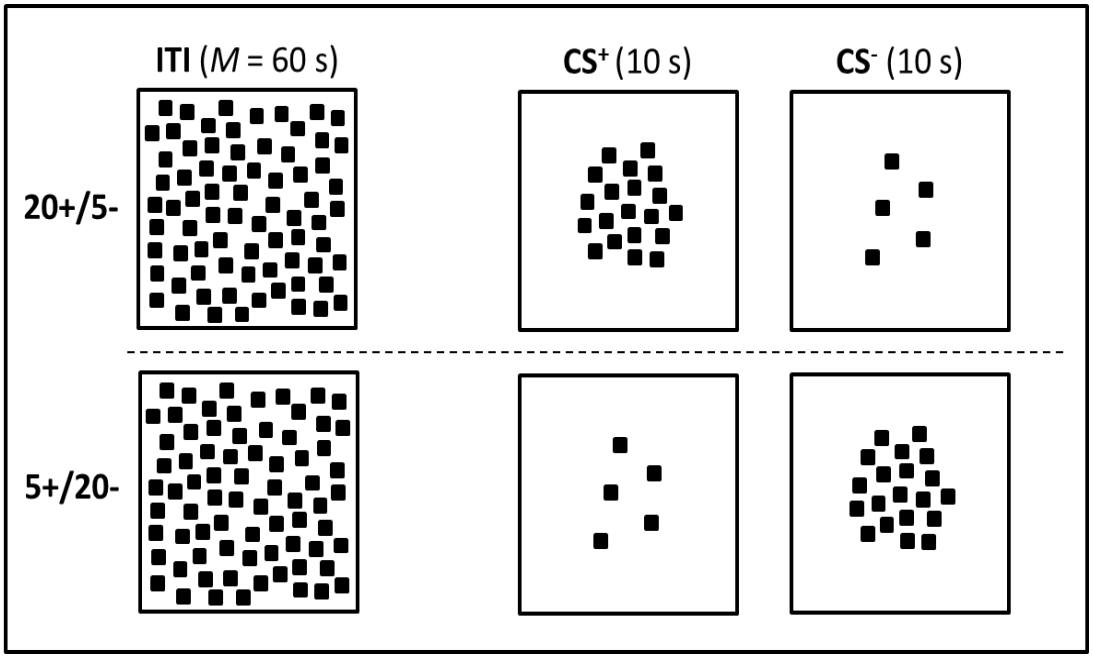

Figure 20. The stimuli used for Experiment 7. The figure is for illustrative purposes and does not depict accurately the images used in the experiments.

Of course, it is possible that the proposals of Kosaki et al. (2013) do not apply to rats, and the asymmetry in Experiment 6 may have occurred for a quite different reason to the one they put forward. Indeed, this would be entirely consistent with the failure to observe the reversed asymmetry with auditory stimuli in Experiment 5. For example, the salience of 20 squares might be regarded as being greater than of 5 squares. Theories of conditioning (e.g. Rescorla & Wagner, 1972) would then predict that excitatory conditioning with 20 squares will progress more rapidly than with 5 squares and, irrespective of the stimulation present during the ITI, the 20+/5- discrimination will be acquired more readily than 5+/20-.

Method

Subjects, apparatus, and procedure. The subjects were 16, experimentally naïve, male rats which were housed in the same manner as Experiment 6. Their

mean free-feeding weight was 256 g. They were gradually reduced to between 80 and 85% of their free feeding weights prior to training and maintained at these weights by being fed a restricted diet after each experimental session. The experiment was conducted in the same conditioning chambers as described in Experiment 6. The rats were assigned to the two groups randomly and in equal numbers (n = 8) prior to the start of the experiment.

In the first two sessions rats were trained to approach the food well in order to obtain sucrose solution in the manner described for Experiment 6, except that the screen was filled with alternating black and white horizontal bars that were each 4 cm high. The modification was made with the intention of ensuring that any associations formed during magazine training involved black and white stimuli equally. These associations, if they formed, were then expected to exert a similar influence on responding during the various patterns introduced at the outset of discrimination training.

After magazine training the rats were given 14 sessions of discrimination training with the same patterns containing 5 and 20 squares as for Experiment 6. Sucrose was presented after the patterns with the 20 squares, but not 5 squares, for the 20+/5- group, whereas for the 5+/20- group sucrose was presented after displays with 5, but not 20 squares on the screen. Throughout every ITI the white screen displayed 80 black squares scattered throughout a notional rectangle 28 cm wide and 22 cm high located in the centre of the screen. The squares, which were identical to those for the training stimuli, were separated by a mean distance of 1.8 cm (SD = 0.8 cm). The same pattern of 80 squares was used for every ITI. The remaining details

concerning the training protocol, and the recording of the results, were the same as for Experiment 6.

Results

The mean rates of responding to S+ and S- over the 14 sessions of discrimination training can be seen for the two groups on the upper panels of Figure 21. It is apparent that neither group found the discrimination easy to solve, but towards the end of training the rate of responding was faster during S+ than S- for the 5+/20- group (right-hand panel), whereas a similar difference failed to emerge in the 20+/5- group (left-hand panel), even by the end of training. This observation was supported by a 3-way ANOVA of the individual mean rates of responding during S+ and S- for each of the 14 sessions, which revealed a significant Stimulus × Group interaction, $F(1, 14) = 9.58$, $p = .008$, $\eta_p^2 = .41$. This ANOVA also revealed a significant effect of session, $F(13, 182) = 17.26$, $p < .001$, $\eta_p^2 = .55$, and group, $F(1, 14) = 6.96$, $p = .019$, $\eta_p^2 = .33$, but not stimulus, $F < 1$, and a significant Stimulus × Session interaction, $F(13, 182) = 3.93$, $p < .001$, $\eta_p^2 = .22$. The Session × Group interaction , $F < 1$, and 3-way interaction were not found to be significant, $F(13, 182) = 1.39$, $p > .10$. Tests of simple main effects based on the significant Stimulus × Group interaction revealed a significant difference in the rates of responding to S+ and S- for the 5+/20- group, $F(1, 14) = 7.04$, $p = .019$, $\eta_p^2 = .33$, but not the 20+/5- group, $F(1, 14) = 2.97$, $p > .10$.

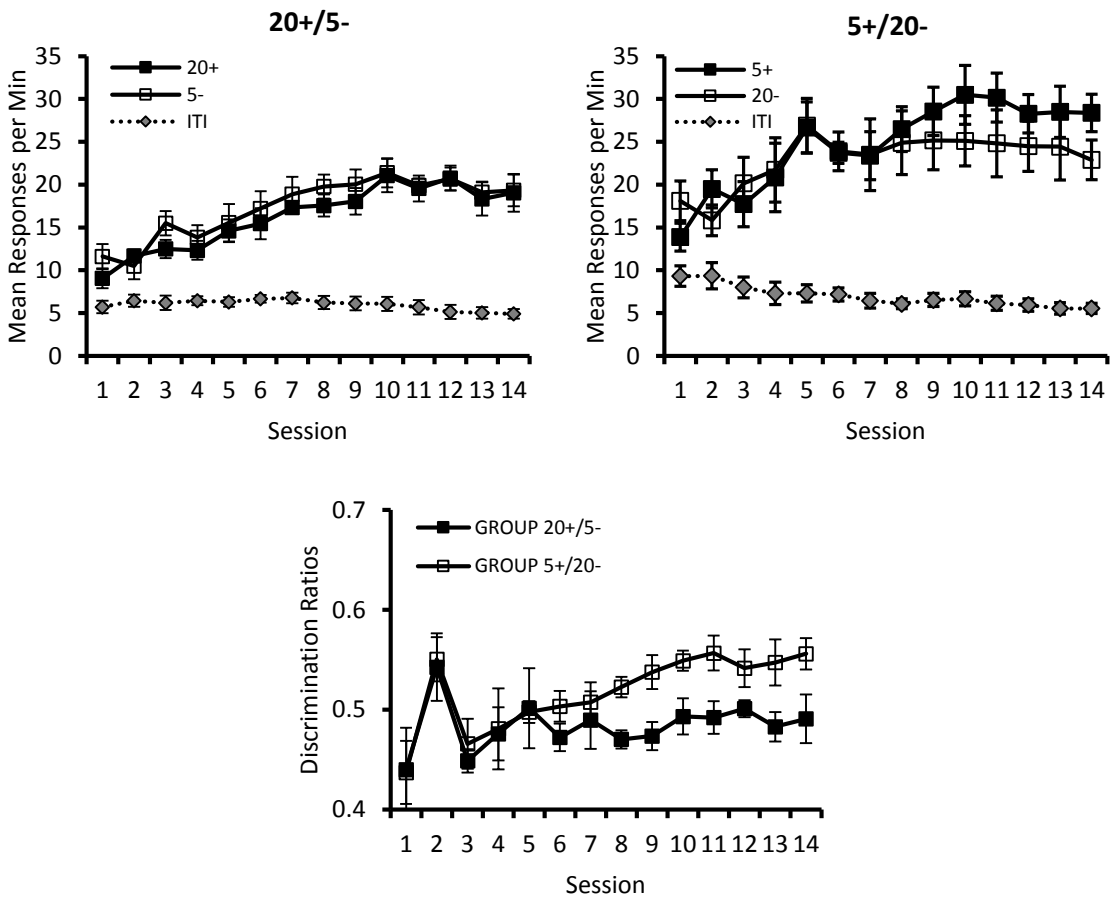

Figure 21. The mean rates of responding to S+ and S- for the 14 sessions of training for the 20+/5- (top left-hand panel) and 5+/20- groups (top right-hand panel) of Experiment 7. The lower panel shows the discrimination ratios derived from the mean rates of responding. Error bars represent ± 1 *SEM*.

The discrimination ratios shown in the lower panel of the figure also support the observation that the 5+/20- group acquired the discrimination more readily than the 20+/5- group. The group mean ratios for the 20+/5- group failed to rise consistently above .50 throughout the experiment, whereas those for the 5+/20- group increased gradually above this value as training progressed. ANOVA conducted on individual mean discrimination ratios for each of the 14 sessions confirmed that there was a significant main effect of group, $F(1, 14) = 7.66$, $p =$

.015, η_p^2 = .35, and a main effect of session, $F(13, 182) = 3.43$, $p < .001$, η_p^2 = .20, but the Session × Group interaction was not significant, $F < 1$.

The upper panels of Figure 21 also show that the mean rates of responding during the ITI were low for both groups. An ANOVA revealed that there was no significant difference between the groups, $F < 1$. The main effect of session, $F(13, 182) = 5.30$, $p < .001$, η_p^2 = .27, and Session × Group interaction, $F(13, 182) = 2.51$, $p = .003$, η_p^2 = .15, were found to be significant. Further analysis of the significant interaction revealed a significant effect of session for the 5+/20- group, $F(13, 182) = 6.31$, $p < .001$, but not the 20+/5- group, $F(13, 182) = 1.50$, $p > .10$. In addition, there was a significant difference between the groups on Sessions 1 and 2, Fs(1, 196) > 5.98.

Discussion

As predicted from the proposals of Kosaki et al. (2013), the presence of numerous black squares on the white screen of the monitor during the ITI resulted in the discrimination between 5 and 20 black squares being easier when the delivery of the reinforcer was signalled by 5 squares in the 5+/20- group than by 20 squares in the 20+/5- group. Indeed, there was no hint that the 20+/5 group was able to solve the discrimination. This pattern of results contrasts with those from Experiment 5 in Chapter 2 which indicated that making the intensity of the ITI more similar to the high-intensity clicker than the low-intensity clicker did not reverse the asymmetry. A discussion of why the pattern of results for auditory stimuli and numerical stimuli are so different will be withheld until Chapter 5.

An unexpected finding from the experiment was the failure of the 20+/5- group to show any indication of solving its discrimination. One might argue that this failure occurred because the group was unable to tell the difference between the training stimulus with 20 squares that signalled food, and the pattern of stimulation that was present during the ITI. Such an explanation, however, is challenged by the finding that throughout the experiment responding during the ITI was extremely slow, and consistently slower than during either of the training stimuli. Some other explanation is thus needed in order to explain the failure of the 20+/5- group to solve the discrimination. Apart from suggesting that insufficient sessions were administered, I am at a loss to suggest what this explanation might be.

Experiment 8

The asymmetry in the discrimination of quantity that was observed in the first two experiments has been attributed to differences between the similarity of the ITI cues and the stimuli that serve as S+ in the two discriminations (Kosaki et al., 2013). If this proposal is correct, then it should be possible to prevent the asymmetry from developing by presenting S+ and S- of the discrimination without them being separated by an ITI. Experiment 8 was conducted with this rationale in mind.

In order to be consistent with Experiment 2 in which the aforementioned prediction was tested with auditory stimuli, the same experimental procedure was used. The two groups of rats that were trained without an ITI received alternating exposure to patterns with either 5 or 40 black squares[4] throughout each experimental

[4] This experiment used 40 rather than 20 black squares in attempt to facilitate the discrimination that, in the previous experiment, was observed to be difficult to acquire by the rats.

session. Sucrose was signalled by patterns with 40 squares, but not 5, in the 40+/5-/no ITI group, and by patterns with 5 squares, but not 40, in the 5+/40-/no ITI group. On the basis of the proposals of Kosaki et al. (2013), the acquisition of this discrimination was expected to progress at the same rate in both groups.

Two additional groups for which S+ and S- were separated by an ITI were also included in the experiment. For the 40+/5-/ITI group sucrose was presented during trials with the 40-square patterns, but not the 5-square patterns, whereas for the 5+/40-/ITI group sucrose was presented during patterns with 5, but not 40 squares. The presence of the white screen during the ITI was expected to influence the two discriminations in the manner predicted by Kosaki et al. (2013), with the result that the discrimination will be acquired more rapidly by the 40+/5-/ITI than the 5+/40-/ITI group.

Method

Subjects and apparatus. The experiment used 64 experimentally naïve male rats housed in the same manner as the previous experiments. They were reduced to between 80 and 85% of their free feeding weights (*M* = 234 g) and maintained at this level by being fed a restricted diet after each experimental session. The experiment was conducted in the chambers used in Experiment 6.

Procedure. Rats were first given two sessions in which they were trained to retrieve sucrose when it was delivered into the food well. The full procedural details are described in Experiment 6. For the duration of each of these sessions the computer monitor was turned off. They were subsequently given ten further sessions in which they were trained to discriminate between 40- and 5-square stimuli. The 5-

square stimuli were the same as those for Experiments 6 and 7. The 40-square stimuli consisted of 40 black squares, with sides of 1 cm. The mean distance between the squares was 2.1 cm (SD = 0.6). Each session comprised nine presentations of S+ and nine presentations of the S- in an alternating fashion with each presentation of S+ and S- lasting 73 s. The stimulus type presented first was determined at random but kept the same across groups for each session. For all groups, during the presentation of each S+ sucrose was delivered at a randomly determined time between 1 and 20 s within each of 3 successive 20-s periods. The first of these 20-s periods occurred after a 10-s interval during which no sucrose was delivered. Each 20-s period was separated by 1 s in which no sucrose was delivered. Thus, each S+ presentation lasted 73 s. For the two groups trained with an ITI, the duration of each trial with S- was 73 s, and each trial was separated by an interval of 73 s when the screen was entirely white. For the groups trained without an ITI, the duration of each trial with S- was 219 s, and there was no interval between successive trials.

The number of snout entries prior to the first delivery of sucrose (pre-US) was recorded for every S+ trial. The timer that was in operation during S+, to determine when sucrose was first presented on each trial, was also in operation during S-, but its purpose was solely to identify the end of the interval during which responding was recorded for the nonreinforced trials. Responses were also recorded throughout every ITI for the two groups trained with an ITI.

Results

Figure 22 presents the mean rates of pre-US responding to S+ and S- across the 10 sessions of training for the four groups. The discrimination was solved successfully by each group. The results displayed in the upper two panels of the

figure are for the groups trained with an ITI, and it is evident that the group who received the 40+/5- discrimination mastered its problem more successfully than the group receiving the 5+/40- discrimination. Turning now to the lower two panels of Figure 22, which display the results from the two groups trained without an ITI, the course of acquisition of the two discriminations is remarkably similar. In support of these observations, a 4-way ANOVA with the factors group (40+/5- or 5+/40-), ITI (present or absent), stimulus and session, revealed a Group × ITI × Stimulus interaction, $F(1, 60) = 5.35$, $p = .024$, $\eta_p^2 = .08$. Tests of simple main effects based on this significant interaction revealed that the 40+/5-/ITI group acquired the discrimination more readily than the 5+/40-/ITI group, $F(1, 60) = 9.29$, $p = .003$, $\eta_p^2 = .13$, but there was no significant difference between the 40+/5-/no ITI and 5+/40-/no ITI groups, $F < 1$. The remaining findings of the overall 4-way ANOVA were significant effects of ITI, $F(1, 60) = 4.43$, $p = .04$, $\eta_p^2 = .07$, stimulus, $F(1, 60) = 36.76$, $p < .001$, $\eta_p^2 = .38$, and session, $F(9, 540) = 11.49$, $p < .001$, $\eta_p^2 = .16$, but not group, $F(1, 60) = 2.25$, $p > .10$, and significant ITI × Stimulus, $F(1, 60) = 7.16$, $p = .010$, $\eta_p^2 = .11$, and Stimulus × Session, $F(9, 540) = 36.06$, $p < .001$ interactions, $\eta_p^2 = .38$. The Group × Stimulus interaction was marginally significant, $F(1, 60) = 4.00$, $p = .050$. The Group × ITI, $F < 1$, Group × Session, $F < 1$, ITI × Session, $F(9, 540) = 1.07$, $p > .10$, Group × ITI × Session, $F(9, 540) = 1.64$, $p > .10$, Group × Stimulus × Session, $F(9, 540) = 1.45$, $p > .10$, ITI × Stimulus × Session, $F(9, 540) = 1.42$, $p > .10$, and 4-way interaction, $F(9, 540) = 1.13$, $p > .10$, were not significant.

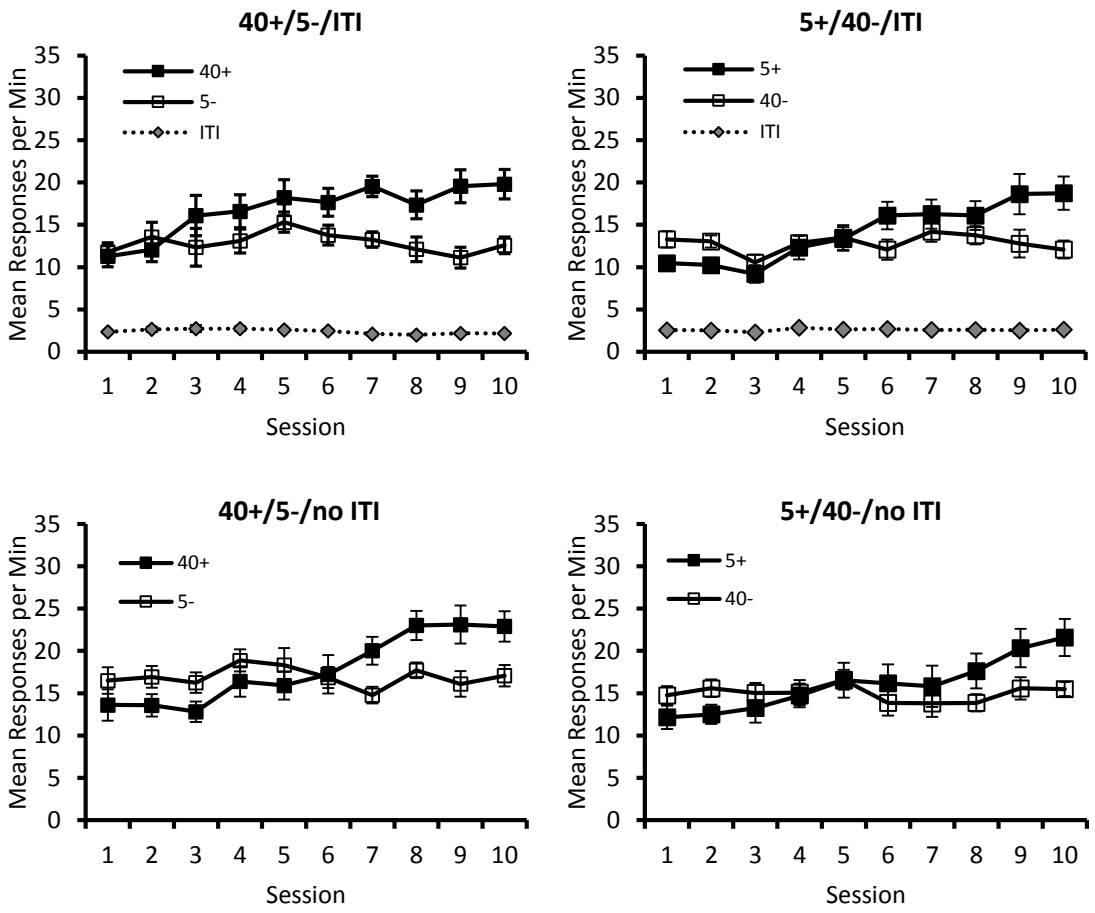

Figure 22. The mean rates of responding to S+ and S- for the 10 sessions of training for the 40+/5- (left-hand panels) and 5+/40- groups (right-hand panels) of Experiment 8. Top panels present responding for the ITI groups and lower panels present responding for the no ITI groups. Error bars represent ± 1 *SEM*.

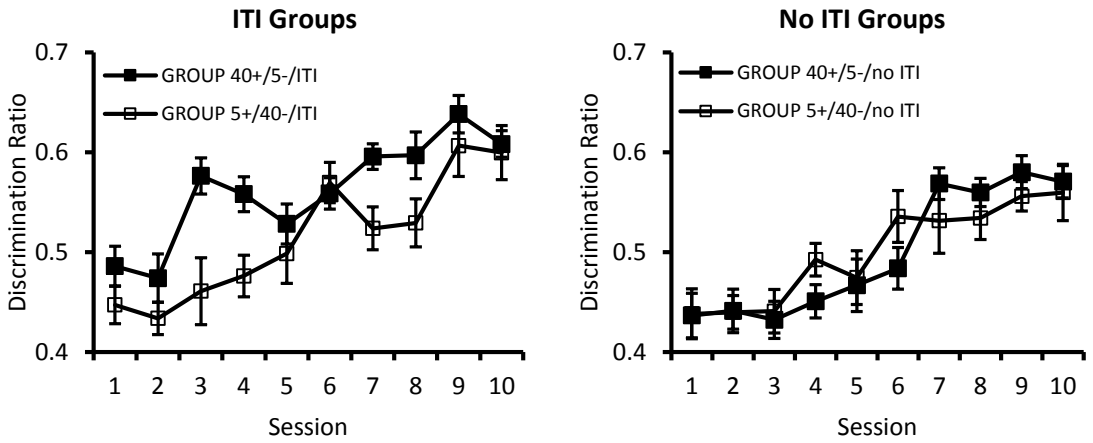

Figure 23. Discrimination ratios for the ten sessions of training for the ITI (left-hand panel) and no ITI groups (right-hand panel) in Experiment 8. Error bars represent ± 1 *SEM*.

In keeping with the observations made of Figure 22, the group mean discrimination ratios shown in the left-hand panel of Figure 23 were larger for the 40+/5-/ITI group than the 5+/40-/ITI group for all but one session of training. There was also not a substantial difference between the discrimination ratios for the 40+/5-/ no ITI group and the 5+/40-/ no ITI group, as can be seen in the right-hand panel of the same figure. A 3-way ANOVA of individual mean discrimination ratios for each of the 10 sessions, with the factors of ITI, group, and session revealed a significant main effect of ITI, $F(1, 60) = 16.59$, $p < .001$, $\eta_p^2 = .22$, and group, $F(1, 60) = 5.94$, $p = .018$, $\eta_p^2 = .09$, and importantly a significant ITI × Group interaction, $F(1, 60) = 6.61$, $p = .013$, $\eta_p^2 = .10$, supporting the observation that there was an asymmetry in the acquisition of the discriminations for the groups trained with an ITI, but not for those trained without an ITI. Further investigation of this significant interaction revealed a significant effect of group for rats trained with an ITI, $F(1, 60) = 21.89$, $p < .001$, $\eta_p^2 = .27$, but not the rats trained without an ITI, $F(1, 60) = 1.09$, $p > .10$. There was also a significant effect of ITI for the groups trained with 40 squares as S+, $F(1, 60) = 12.58$, $p = .001$, $\eta_p^2 = 17$, but not for those trained with 5 squares as S+, $F < 1$. The remaining results from the 3-way ANOVA were a significant effect of session, $F(9, 540) = 26.76$, $p < .001$, $\eta_p^2 = .31$, but the interactions of Session × ITI, $F(9, 540) = 1.18$, $p > .10$, Session × Group, $F(9, 540) = 1.51$, $p > .10$, and Session × ITI × Session were not significant, $F(9, 540) = 1.05$, $p > .10$.

To explore further the lack of significant difference between the ratios for group 40+/5-/no ITI and group 5+/40-/no ITI a Bayesian analysis was conducted of the mean discrimination ratios across the 10 sessions. This analysis tells us if the data favour the null hypothesis or the alternative hypothesis. A value above 3 has

been suggested to be the cut off for accepting whether the results substantially favour the null hypothesis and suggests that the null hypothesis is three times more likely than the alternative hypothesis given the data and priors (Rouder, Speckman, Sun, Morey, & Iverson, 2009). The analysis revealed a Bayes Factor of 2.97, which is just short of the conventional cut off value and is thus suggestive of a pattern of results in favour of the null hypothesis.

Figure 22 also presents the rate of responding during the ITI for the ITI groups. An ANOVA revealed that there was no difference in the rate of responding between groups, $F < 1$, although the effect of session, $F(9, 270) = 2.42$, $p = .012$, $\eta_p^2 = .07$, and the Session × Group interaction were significant, $F(9, 270) = 2.50$, $p = .009$, $\eta_p^2 = .08$. Further analysis of the significant interaction revealed an effect of session for the 40+/5- group, $F(9, 270) = 3.91$, $p < .001$, but not the 5+/40- group, $F(9, 270) = 1.02$, $p = .42$. There was also a significant effect of group at Session 8, $F(1, 300) = 4.08$.

Discussion

Despite the substantial differences between the treatments given to the two groups trained with an ITI in the present experiment, and that given to the two groups of Experiment 6, the results from both pairs of groups were similar. In each experiment an asymmetry was observed, with a magnitude discrimination based on quantity being acquired more readily when sucrose was signalled by the larger rather than the smaller of the two stimuli. The new finding to emerge from the present experiment is that this asymmetry is less likely to occur when there is no interval between successive trials of the discrimination.

The design of the present experiment ensured that the four groups experienced the same amount of exposure to S+, the same pattern of delivery of sucrose throughout the experimental session, and the same duration of each experimental session. As a consequence, there was a difference between the groups not only in whether they experienced an ITI, but in the amount of exposure to S-. The groups trained without an ITI received three times as much exposure to S- as the groups trained with an ITI. It is thus conceivable that the failure of the 40+/5-/no ITI group to acquire its discrimination more readily than the 5+/40-/no ITI group was not due to the absence of the ITI but to the excessive exposure to S-. I am unable to rule out fully this explanation for my results, but the lack of any theoretical account for why excessive exposure to S- should exert such an effect makes the explanation unlikely.

The failure to find an asymmetry in the discrimination of magnitude in the groups trained without an ITI stands in contrast to findings reported by Bouton and Hendrix (2011, see also Bouton & García-Gutiérrez, 2006) and to results in Chapter 2. Rats received appetitive Pavlovian conditioning in which a single CS was preceded by either a short or a long duration ITI. For one group, food was presented after the CS when it was preceded by a long, but not a short ITI, and for a second group food was delivered when the CS was preceded by a short but not a long ITI. The difference between the rates of responding on reinforced and nonreinforced trials during the CS was considerably greater when the delivery of food followed the longer rather than the shorter ITI. These results thus provide a further demonstration of the asymmetry observed in the acquisition of magnitude discriminations. However, because the cues that signalled whether food would be presented were

present throughout the ITI, there was no period of exposure to context in the absence of the cues on which the discrimination was based. Contrary to Experiment 12, therefore, this result demonstrates that an ITI is not essential for an asymmetry in the acquisition of a magnitude discrimination to be observed.

It is difficult to offer a satisfactory explanation for the discrepancy between the present results and those of Bouton and Hendrix (2011) and Bouton & García-Gutiérrez (2006), apart from suggesting that the asymmetry in discriminations based on stimulus magnitude is a result of more than one influence. Perhaps the principles responsible for the asymmetry found with magnitude discriminations based on quantity are different to those based on temporal duration. Why this should be the case remains to be determined.

General Discussion

The experiments have demonstrated for the first time with rats an asymmetry in the acquisition of a magnitude discrimination based on quantity. Experiments 6 and 8 revealed that a discrimination based on appetitive conditioning was relatively easy to acquire when the outcome was signalled by a large, but not a small number of identical objects, and relatively hard to acquire when the outcome was signalled by a small but not a large number of objects. The principal purpose of the experiments has been to determine if this pattern of results occurs because the ease with which a magnitude discrimination is solved is related to the similarity between S+ and the stimuli present during the ITI. When both sets of cues are similar then, according to Kosaki et al. (2013), the discrimination will be harder to solve than when they are different. Experiment 7 lent support to this possibility by showing that when the cues during the ITI were more similar to the CS composed of a large

rather than a small number of objects, then the asymmetry was reversed. Experiment 8 further demonstrated the importance of the role played by cues present during the ITI for the asymmetry in magnitude discriminations, by showing that the asymmetry was eliminated when training took place in the absence of an ITI.

It is interesting to note that the pattern of results obtained in the present chapter contrasts significantly with the results from Chapter 2. Experiment 7 indicates that the asymmetry is reversed when the cues during the ITI are more similar to the high-magnitude stimulus than the low-magnitude stimulus. Experiments 4 and 5 on the other hand demonstrated that the asymmetry was not reversed with the same manipulation. In a similar manner, Experiment 8 demonstrates that removal of the ITI eliminated the asymmetry, whilst Experiments 2 and 3 demonstrated that the asymmetry persisted. A full consideration of why the pattern of results might differ dependent on the stimulus type will be withheld until the general discussion in Chapter 5.

The experiments show for the first time that rats can discriminate between different quantities of the same visual stimulus, a black square, which raises the question of how the discrimination was solved. One possibility is that the discrimination was solved by referring to the number of squares that were displayed on the screen. A second possibility is that the discrimination was based on the distance between the squares, which was greater for the 5-square than the 20-square patterns (although if this was the case then the rats should not have responded more slowly during the 80-square ITI than the 20-square stimuli, which shared the same mean square density of 1.8 cm). Finally, there is the possibility that the discrimination was based on the amount of black, or the amount of white,

stimulation provided by the patterns. The 20-square patterns contained more black, and less white, than the 5-square patterns. On the basis of the present evidence, it is not possible to choose between these alternatives (although some research has stated that rats preferentially discriminate based on luminance, e.g. Minini & Jeffery, 2006), but that does not detract from the principal conclusion I wish to draw from the experiments. The stimulation that is present during the ITI plays a role in asymmetry that is found with discriminations based on stimulus magnitude, but a full theoretical understanding of how this role is effective remains to be developed.

CHAPTER 4

Asymmetry in the discrimination of quantity by pigeons:

The initial purpose of the experiments presented in this chapter was to determine if the results obtained with rats in Chapter 3 can also be found with pigeons. For the first experiment, Experiment 9, two groups of pigeons received autoshaping in which the conditioned stimuli (CS) were patterns containing either 5 identical black squares or 20 identical black squares (see Figure 24). The squares were presented against a white background on a television screen that was illuminated white throughout the inter-trial interval (ITI). Food was presented after patterns with 5 squares, S+, but not after patterns with 20 squares, S-, for the 5+/20- group. For the 20+/5- group there were 20 squares in S+ and 5 squares in S-. According to the proposals of Kosaki et al. (2013), the background cues that are present during the ITI, such as the white television screen, will enter into inhibitory associations by virtue of being present for prolonged periods in the absence of food. This inhibition will then generalise to the training stimuli, but it is likely that the extent of this generalisation will be greater to the pattern containing 5 rather than 20 squares, as the former will contain a larger proportion of background cues than the latter. As a consequence, the generalisation of inhibition from background cues is then predicted to facilitate the acquisition of the 20+/5- discrimination, and disrupt the acquisition of the 5+/20- discrimination. The results, consistent with Chapter 3, confirmed this prediction.

Experiment 10 was then conducted to assess the role played by cues present during the ITI on the asymmetry revealed in Experiment 9. Here the cues present during the ITI were made more similar to the large quantity, rather than the small quantity conditioned stimulus, in order to determine if this manipulation would reverse the asymmetry. Consistent with the pattern of results reported in Chapter 3,

this experiment revealed that the 5+/20- group acquired the discrimination more rapidly than the 20+/5- group. These results, in combination with those from Chapter 3, indicate that the asymmetry in quantity discriminations stems from generalisation between background cues present during the ITI and those present during S+ and S-. Experiments 11, 12 and 13 were therefore conducted in order to identify the dimension along which generalisation in quantity discriminations takes place.

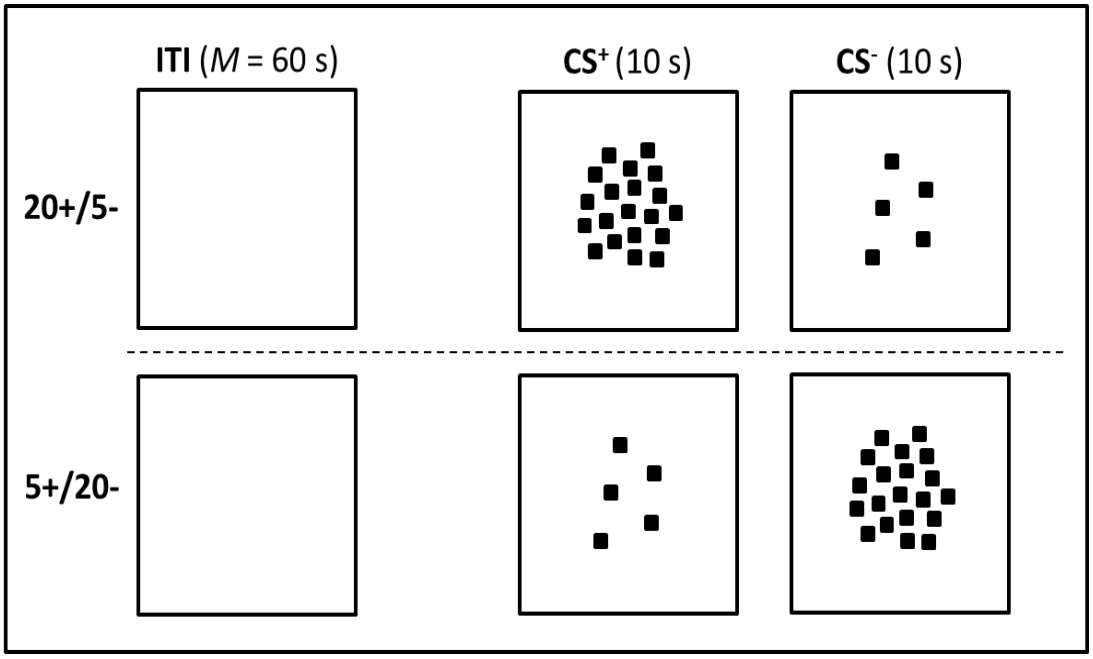

Figure 24. The stimuli for Experiment 9 (the figure is for illustrative purposes and does not depict accurately the images displayed to the pigeons).

Experiment 9

Method

Subjects. The subjects were 16 experimentally naïve adult homing pigeons (*Colomba livia*). They were housed in pairs in a temperature-controlled colony room (approximately 20°C) that was continuously illuminated for 14.5 hours per day, with

lights on at 07:00. They had access to water and grit *ad libitum* but were food deprived and reduced to between 80 - 85% of their free feeding weights (*M* = 411 g) prior to the start of the experiment. They were maintained at this weight by being fed a restricted diet after each experimental session. They were randomly assigned to the two groups in equal numbers.

Apparatus. Eight operant chambers (30.0 × 33.0 × 35.0 cm, L × W × H) were used. Each chamber was constructed from three aluminium walls, an aluminium ceiling, and a clear acrylic door serving as the fourth wall. A wire mesh served as the floor for these chambers. A tray lined with absorbent, odour – removing paper served to collect waste below the mesh. The left-hand wall looking into the chambers contained a clear acrylic response key (8.3 cm × 6.3 cm), which was hinged at the top. The midpoint of this key was 24 cm from the chamber floor and situated halfway between the two side walls. Pecks on this panel were detected by a reed relay, which was operated whenever a magnet on the bottom of the key was displaced by a distance greater than 1 mm. A color, thin-film transistor TV (Saka; 15.5 × 8.7 cm) was used to present the stimuli. Food was delivered into a food well (4.6 × 5.4 cm) which was located in the same wall. The midpoint of the entrance to this food well was 9.0 cm from the chamber floor and 7.0 cm to the left of the midline of the wall. Conditioning seed (Bucktons®) was made available inside the food well via a grain feeder (Colbourn Instruments, Lehigh Valley, PA). A PC with Whisker software, and programmed in Visual Basic 6.0, controlled the experimental events and recorded the number of pecks made on the key. Each chamber was contained in an individual sound attenuating chamber which was shut during the experimental session. Throughout each experimental session test

chambers were illuminated by a single 2.8-W bulb, operated at 24 V and located in the chamber ceiling.

Procedure. The subjects first received six sessions of pre-training, each lasting 60 min, in which they were trained to retrieve food from the food well whenever the grain feeder was operated. The televisions behind the response keys were turned off for this stage of the experiment. In order to encourage them to attend to the television screen, the birds then received 13 sessions of autoshaping in which a coloured cross was presented on the screen against a background of a different colour. For half the birds in each group the cross was red and the background was green, whereas for the remainder the cross was green and the background was red. The cross was 40 mm wide and 40 mm high with lines that were 7 mm thick. The cross was presented for 10 s and food was presented for 4 s as soon as the cross was removed from the screen. Throughout the ITI, the entire television screen was the same colour as the background for the conditioning trials. The duration of the ITI was increased gradually from 40 to 60 s during this stage.

Throughout the following eight sessions, the television screen was entirely white during every ITI. For each conditioning trial the screen was again white but also displayed either 5 or 20 black squares (3 mm × 3 mm). The squares were arranged randomly within a notional circle of diameter 4.1 cm, the centre of which was coincident with the centre of the television screen. There were ten different variants of the 5-square and the 20-square stimuli. The stimuli were presented in a random sequence with the constraint that no more than two trials with the same numbers of squares could occur in succession. The duration of each stimulus was 10 s, and the ITI was 40, 60 or 80 sec (M = 60 sec) determined randomly for each trial.

There were 20 trials with each of the two stimuli within each session. Food was not presented after any stimulus during the first two sessions of this stage. The purpose of these extinction sessions was to reduce responding during the experimental stimuli to a low rate, and thereby make it possible to observe differences in the acquisition of the discrimination when one of the stimuli, but not the other signalled food during the remaining six sessions of the experiment. During the final six sessions, every presentation of a 5-square stimulus, but not a 20-square stimulus, was followed by the delivery of food for 4 s for the 5+/20- group, whereas food was presented after the 20-square but not the 5-square stimulus for the 20+/5- group.

Data Analysis. Individual mean rates of responding, in responses per min, were recorded for all trials in the final extinction session, and for each of the six sessions of discrimination training. The rates of responding during an interval of 10-s before every trial were also recorded. When the majority of response rates were at zero, or close to zero, which was the case for the data analysed in the final extinction session, and those recorded during the pre-CS intervals, then non-parametric statistical tests were used. Where omitted, the details concerning data analysis are the same as for Chapters 2 and 3.

Results

The mean rates of responding to the 5-square and 20-square stimuli during the second session of extinction training were very low. For group 5+/20- the mean rate of responding to the 5-square stimuli was 0.9 responses per min and to the 20-square stimuli it was 1.5 responses per min. The equivalent results for the 20+/5- group were, respectively, 1.1 and 1.0 responses per min. The difference between the rates of responding to the 5- and 20-square stimuli was not significant in either

group, Wilcoxon *zs*(6) < 1.16, *p*s > .10. A between-group comparison of the mean rates of responding to both stimuli combined also revealed a non-significant difference, $U(8,8) = 23$, $p > .10$.

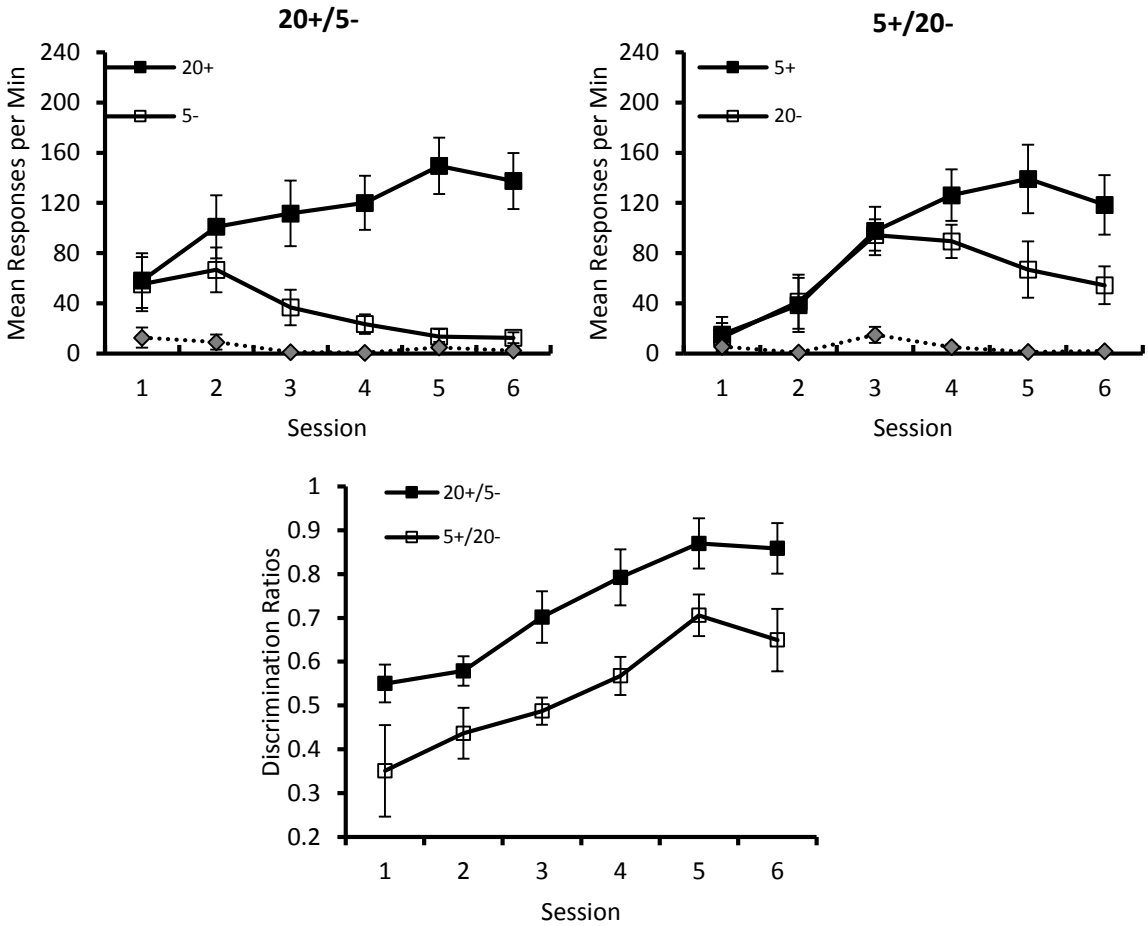

Figure 25. The mean rates of responding to S+ and S- for the six sessions of training for the 20+/5- (upper left-hand panel), and 5+/20- (upper right-hand panel) groups of Experiment 9. The lower panel presents the discrimination ratios for both groups. Error bars represent ± *SEM*.

The upper half of Figure 25 shows the group mean rates of responding by the two groups during the reinforced, S+, and nonreinforced, S-, stimuli, and during the pre-CS intervals, for the six sessions of discrimination training. The discrimination was acquired more readily by the 20+/5- than the 5+/20- group. These observations were supported by a 3-way ANOVA of individual mean rates of responding during S+ and S- for each of the six sessions, which revealed a significant, Stimulus ×

Group interaction, $F(1, 14) = 6.07$, $p = .027$, $\eta_p^2 = .30$. The remaining findings from the ANOVA were a significant effect of stimulus, $F(1, 14) = 29.38$, $p < .001$, $\eta_p^2 = .68$, and session, $F(5, 70) = 6.19$, $p < .001$, $\eta_p^2 = .31$, but not group, $F < 1$. The interactions of Session × Group, $F(5, 70) = 4.0$, $p = .003$, $\eta_p^2 = .22$, and Stimulus × Session, $F(5, 70) = 21.90$, $p < .001$, $\eta_p^2 = .61$, were significant, but the three-way interaction, $F(5, 70) = 2.20$, $p = .064$, was not significant. Tests of simple main effects based on the Stimulus × Group interaction revealed a significant difference in the rates of responding to S+ and S- for the 20+/5- group, $F(1, 14) = 31.08$, $p < .001$, $\eta_p^2 = .69$, but not the 5+/20- group, $F(1, 14) = 4.37$, $p = .055$.

The upper panels of Figure 25 also show that pigeons made few responses on the illuminated key during the 10-s pre-CS period. A between-group comparison of individual mean rates of responding for the six sessions combined revealed that they were not significantly different, $U(8, 8) = 26$, $p > .10$.

The lower panel of Figure 25 clearly indicates that the discrimination ratios for the 20+/5- group were higher than for the 5+/20- group. A 2-way ANOVA was conducted for these ratios and this analysis revealed a significant main effect of group, $F(1, 14) = 11.75$, $p = .004$, $\eta_p^2 = .46$, supporting the observation that the 20+/5- group acquired the discrimination more readily than the 5+/20- group. The remaining findings of this ANOVA were a significant effect of session, $F(5, 70) = 16.08$, $p < .001$, $\eta_p^2 = .53$ but the Group × Session interaction was not significant, $F < 1$.

Discussion

The more rapid acquisition of the discrimination by the 20+/5- group than the 5+/20- group confirms that discriminations in which reward and nonreward are signalled by different numbers of identical objects are asymmetrical. Moreover, in keeping with discriminations involving stimuli that differ either in physical length, or temporal duration, the asymmetry favoured the discrimination in which reward was signalled by the stimulus that is of larger rather than smaller magnitude. I noted in the Introduction to this chapter that the asymmetry may be a consequence of the generalisation of inhibition from cues present during the ITI to the cues used to signal the delivery and absence of food. If it is accepted that the extent of this generalisation is greater to the smaller than the larger of the two cues, then it would follow that the 20+/5- discrimination will benefit the generalisation of inhibition from the ITI cues, whereas the 5+/20-discrimination will be disrupted by this generalisation. Experiment 10 was conducted in order to evaluate this explanation for the results from Experiment 9.

Experiment 10

The experiment contained two groups who received the same discrimination as their namesakes in Experiment 9. In contrast to the first experiment, the screen during the ITI was not uniformly white but, for both groups, it consisted of a white background with 288 black squares randomly distributed over the entire screen (see Figure 26). These squares were identical to those used to create the experimental stimuli. In terms of the proposals put forward by Kosaki et al. (2013), the screen displaying a large number of squares during the ITI will enter into an inhibitory association, the effects of which will generalise to the experimental stimuli. The

extent of this generalisation is likely to be greater to the stimulus comprising 20 squares than the one comprising 5 squares. As a consequence, the influence of the inhibition associated with the cues present during the ITI will be to promote the acquisition of the 5+/20- discrimination, by augmenting the effects of the nonreinforced trials with the 20-square stimulus, and disrupt the acquisition of the 20+/5- discrimination, by counteracting the effects of excitatory conditioning with the 20-square stimulus. Thus the proposals of Kosaki et al. again predict there will be an asymmetry in the acquisition of the discriminations by the two groups but, on this occasion, performance will be superior by the group for which food is signalled by the 5-square rather than the 20-square stimulus.

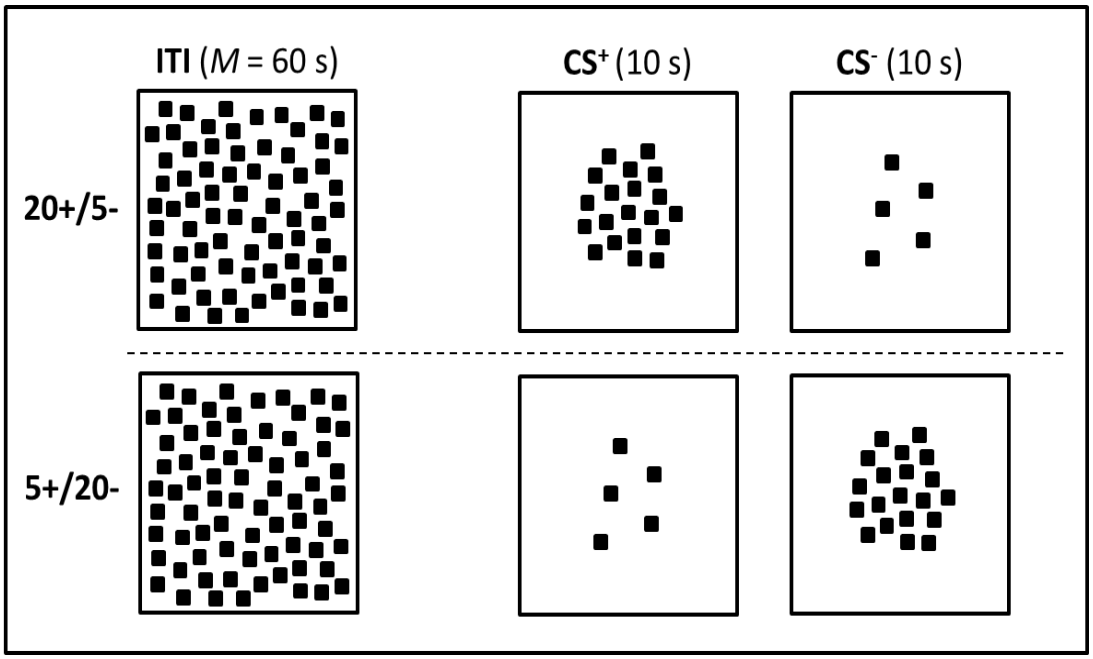

Figure 26. The stimuli for Experiment 10 (the figure is for illustrative purposes and does not depict accurately the images displayed to the pigeons).

Method

Subjects, apparatus and procedure. The subjects were 16 experimentally naïve homing pigeons with a mean free-feeding weight of 471 g. Their housing, and the method of food deprivation was the same as for Experiment 9. The birds were assigned at random to the two groups at the start of the experiment. The apparatus was the same as for Experiment 9.

Pretraining was the same as for Experiment 9, except that there were four sessions of magazine training and five sessions of autoshaping. For each of the final nine sessions, the television screen during every ITI was a white background covered with 288 squares. The squares were identical to those used to create the test patterns and were spread with approximately equal spacing but randomly distributed across the entire screen. The same pattern of 288 squares was used for every ITI. The experimental stimuli presented after each ITI consisted of patterns comprising either 5 or 20 black squares against a white background. These patterns were not followed by food for three extinction sessions. For the remaining six sessions, patterns with 5 squares were followed by food, and patterns with 20 squares were not followed by food for the 5+/20- group, whereas for the 20+/5- group food was presented after patterns with 20 but not 5 squares. Procedural details that have been omitted were the same as for Experiment 9.

Results

One pigeon from the 20+/5- group failed to respond during any session of discrimination training and was therefore excluded from the experiment. The average rates of responding per min to stimuli during the final session of extinction

was low for all groups. For the 20+/5- group the mean rate of responding to the 20-square stimuli was 3.0 responses per min and 1.9 responses per min to the 5-square stimuli. This difference was not significant, Wilcoxon $z(6) = .11$, $p > .10$. For the 5+/20- group the rates of responding were 2.0 and 2.3 responses per minute for the 20- and 5- square stimuli respectively. Again, this difference was not significant, Wilcoxon $z(7) = 1.69$, $p > .10$. Between group comparisons of the mean rates of responding to both stimuli also revealed a non-significant difference, $U(7, 8) = 26.5$, $p > .10$.

The results from the six sessions of discrimination training can be seen in Figure 27, which shows that the 5+/20- group acquired its discrimination more readily than the 20+/5- group. Indeed, the latter group showed no sign of mastering the discrimination, even after six sessions of training. A three-way ANOVA with the factors of group, session and stimulus (S+ vs S-) confirmed these observations revealing a significant three-way interaction, $F(5, 65) = 7.92$, $p < .001$, $\eta_p^2 = .38$, and significant Group × Stimulus interaction, $F(1, 13) = 10.32$, $p = .007$, $\eta_p^2 = .44$. The remaining findings were a significant effect of session, $F(5, 65) = 26.70$, $p < .001$, and Stimulus × Session interaction, $F(5, 65) = 8.25$, $p < .001$, $\eta_p^2 = .39$, but no significant effect of group, $F(1, 13) = 1.82$, $p < .10$, and no Group × Session interaction, $F < 1$. Tests of simple main effects based on the significant three-way interaction revealed a significant Group × Stimulus interaction from session three onwards, $Fs(1, 78) > 4.26$, $p < .043$, $\eta_p^2 = .05$. In addition, this analysis revealed a significant effect of stimulus from session three onwards for the 5+/20- group, $Fs(1, 78) > 6.36$, $ps < .014$, $\eta_p^2 = .08$, but no effect of stimulus at any session for the 20+/5- group, $Fs < 1$.

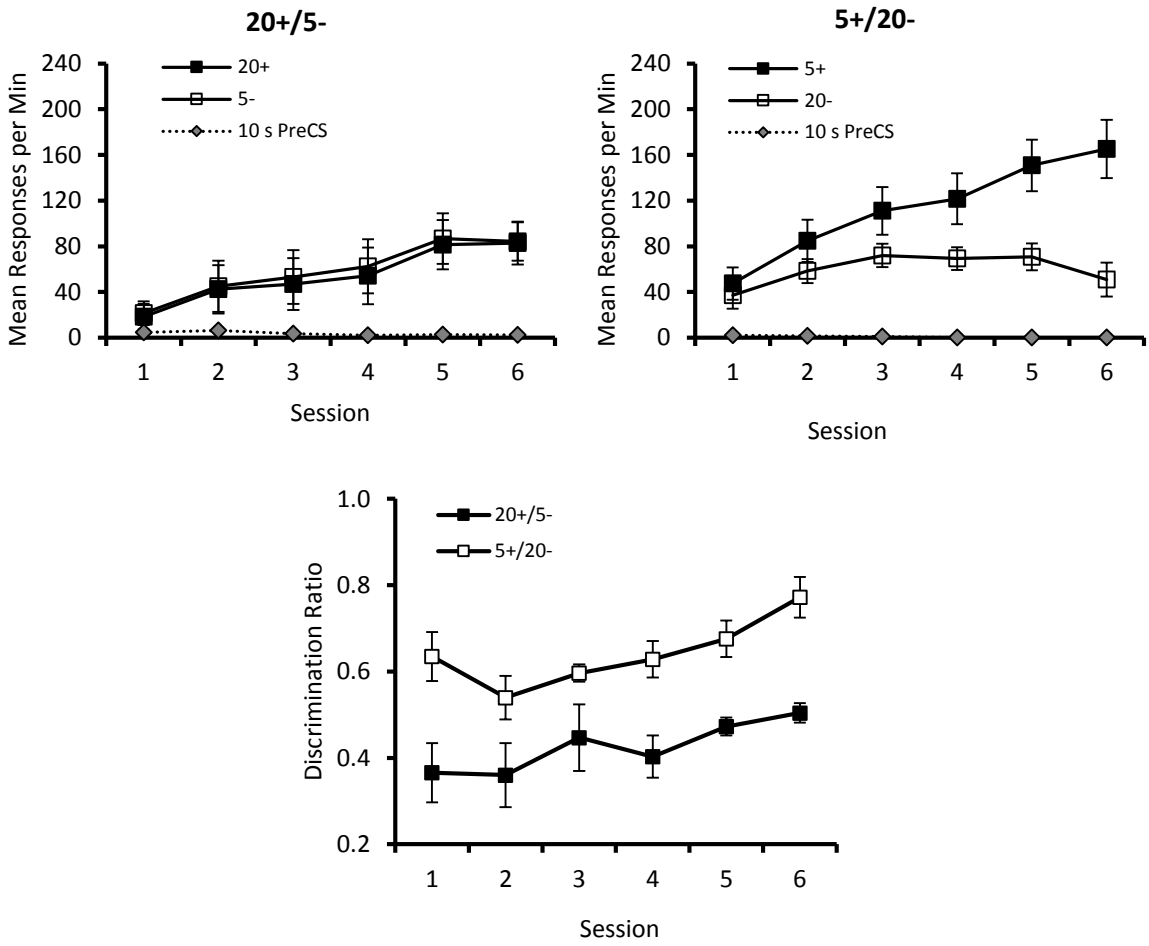

Figure 27. The mean rates of responding to S+ and S- for the six sessions of training for the 20+/5- (upper left-hand panel), and 5+/20- (upper right-hand panel) groups of Experiment 10. The lower panel presents the discrimination ratios for both groups. Error bars represent ± *SEM*.

Figure 27 also shows that throughout the final six sessions, responding during the pre-CS periods was at a very low rate for both groups. A comparison of individual mean rates of responding for the six sessions combined revealed that they were not significantly different, ($U(7, 8) = 14$, $p > .10$).

The lower panel of Figure 27 presents the discrimination ratios for the two groups presented in the upper panels. The figure reveals that the ratios for the 5+/20- group were higher for all sessions than the 20+/5- group. A 2-way ANOVA on these

ratios revealed a significant main effect of group, $F(1, 13) = 32.23$, $p < .001$, $\eta_p^2 = .71$, confirming this observation. This analysis also revealed a significant main effect of session, $F(5, 65) = 3.84$, $p = .004$, $\eta_p^2 = .23$, but the Group × Session interaction was not significant, $F < 1$.

Discussion

The results from the two groups were as anticipated. Thus the group for whom 5 squares signalled food, and 20 squares signalled the absence of food, readily acquired the discrimination whereas this was not the case for the group for whom 20 squares signalled food, and 5 squares signalled the absence of food. This pattern of results strongly suggests that one cause of the asymmetry in magnitude discriminations is the nature of the stimulation present during the ITI. If this stimulation is more similar to the small than the large training stimulus then the discrimination with the small stimulus as a signal for reward will be acquired with more difficulty than when the large stimulus signals reward. On the other hand, when the stimulation during the ITI is more similar to the large than the small stimulus, then the discrimination with food signalled by the small rather than the large stimulus will be acquired more readily.

As noted earlier, the design of the present experiment is very similar to one conducted with rats in Chapter 3. In keeping with the present results, they found that the group trained with 5 squares as S+ solved the discrimination, albeit slowly, but the group trained with 20 squares as S+ failed completely to solve the discrimination. It is always dangerous to place too much emphasis on a null result, but according to the proposals of Kosaki et al. (2013) there is no good reason why the 20+/5- group in the present study, and the equivalent group in the study in

Chapter 3, should fail to solve its discrimination. The results from the two experiments, therefore, may be regarded as posing a potential challenge to this account for the asymmetry that is seen with discriminations based on stimulus magnitude. Even though it may not be possible to provide a wholly satisfactory account for the present results (although see Chapter 5 for further discussion), their close similarity with those reported in Chapter 3 confirms their reliability and generality.

Experiment 11

The results from Experiment 10 point to the importance of the stimuli present during the ITI for the asymmetry between the acquisition of a 5+/20- and a 20+/5- discrimination. We have argued that this asymmetry stems from generalisation between the cues present during the ITI and those present during trials with S+ and S-. The purpose of the present experiment was to identify the dimension along which this generalisation takes place. Inspection of Figures 23 and 25 reveals that generalisation could be based on at least two possible dimensions. One dimension will be referred to as brightness. In Experiment 9 the overall brightness of the screen on which the stimuli were displayed was maximal during the ITI, not quite so bright for the trials with five squares, and least bright during trials with 20 squares. On this basis there was more scope for generalisation of inhibition from the ITI to the patterns with 5 rather than 20 squares, which would then account for the asymmetry that was observed. In Experiment 10, the numerous squares present during the ITI would mean that the screen was at its darkest during these intervals, and any inhibition associated with it would generalise to a greater extent to patterns with 20

rather than 5 squares, and result in the opposite asymmetry to that seen in Experiment 9.

The second possible dimension will be referred to as number. In Experiment 9, there were no squares present during the ITI, and either 5 or 20 squares present during the conditioning trials. If it is accepted that the absence of squares serves as an anchor representing zero on the dimension of number (e.g. Perkins, 1953), then it follows there will be more generalisation of inhibition from the ITI to the 5-square than the 20-square patterns in Experiment 9. These differences in generalisation of inhibition will then disrupt the 5+/20- discrimination and facilitate the 20+/5- discrimination. Conversely, there will be more scope for the generalisation of inhibition from the ITI patterns composed of 288 squares to the 20-square than 5-square patterns and thus the 20+/5- discrimination will be disrupted to a greater extent than the 5+/20- discrimination. By appealing to either the dimension of brightness or number, therefore, it is possible to explain the results thus far. The purpose of the present experiment was to identify which of these dimensions was used to solve the discriminations.

The two groups of Experiment 11 received the same discrimination training as for the previous experiments, except that the television screen was entirely black during the ITI (see Figure 28). If generalisation between the experimental stimuli and the cues present during the ITI is based on the dimension of brightness, then there will be more scope for generalisation between the black television screen of the ITI and the patterns containing 20 black squares, than the patterns containing 5 black squares. On this basis, therefore, the proposals of Kosaki et al. (2013) predict that a group receiving a 5+/20- discrimination will acquire it more readily than one

receiving a 20+/5- discrimination. In other words, using a black screen during the ITI should reverse the asymmetry that was seen in Experiment 9, in much the same manner as the 288 black squares that were presented on the white screen during the ITI in Experiment 10. A different outcome is predicted if generalisation between the stimuli used in the experiment is based on number. The absence of any small squares on the black screen during the ITI might result in it being treated as zero on the dimension of number and result in more generalisation of inhibition from this cue to the patterns displaying 5 rather than 20 squares. As a consequence, despite the very different stimulation provided by the television screen during the ITI in the present experiment, and Experiment 9, the outcome of both experiments is predicted to be the same. The 20+/5- group should acquire its discrimination more readily than the 5+/20- group.

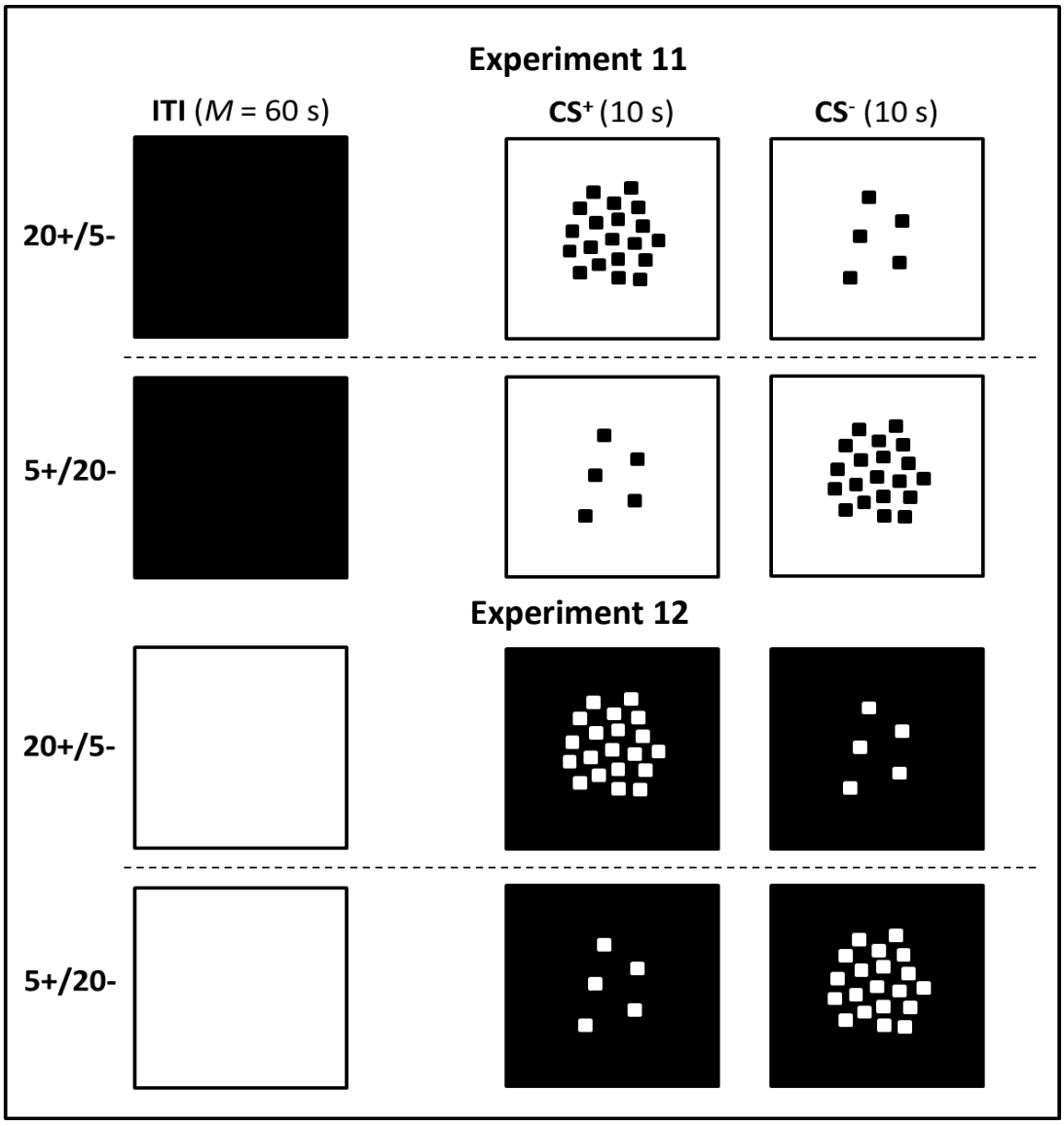

Figure 28. The stimuli for Experiments 11 and12 (the figure is for illustrative purposes and does not depict accurately the images displayed to the pigeons).

Method

Subjects, apparatus and procedure. Sixteen experimentally naïve pigeons with a mean free-feeding weight of 485 g were used. They were from the same stock and housed in the same manner as for Experiment 9. The method of food deprivation, the apparatus, and the procedural details concerning pretraining were the same as for Experiment 9.

The pretraining was followed by three sessions of extinction and six sessions of discrimination training in which the pigeons were presented with the same 20- and 5- square stimuli as in Experiment 10. For the duration of each ITI the screen was entirely black. In all other aspects the procedural details were identical to Experiments 9 and 10.

Results

The average rate of responding per minute to stimuli during the final session of extinction was low for all groups. For group 20+/5- the mean rate of responding to the 20-square stimuli was 1.3 responses per min and 2.1 responses per min to the 5-square stimuli. This difference was not significant, Wilcoxon $z(4) = 1.84$, $p > .05$. For group 5+/20- the equivalent rates of responding were 1.0 and .5 responses per minute, respectively. Again this difference was not significant, Wilcoxon $z(7) = 1.27$, $p > .10$. Between group comparisons of the mean rates of responding to both stimuli also revealed a non-significant difference, $U(8, 8) = 17.5$, $p > .10$.

The results from the six sessions of discrimination training can be seen in Figure 29, which shows that the 20+/5- group acquired its discrimination more readily than the 5+/20- discrimination. In support of this observation, a three-way ANOVA with the factors of group, session and stimulus (S+ vs S-) revealed a Group × Stimulus interaction, $F(1, 13) = 10.29$, $p = .006$, $\eta_p^2 = .44$. The three-way interaction fell short of the accepted level of significance, $F(5, 70) = 2.21$, $p = .063$, $\eta_p^2 = .14$. The remaining findings were a significant effect of session, $F(5, 70) = 6.95$, $p < .001$, $\eta_p^2 = .33$, a significant Stimulus × Session interaction, $F(5, 70) = 28.46$, $p < .001$, $\eta_p^2 = .67$, and a significant Group × Session interaction, $F(5, 70) = 5.56$, $p < .001$, $\eta_p^2 = .28$, but no significant effect of group, $F < 1$. Tests of simple

effects on the significant Group × Stimulus interaction revealed a significant effect of stimulus for the 20+/5- group, $F(1, 14) = 50.84$, $p < .001$, $\eta_p^2 = .78$, and the 5+/20- group, $F(1, 14) = 6.73$, $p = .021$, $\eta_p^2 = .32$. There was no significant effect of group for either the S+ or S-, Fs$(1, 28) < 1.66$, $ps > .10$.

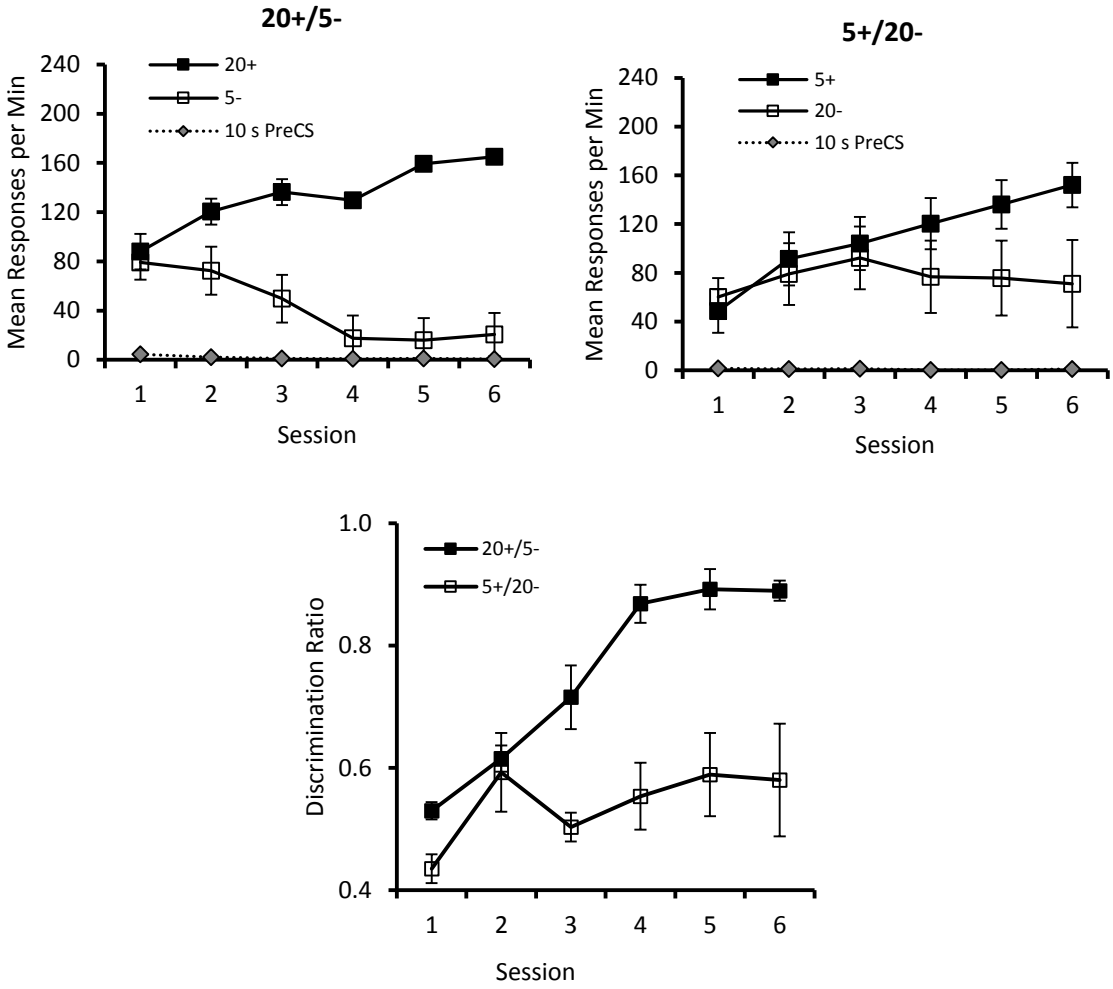

Figure 29. The mean rates of responding to S+ and S- for the six sessions of training for the 20+/5- (upper left-hand panel), and 5+/20- (upper right-hand panel) groups of Experiment 11. The lower panel presents the discrimination ratios for both groups. Error bars represent ± *SEM*.

Figure 29 also shows that responding during the pre-CS periods was at a very low rate for both groups throughout the final six sessions. A comparison of

individual mean rates of responding for the six sessions combined revealed that they were not significantly different for the two groups, $U(8, 8) = 30$, $p > .10$.

The lower panel in Figure 29 shows the discrimination ratios across the six sessions for the two groups. The figure reveals that from session 3 onwards the ratios for the 20+/5- group are higher than those for the 5+/20- group. An ANOVA on these ratios supported the above observation, revealing a significant main effect of group, $F(1, 14) = 35.33$, $p < .001$, $\eta_p^2 = .72$. The remaining findings were a significant effect of session, $F(5, 70) = 10.40$, $p < .001$, $\eta_p^2 = .43$, and Group × Session interaction, $F(5, 70) = 4.0$, $p = .003$, $\eta_p^2 = .22$. Further tests of simple effects for the significant interaction revealed an effect of session for the 20+/5- group, $F(5, 70) = 12.51$, $p < .001$, $\eta_p^2 = .47$, but no the 5+/20- group, $F(5, 70) = 1.96$, $p = .096$. The analysis also revealed a significant effect of group from session 3 onwards, Fs$(1, 84) > 9.89$, ps $< .002$, $\eta_p^2 = .11$.

Discussion

Despite the television screen being black during the ITI of the present experiment, as compared to white for Experiment 9, the results from both studies were remarkably similar. The 20+/5- discrimination was acquired more readily than the 5+/20- discrimination. If this asymmetry in the acquisition of the two discriminations is a consequence of generalisation of inhibition from the ITI being greater to one pattern than the other, then the present results indicate that the dimension along which generalisation takes place is unlikely to be brightness. For reasons noted in the introduction to the experiment, generalisation of inhibition along this dimension would result in the opposite pattern of results to that obtained. In contrast, provided it is accepted that a blank screen that is entirely black

represents zero small squares, then the above results can be explained by assuming that generalisation took place along the dimension of number. A greater amount of inhibition that would then generalise from the ITI to patterns with 5, rather than 20 squares, and thus favour the 20+/5- over the 5+/20- discrimination.

Experiment 12

The purpose of Experiment 12 was to test the theoretical conclusions that were drawn from the previous experiment. Training was similar to that for Experiment 11, but the screen was entirely white during the ITI, and the patterns consisted of white squares on a black background (see Figure 28). If birds rely on the dimension of number to solve the discrimination then a similar result to that observed in Experiments 9 and 11 will be found. The absence of squares during the ITI will ensure that the value of no squares will enter into an inhibitory association, which will generalise more strongly to the pattern displaying 5 rather than 20 squares and thereby help the 20+/5- discrimination and hinder the 5+/20- discrimination. On the other hand, if birds rely on the dimension of brightness to solve the discrimination, the high level of brightness of the screen during the ITI will enter into an inhibitory association. The effects of this association will then generalise more strongly to patterns with 20 squares than with 5 squares, because of the greater brightness of the former than the latter, and result in 20+/5- discrimination being harder to acquire than the 5+/20- discrimination.

Method

Subjects, apparatus and procedure. The subjects were 16 experimentally naïve, adult homing pigeons with a mean free-feeding weight of 422 g. They were

food deprived and housed in the same manner as for Experiment 9. At the start of the experiment they were randomly assigned to two groups. The apparatus was the same as for Experiment 9.

The pretraining was the same as for Experiment 9, except that there were seven sessions of magazine training and six sessions of autoshaping. Both groups then received two sessions of extinction training and then six sessions of discrimination training. For these final eight sessions the screen was white during the ITI and black with either 5 or 20 white squares for the training trials. Procedural details that have been omitted were the same as for Experiment 9.

Results

The average rate of responding per minute to both types of stimuli on the final day of extinction was low for all groups. For the 20+/5- group the mean rate of responding to the 20-square stimuli was .5 responses per min and .3 responses per min to the 5-square stimuli. This difference was not significant, Wilcoxon $z(4) = 1.13$, $p > .10$. For the 5+/ 20- group the equivalent values were .3 and .4 responses per min respectively. Again, this difference was not significant, Wilcoxon $z(4) = .68$, $p > .10$. Between group comparisons of the mean rates of responding to both stimuli also revealed a non-significant difference, $U(8, 8) = 24.5$, $p > .10$.

The results from discrimination stage can be seen in Figure 30, which shows both groups were able to discriminate successfully between visual arrays consisting of 5 and 20 white squares on a black background. It is also evident from the figure that the 20+/5- discrimination was acquired more readily than the 5+/20- discrimination.

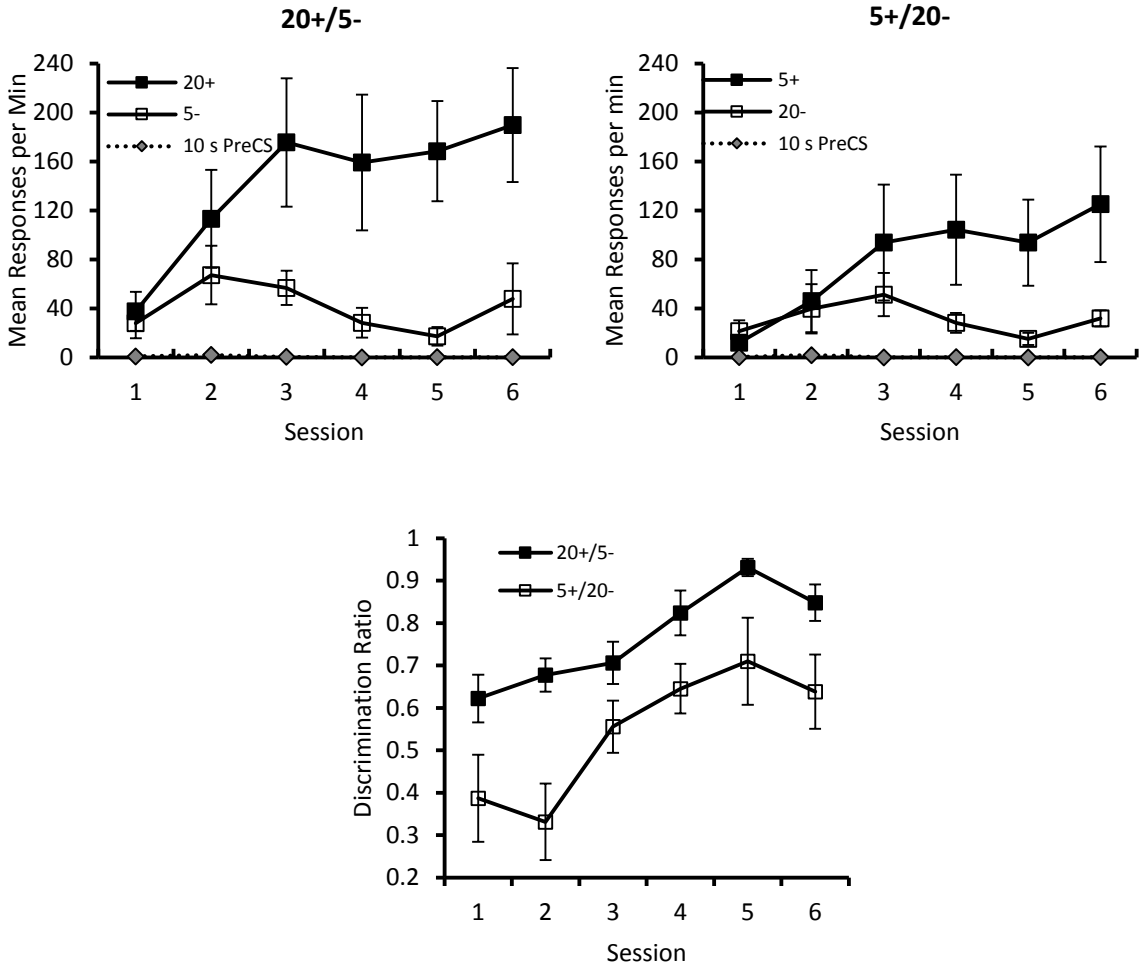

Figure 30. The mean rates of responding to S+ and S- across the six sessions of training for the 20+/5- (left-hand panel) and 5+/20- (right-hand panel) groups of Experiment 12. Error bars represent ± *SEM*.

A 3-way ANOVA with factors of group, stimulus and session confirmed that overall responding to the reinforced stimuli was significantly faster than to the non-reinforced stimuli, $F(1, 14) = 13.72$, $p = .002$, $\eta_p^2 = .49$. However, the critical Stimulus × Group, $F(1, 14) = 1.69$, $p = .21$, and Stimulus × Group × Session $F < 1$, interactions were not significant. The analysis also revealed a significant effect of session, $F(5, 70) = 11.31$, $p < .001$, $\eta_p^2 = .45$, and a significant Stimulus × Session interaction, $F(5, 70) = 9.71$, $p < .001$, $\eta_p^2 = .41$. The effect of group, $F(1, 14) = 1.17$, $p = .30$, and the Group × Session, $F < 1$, interaction were not significant. A separate analysis conducted on individual mean rates of responding during the 10-s pre-CS

periods for the six sessions combined. The analysis revealed that the slow rates of responding during these periods were not significantly different between groups, $U(8, 8) = 24.5$, $p > .10$.

The lower panel of Figure 30 presents the group mean ratios for each of the six sessions of discrimination training and shows that the discrimination was mastered more successfully by the 20+/5- group than the 5+/20- group. In support of this observation, analysis of the discrimination ratios with a two-way ANOVA revealed significant effects of group, $F(1, 14) = 10.60$, $p = .006$, $\eta_p^2 = .43$. There was also a significant effect of session, $F(5, 70) = 12.47$, $p < .001$, $\eta_p^2 = .47$, but the interaction was not significant, $F < 1$.

Discussion

The results from the experiment are entirely consistent with the claim that pigeons relied on information about the number of squares on the screen in order to solve the discriminations. The asymmetry that was observed across discrimination training can be understood if generalisation of inhibition was based on information about the number of squares on the screen, 0, 5 or 20. At the same time, the pattern of results is opposite to that predicted if stimulus generalisation was based on the overall illumination of the screen

In contrast to the previous experiments, the analysis of the rates of responding during the reinforced and nonreinforced stimuli of the training stage failed to reveal a significant interaction with the effect of group. From Figure 30 it is evident that numerically 20+/5- discrimination was acquired more readily than the 5+/20- discrimination. It is also evident from the error bars in the figure that the

within-group variation of response rates was considerable, which might explain the failure on this occasion to obtain a significant interaction involving the factors of group and stimulus. In support of this argument we can note that when the results were analysed in terms of discrimination ratios, in order to reduce the within-group variation in the data, then a significant asymmetry in the acquisition of the discriminations by the two groups was observed. A surprising aspect of the discrimination ratios plotted in Figure 30, is that for the first two sessions, the mean ratios were less than .50 for the 5+/20- group. This effect is a consequence of several birds in this group failing to respond on any conditioning trial and thus resulting in them being assigned a score of 0. A similar effect, for the same reason, was also observed in Experiment 11.

Experiment 13

The asymmetry revealed in each of the previous experiments has been explained by assuming there was generalisation of inhibition from the ITI to the experimental stimuli along the dimension of number. This conclusion then raises the question of how the dimension of number should be conceptualised. The dimension could be concrete and thus tied closely to the physical properties of objects that differ in number. Such a conceptualisation would result in generalisation from one quantity to another, but only when the objects in the two quantities are identical. Alternatively, the dimension could be more abstract and represent the number of squares, without regard to their physical properties (e.g. Dumont, Jones, Pearce, & Kosaki, 2015). According to this proposal, generalisation based on number might take place even when there is a change in a feature of the objects, such as their color (black or white). In order to choose between these possibilities, the pigeons from

Experiment 9 were used for one further experiment. The experiment commenced with a period in which the two groups of that experiment continued with their original training: 5+/20- and 20+/5-. The television screen was therefore white for the ITI, and white with black squares for the experimental stimuli. After this training the birds received a new discrimination for which the screen was again white for the ITI, but it was now black with either 5 or 20 white squares for the experimental stimuli. Thus the patterns were the negatives of those used for the initial training and were similar to those shown in the lower half of Figure 28. If the original discrimination was based on the concrete properties of the black squares, then the effects of the training with the original stimuli will not transfer to the white squares of the new discrimination. On the other hand, if the representation of the number of squares on the screen is more abstract, then transfer from the old to the new discrimination may well take place.

In order to evaluate these predictions, the two original groups were divided into four groups. The 5+/20-/ Same group was composed of four birds from the original 5+/20- group and received a 5+/20- discrimination with the new stimuli. Likewise, the 20+/5-/Same group was composed of four birds from the original 20+/5- group and received a 20+/5- discrimination with the new stimuli. The remaining two groups received the opposite training in the final stage to that administered originally. The 20+/5-/ Diff group initially received a 5+/20- discrimination but was then given a 20+/5- discrimination, while the 5+/20-/ Diff group received a 20+/5- discrimination followed by 5+/20-. If the original discrimination was solved by learning about the significance of different numbers of black squares, then performance on the new discrimination should not be influenced

by the effects of the original training. On this basis, the 20+/5-/Same and the 20+/5-/Diff groups should perform similarly on the new discrimination, and so too should the 5+/20-/Same and the 5+/20-/Diff groups. If, however, the original training resulted in the number of squares being represented in a more abstract fashion, then the acquisition of the new discrimination should be affected by the original training. That is, the two groups receiving the same discrimination in both stages would be expected to acquire the new discrimination more rapidly than the groups receiving different discriminations.

Whatever the fate of the foregoing predictions, a further prediction concerning the experiment is that the performance of the groups receiving the 20+/5- discriminations in the final stage will be superior to groups receiving the 5+/20- discriminations. Given that the screen was entirely white during the ITI, the absence of any small squares would be expected to enter into an inhibitory association. Generalisation of this inhibition would then disrupt the acquisition of the 5+/20- discrimination to a greater extent than the 20+/5- discrimination, and result in the former being acquired more slowly than the latter.

Method

Subjects, apparatus and procedure. The subjects were the same 16 pigeons used in Experiment 9. The apparatus was the same as for Experiment 9.

For Stage 1, the birds received six sessions of discrimination training, the details of which were the same as for Experiment 9. At the outset of this training, each of the two groups from Experiment 9 was divided at random into two groups with four birds in each group. On the day following the completion of Stage 1, the

four groups received six sessions of training in Stage 2. The ITI was again white but, during the trials, white squares were presented against a black background. The 20+/5-/Same group and the 5+/20-/Same group received the same discrimination in Stage 2 that they were given in Stage 1, whereas the 20+/5-/Diff group and 5+/20-/Diff groups received, respectively, a 20+/5- and a 5+/20- discrimination in Stage 2 and the opposite discrimination in Stage 1. Procedural details that have been omitted were the same as for Experiment 9

Results

All four groups showed consistently faster responding to S+ than S- throughout the six sessions of training in Stage 1. The mean rates of responding to S+ across the six sessions were 129.4 (*SE* = 38.5), 170.2 (*SE* = 31.5), 142.0 (*SE* = 42.0) and 159.3 (*SE* = 28.3) responses per minute for the 20+/5-/Same, 5+/20-/Same, 20+/5-/Diff and 5+/20-/Diff groups respectively. The mean rates of responding to S- for the same groups were 11.8 (*SE* = 1.0), 30.1 (*SE* = 10.7), 53.9 (*SE* = 22.2) and 12.7 (*SE* = 6.9) responses per minute. A three-way ANOVA with the factors of stimulus (S+ or S-) congruence (whether the final discrimination was the same or different to that for training), and discrimination (whether the discrimination was 20+/5- or 5+/20-) indicated that there was a significant effect of stimulus, $F(1, 12) = 81.00$, $p < .001$, but found no significant differences based on discrimination, $F < 1$, or congruence, $F < 1$, and no significant Discrimination × Congruence, $F(1, 12) = 1.03$, $p > .10$, Stimulus × Discrimination, $F(1, 12) = 2.16$, $p > .10$, Stimulus × Congruence, $F < 1$, or thee-way interaction, $F < 1$.

The top panels of Figure 31 show the results during Stage 2 for the 20+/5-/Same and the 5+/20-/Same groups, who received the same discriminations, in terms

of the relationship between the number of squares and the outcomes they signalled, in both stages. It is evident that changing from black squares on a white background, to white squares on a black background had rather little impact on the 20+/5- discrimination and the 5+/20- discrimination. Conversely, when the new discrimination was the opposite of that administered in the initial training, the acquisition of the new task was slower. The bottom left-hand panel indicates that the 20+/5-/Diff group, which was trained with the 5+/20- discrimination and then transferred to the 20+/5- discrimination found it difficult to acquire the new task, as compared with the 20+/5-/Same group. Likewise, the bottom right-hand panel indicates that the 5+/20-/Diff group, which was trained with the 5+/20- discrimination and then transferred to the 20+/5- discrimination, found the new task considerably more difficult than the 5+/20-/Same group.

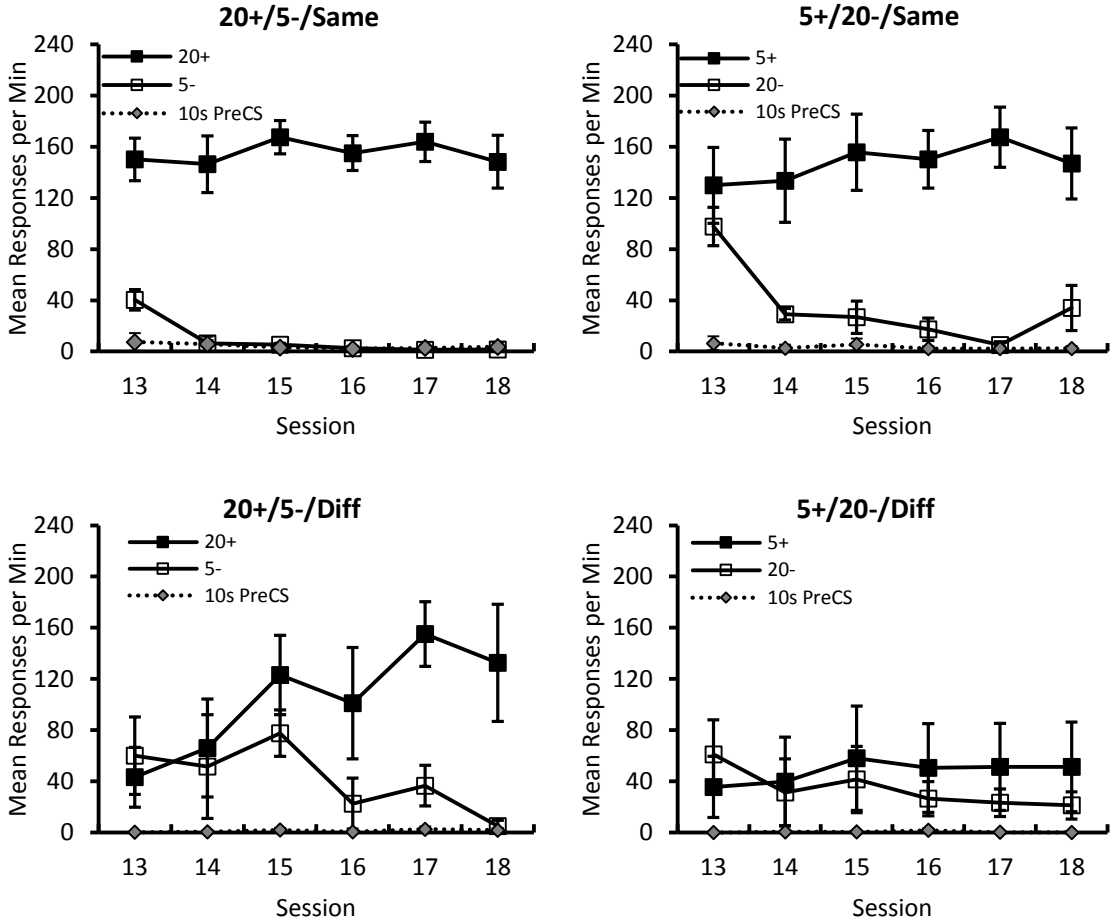

Figure 31. The mean rates of responding to S+ and S- for the 20+/5/Same (top left panel), 5+/20-/Same (top right panel), 20+/5-/Diff (bottom left panel) and the 5+/20-/Diff (bottom right panel) groups during Stage2 of Experiment 13. Error bars represent ± *SEM.*

The results from the experiment were analysed with a 4-way ANOVA with the factors of Stage-2 discrimination (20+/5- or 5+/20-), congruence (whether the Stage-2 discrimination was the same or different to the original discrimination), stimulus (S+ or S-) and session. The analysis revealed that the four-way interaction was not significant, $F < 1$, but there was a significant Stimulus × Congruence interaction, $F(1, 12) = 30.39$, $p < .001$, $\eta_p^2 = .72$, which indicates that the discrimination was acquired more readily by the groups receiving the same discrimination in both stages than those receiving different discriminations in both

stages. To explore this interaction further, tests of simple main effects were conducted. These tests revealed a significant difference between the overall rates of responding to S+ and S- for the two groups trained with the same discriminations in both stages, $F(1, 12) = 123.01$, $p < .001$, $\eta_p^2 = .91$, and for the two groups trained with different discriminations in both stages, $F(1, 12) = 10.86$, $p = .006$, $\eta_p^2 = .48$. In addition, there was a significant effect of congruence for S+, $F(1, 24) = 20.09$, $p < .001$, $\eta_p^2 = .46$, but not S-, $F < 1$.

To return to Figure 31, the asymmetry revealed in the previous experiments was again evident in the present study. Thus the discrimination by the 20+/5-/Same group was acquired more readily than by the 5+/20-/Same group, and the discrimination by the 20+/5-/Diff group was acquired more readily than by the 5+/20-/Diff group. In support of these observations, the four-way ANOVA also revealed a significant Stimulus × Session × Stage-2 Discrimination interaction, $F(5, 60) = 5.33$, $p < .001$, $\eta_p^2 = .31$.

The remaining results from the four-way ANOVA were as follows. The effects of session, $F(5, 60) = 2.17$, $p = .069$, congruence, $F(1, 12) = 4.12$, $p = .065$, group, $F(1, 12) = 1.24$, $p > .10$, and did not reach significance, but there was a significant effect of stimulus, $F(1, 12) = 103.48$, $p < .001$, $\eta_p^2 = .90$. There were significant interactions of Stimulus × Session, $F(5, 60) = 27.13$, $p < .001$, $\eta_p^2 = .69$, but the remaining interactions were not significant. Stimulus × Session × Congruence, $F(5, 60) = 2.12$, $p > .10$, Session × Congruence × Stage-2 Discrimination, $F < 1$, Session × Stage-2 Discrimination, $F < 1$, Session × Congruence, $F(5, 60) = 1.65$, $p > .10$, Stimulus × Congruence × Stage-2

Discrimination, $F(1, 12) = 4.08$, $p = .07$, Congruence × Stage-2 Discrimination, $F < 1$, and Stimulus × Stage-2 Discrimination, $F < 1$.

Figure 31 also reveals that pigeons in all groups made very few responses during the pre-CS period. Comparisons between the mean rates of responding across the six sessions revealed that there were no significant differences between the 20+/5-/Same and 20+/5-/Diff groups, $U(4, 4) = 6$, $p > .10$, or the 5+/20-/Same and 5+/20-/Diff groups, $U(4, 4) = 5.5$, $p > .10$. Additionally, when data was combined across Stage-2 discrimination (20+/5- vs. 5+/20) there was no significant effect of congruence, $U(4, 4) = 25$, $p > .10$, whilst when the data were combined across congruence there was no effect of Stage-2 discrimination, $U(4, 4) = 27.5$, $p > .10$.

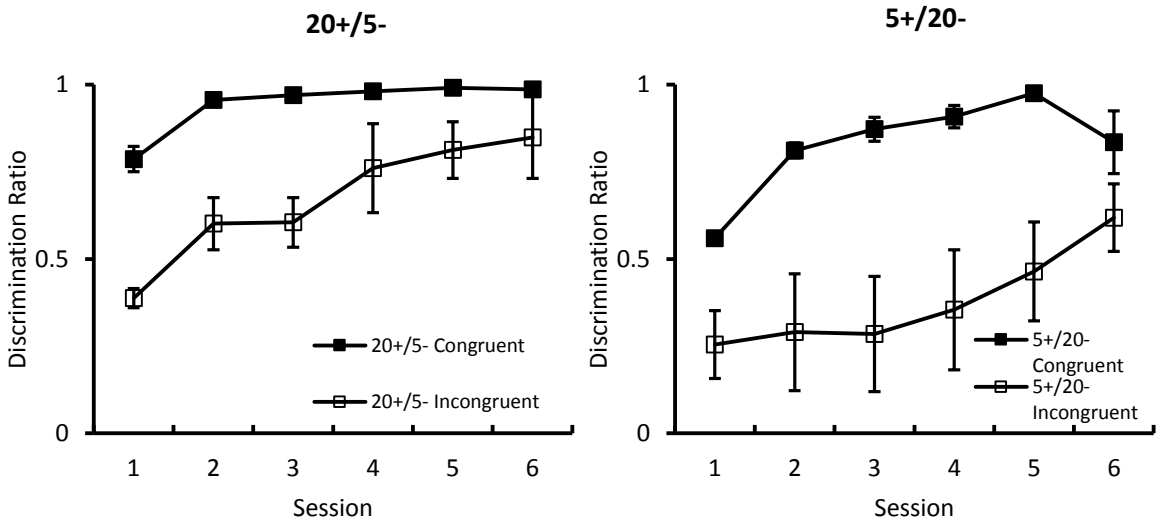

Figure 32. Discrimination ratios for the groups in Experiment 13. The left-hand panel represents subjects initially trained with a 20+/5- discrimination in Experiment 9 and the right-hand panel represents subjects initially trained with a 5+/20- discrimination in Experiment 9. Error bars represent ± *SEM*.

Figure 32 presents the discrimination ratios for the four groups in the stage-2 discrimination. Both panels present higher discrimination ratios for the congruent

groups than the incongruent groups. A 3-way ANOVA for these ratios with the factors stage-2 discrimination, congruent and session confirmed this observation, revealing a significant main effect of stage-2 discrimination, $F(1, 12) = 7.60$, $p = .017$, $\eta_p^2 = .39$. The analysis also revealed a significant effect of session, $F(5, 60) = 22.16$, $p < .001$, $\eta_p^2 = .65$, and congruence, $F(1, 12) = 23.72$, $p < .001$, $\eta_p^2 = .66$, and significant Session × Congruence interaction, $F(5, 60) = 4.14$, $p = .003$, $\eta_p^2 = .26$. Tests of simple effects for this significant interaction revealed an effect of session for congruent and incongruent groups, Fs$(5, 60) > 9.26$, ps $< .001$, $\eta_p^2 = .44$, and effect of congruence for sessions 1 to 5, Fs$(1, 72) > 15.77$, ps $< .001$, $\eta_p^2 = .18$. The effect of congruence at session 6 did not reach significance, $F(1, 72) = 3.91$, $p = .052$. The remaining findings from the 3-way ANOVA were a non-significant Session × Group interaction, $F < 1$, Stage-2 Discrimination × Congruence interaction, $F(1, 12) = 1.36$, $p > .10$, and 3-way interaction, $F(2.7, 32.9) = 2.56$, $p = .076$.

Discussion

The superior Stage-2 performance by the groups receiving the same discrimination in both stages, relative to the groups receiving different discriminations, indicates there was a degree of generalisation from the stimuli used in Stage 1 to those used in Stage 2. This generalisation must have been based on the number of squares within each stimulus but, because of the differences between the stimuli used for the two stages, it follows that the dimension representing number was not confined to squares of a specific colour. Instead, the dimension of number permitted generalisation between similar quantities, even though the characteristics of the objects belonging to those quantities were physically very different. Such a conclusion indicates that the representation of quantity is not tied to the concrete

properties of the training stimuli but, instead, can mediate generalisation in a more abstract manner. In the absence of further evidence, it is not possible to specify the nature of the abstract manner in which pigeons represent information about different quantities, but I would not want to argue that they do so by means of counting.

In keeping with conclusions drawn from Experiments 11 and 12, it is hard to explain the present pattern of results if the original discrimination was based on the overall brightness of the training patterns for two reasons. First, the use of the negatives of the training patterns in the final stage ensured that the level of brightness of the patterns in the two stages was very different, and any generalisation between the patterns based on absolute levels of illumination would be slight. Second, if the original discrimination was based on the relative levels of brightness of the two classes of pattern, then transfer to the test patterns should have been superior in the two groups trained with the opposite discriminations in the test than the training stages. This prediction follows because for these groups the brighter patterns, and the dimmer patterns, in both stages signalled the same outcome.

General Discussion

The experiments confirm that the asymmetry observed in magnitude discriminations involving time, auditory intensity, and the length of an object, can also be found with discriminations in which different numbers of identical objects signal the presence or absence of reward. It would thus appear that the benefit of using a large magnitude to signal reward, and a small magnitude to signal the absence of reward, relative to when the opposite is true, may well be a characteristic of magnitude discriminations in general.

The results from Experiment 10, consistent with the pattern of results from Chapter 3 with rats, demonstrate that by manipulating the stimulation during the ITI it is possible to reverse the asymmetry that is normally found with magnitude discriminations. Thus, rendering the ITI more similar to the larger of the two training stimuli resulted in the 5+/20- discrimination being acquired more readily than 20+/5-. This pattern of results is entirely in keeping with the proposal of Kosaki et al. (2013) that inhibition generalizing from the stimuli present during the ITI to S+ and S- is responsible for the asymmetry observed with magnitude discriminations.

Despite their consistency with the proposals of Kosaki et al. (2013), the results from the present chapter contrast considerably with those from Chapter 2. In particular, Experiments 4 and 5 showed that manipulating the stimulation during the ITI, so that this period was of greater intensity than the experimental stimuli, did not reverse the asymmetry. In these instances, presenting a loud clicker during the ITI did not result in a no clicker+/soft- discrimination being acquired more readily than a soft+/no clicker- discrimination, or a soft+/medium- discrimination being acquired more readily than a medium+/soft- discrimination. A discussion of why the results from Chapters 3 and 4 show conflicting results with Chapter 2 will be postponed until the General Discussion in Chapter 5.

Another prediction that can be drawn from the proposals of Kosaki et al. (2013), which was addressed in Chapters 2 and 3, is that an asymmetry in a magnitude discrimination will be observed only if there is an opportunity for cues, other than those on which the discrimination is based, to enter into an inhibitory association. This requirement was met in the present experiments. It was also met in the experiments by Kosaki et al., in which rats had to choose between long and short

black panels in order to find a goal. The panels were pasted to the grey walls of a square pool and any approach to the walls by themselves necessarily resulted in failure to find the goal and thus the opportunity for inhibitory conditioning. An obvious way to test this prediction with the dimension of number would be to present pigeons with a discrimination between 40 and 5 black squares using the design used in Experiment 8. For two groups the presentations of stimuli would be separated by an ITI in which no black squares are presented while for two further groups trials would not be separated by an ITI. However, at least two unpublished pilot investigations conducted by Inman found that pigeons made almost no key presses when S+ and S- were not separated by an ITI, and have therefore not been included in this thesis.

CHAPTER 5

General Discussion

As outlined in Chapter 1, many theories of conditioning are based on the assumption that generalisation gradients are symmetrical (e.g. Blough, 1975; Spence, 1936). A clear prediction that follows from this assumption is that the rate at which a discrimination between two stimuli is acquired should be unaffected by which stimulus signals the reinforcer. In contrast to this common prediction, a small body of evidence described in Chapter 1 suggests that there is an asymmetry when subjects are required to discriminate between two stimuli differing in magnitude. The primary purpose of this thesis was therefore to assess the reliability and generality of this asymmetry. Having successfully demonstrated the asymmetry with discriminations based on quantity and auditory intensity the following discussion will mostly deal with one question: How might theories of conditioning account for these findings? For the sake of brevity, only Pearce's (1987, 1994) configural theory and the Rescorla-Wagner (1972) model will be discussed in detail. What will become apparent is that neither model, in fact, can account for all the data presented in Chapters 2 to 4. In view of this disparity, I will then discuss the differences between stimuli differing in auditory intensity and those differing in quantity in terms of how they are represented at the receptor level. Based on these differences, I then propose a novel theory, based on the principles described in Pearce (1994) that may go some way toward offering a unified account for the results presented in this thesis. Finally, I will briefly address how pigeons and rats were representing number in Chapters 3 and 4.

Theoretical Accounts for Data

Thus far the explanation for the results in terms of stimulus generalisation has been derived from an informal account put forward by Kosaki et al. (2013)

which, itself, was based on Spence's (1936, 1937) account of discrimination learning. Is it possible to explain the results more formally in terms of more recent theories of learning? A challenge posed by this question is that to my knowledge a formal account of stimulus generalisation along a dimension of magnitude does not exist. One way of addressing this problem is to follow the proposals of Bouton and Hendrix (2011) concerning the way in which animals solve discriminations between stimuli of different durations. In essence, they suggested that a short duration stimulus is composed of one element, say A, and a long duration stimulus is composed of a succession of elements, say A followed by B. As far as the present experiments are concerned, it might then be suggested that the dimensions of auditory intensity (Chapter 2) and number (Chapters 3 and 4) can be represented by an increasing number of distinctive elements. The dimension of auditory intensity of a clicker, say, would be anchored by one element, A, to represent the cues present during the absence of a clicker, two elements, AB, to represent the additional cues provided by a soft clicker, three elements, ABC, to represent a louder clicker, ABCD representing an even louder clicker, and so on. The number of black squares on a white screen might also be represented in the same manner, with A representing zero black squares, AB representing 5 squares, ABC representing 20 squares, ABCD representing a quantity of squares greater than 20, and so on. Using this characterisation it is then possible to apply current theories of learning to the experiments presented in this thesis. For simplicity I will discuss two of such theories; Pearce's (1994) configural theory and the Rescorla-Wagner (1972) model. Because of its similarity to the proposals put forward by Kosaki et al., the first account to be discussed will be that of Pearce. However, as it will become clear, this

formulation of the experiments may not be satisfactory for providing a unified account for all the results in this thesis.

Configural Theory

Number

In Chapters 3 and 4, subjects were presented with stimuli comprising different quantities of black squares, AB (5 squares) and ABC (20 squares), the presentations of which were separated by an ITI, A. According to the theory of Pearce (1994), the presence of the common element A throughout each session will permit generalisation among the three patterns (the two training stimuli and the ITI) in any experiment, with generalisation being greater for patterns representing magnitudes that are close together. Equation 4 describes how, according to Pearce (1987, 1994) conditioning will progress to a pattern (in this instance AB) on any given conditioning trial during an AB+ ABC- A- discrimination (5+/20-).

$$\Delta V_{AB} = \beta \times (\lambda - ((_{AB}S_A + V_A) + V_{AB} + (_{AB}S_{ABC} \times V_{ABC}))) \quad (4)$$

$$_{AB}S_A = N_C^2/(N_{AB} \times N_A) \quad (5)$$

Here, the change in associative strength to a stimulus, AB, ΔV_{AB}, is determined by the difference between an asymptote in learning, λ, and what has already been learned (the associative strength of the stimulus, V_{AB}, plus the associative strength that generalises from similar stimuli). The generalisation coefficients for similar stimuli (in this example, the similarity of stimulus A to AB) can be calculated using Equation 5, where Nc refers to the number of elements common to A and AB and N_{AB} and N_A, refer to the number of elements in AB and A respectively.

According to this theory, during a discrimination, different patterns of stimulation, including that provided by the ITI, enter into associations with the outcome with which they are paired. Thus a discrimination of the kind AB+ ABC- A- (5+/20-) will result in AB entering in to an excitatory association which will generalise to A because of their similarity. The nonreinforced ITI stimulation, A, will then enter into an inhibitory association that will counteract the excitation that generalises to it. This inhibition will, in turn, generalise to AB and by virtue of disrupting the manifestation of its excitatory properties, hinder the acquisition of the AB+ ABC- discrimination. A similar process will occur in the group trained with ABC+ AB- A- (20+/5-). The considerable difference in similarity between A and ABC will, however, result in relatively little generalisation of excitation from ABC to A which will then need to acquire rather little inhibition. Moreover, little of this inhibition will then generalise to ABC and its disruptive impact on the ABC+ AB- discrimination will be minimal. As a consequence, the theory predicts that the ABC+ AB- discrimination will be acquired more readily than AB+ ABC-. In order to confirm that this prediction does indeed follow from the theory, computer simulations based on Equations 4 and 5 were conducted for a AB+ ABC- A- and a ABC+ AB- A- discrimination. The results from simulations in which the values of .2 and .1 for the learning rate parameter, β, can be seen in Figure 33. In line with the theory, the simulations presented in this figure predict that the difference in associative strength between S+ and S-, despite being small, is greater for the 20+/5- discrimination (right-hand panel) than for the 5+/20- discrimination (left-hand panel).

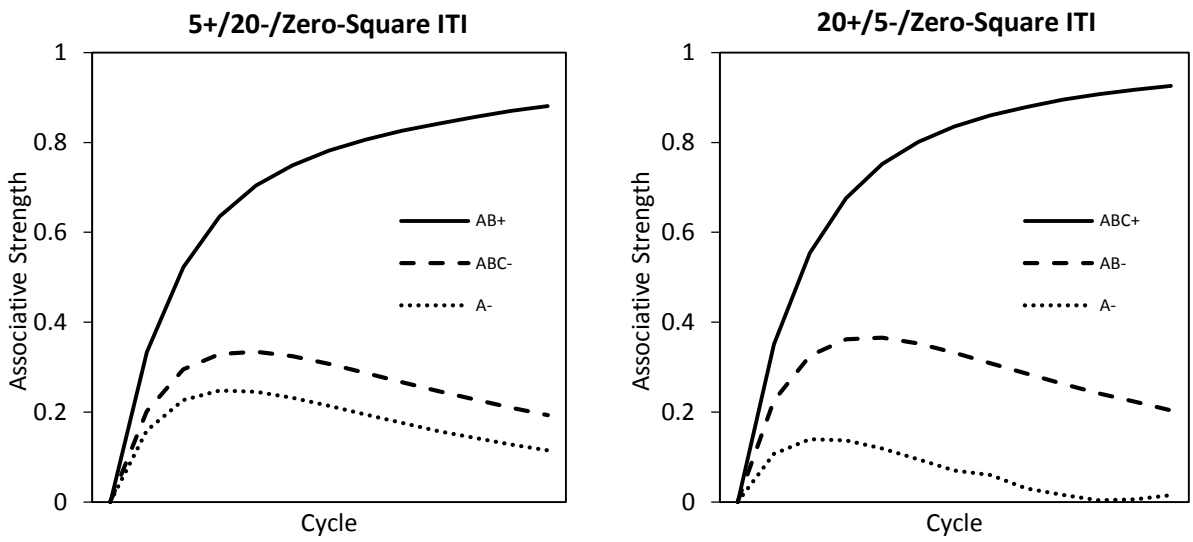

Figure 33. Computer simulations of an ABC+ AB- A- discrimination (left-hand panel), and an AB+ ABC- A- discrimination (right-hand panel) using the equations proposed by Pearce (1987). A- for both of these simulations represents an ITI in which zero black squares are presented.

Pearce's (1987) theory is also able to account for the data obtained from the various manipulations that were performed on the basic experimental design in Chapters 2 and 3. Turning to Experiments 7 (Chapter 3; rats), and 10 (Chapter 4; pigeons), in which the ITI was a white screen covered in a number of black squares greater than that for S+ and S-, the characterisation of S+ and S- would be the same as Experiments 6 (Chapter 3), and 9 (Chapter 4), in which the ITI was a white screen, i.e. AB and ABC. However, in this case the stimulation during the ITI would be ABCD-, rather than A-, to take account of the larger number of squares than during either training stimulus. Once again, computer simulations reveal that the theory of Pearce (1994) correctly predicts the reversal of the asymmetry: The left-hand panel of Figure 34 shows that the AB+ ABC- ABCD- discrimination (5+/20-) is acquired more readily than the ABC+ AB- ABCD- discrimination (20+/5-) seen in

the right-hand panel of the same figure. When it comes to Experiment 8 (Chapter 3; rats) in which subjects were given presentations of S+ and S- without an ITI, training can be characterised as ABC+ AB- and AB+ ABC-. In these circumstances, the theory correctly predicts that both discriminations will be acquired at the same rate. In other words, these computer simulations result in lines that are exactly superimposed.

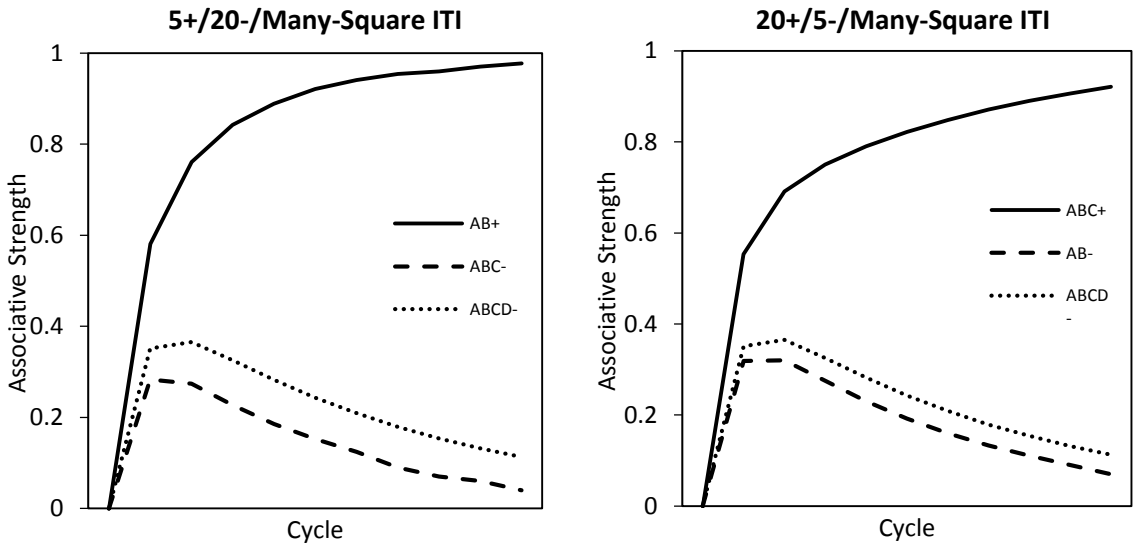

Figure 34. Computer simulation of an AB+ ABC- ABCD- discrimination (left-hand panel), and an ABC+ AB- ABCD- discrimination (right-hand panel) using the equations proposed by Pearce (1987). ABCD- for both these discriminations represents the numerous black squares that were presented during the ITI.

Auditory Intensity

The results from Chapter 2, with the exception of the initial demonstration of the asymmetry in Experiment 1, however, pose a challenge for an analysis in terms of Pearce's (1987) configural theory. In contrast to the findings from the equivalent experiments in Chapters 3 and 4, making the pattern of stimulation during the ITI

more similar to the high-intensity, rather than the low-intensity clicker, did not reverse the asymmetry. Loud+/soft- discriminations remained easier to acquire than soft+/loud- discriminations. Moreover, in contrast to predictions, Experiments 4 and 5 demonstrated that the asymmetry seen in the discrimination between clickers of different intensities was not dependent upon the presence of a nonreinforced ITI. This pattern of results, is not predicted by Pearce´s model; but might it be accounted for in terms of a different theory of conditioning?

Rescorla-Wagner (1972) Theory

Number

A further set of simulations was conducted using the equations proposed by Rescorla and Wagner (1972; see Equation 6) to predict the changes in associative strength of individual elements in the training patterns described above. Here, the change in associative strength of a stimulus, say A, ΔV_A, is the difference between an asymptote in learning, λ, and the sum of the associative strength of all stimuli presented during the trial, V_T. This value is modified by the salience of the CS, α, and the salience of the US, β. The values of β for reinforced and non-reinforced trials were, as above, .2 and .1 respectively, and the ITI was treated as a third CS.

$$\Delta V_A = \alpha.\beta.(\lambda - V_T) \quad (6)$$

Once again, the asymmetry in acquisition observed between 20+/5- (ABC+ AB- A-) and 5+/20- (AB+ ABC- A-) discriminations with an ITI consisting of zero squares in Chapters 3 and 4 was predicted successfully, as can be seen in Figure 35.

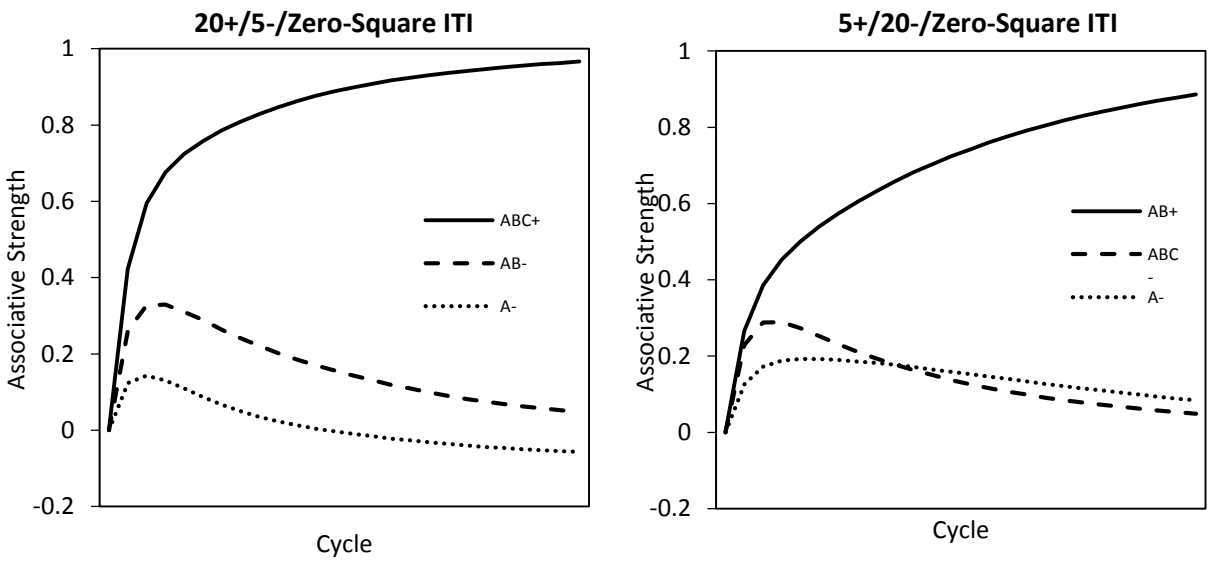

Figure 35. Computer simulations of an ABC+ AB- A- (left-hand panel) and AB+ ABC- A- discrimination (right-hand panel) using the equations proposed by Rescorla and Wagner (1972). A- for both these discriminations represents an ITI in which zero black squares were presented on a white background.

However, in the case of Experiment 7 (Chapter 3; rats) and Experiment 10 (Chapter 4; pigeons), where the ITI was made to be of higher magnitude than S+ and S-, the fit between the predictions and results was less than satisfactory (See Figure 36). During the early stages of training, the 20+/5- discrimination is predicted to be acquired more readily than 5+/20-. As can be seen from the results of Experiments 7 and 10, the opposite pattern was obtained. As far as Experiment 8 (Chapter 3; rats) is concerned where the effect of removing the ITI was examined, the Rescorla-Wagner (1972) theory predicts that an asymmetry should have been observed with the no-ITI groups. The removal of the ITI means the training can be characterised as ABC+ AB-, or AB+ ABC-. The first of these discriminations is essentially a *feature-positive* discrimination and the second a *feature-negative* discrimination. As Bouton & Hendrix (2011) point out, the Rescorla-Wagner theory predicts the former will be

acquired more readily than the latter. The fact that no asymmetry was found therefore poses a significant problem for this account when applied to discriminations based on number.

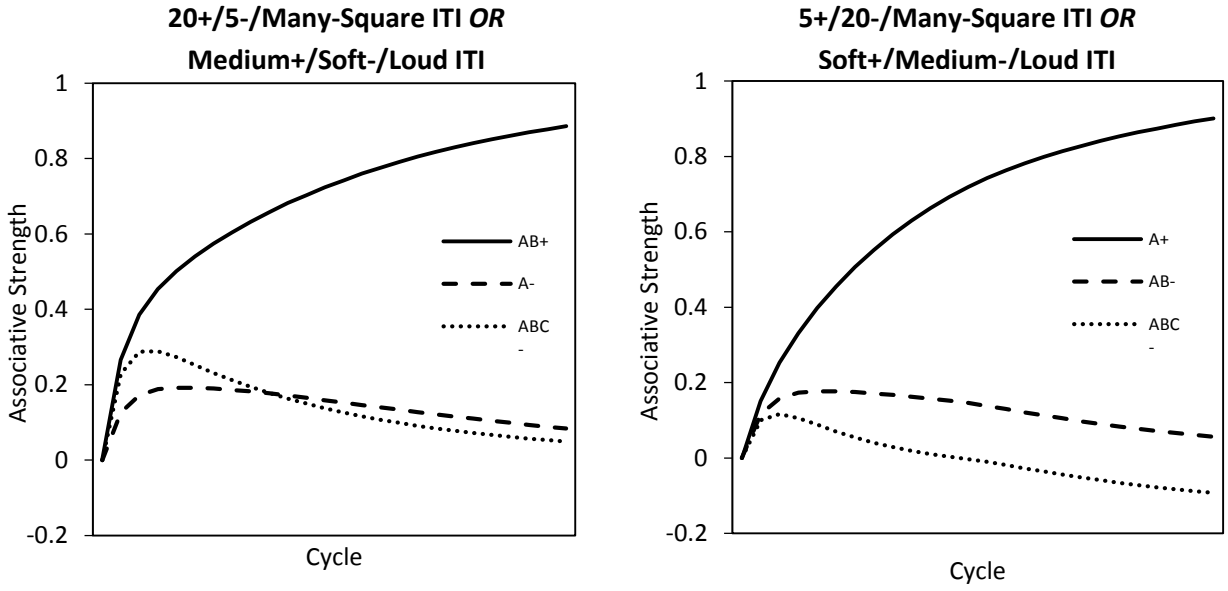

Figure 36. Computer simulations of an A+ AB- ABC- (right-hand panel) and an AB+ A- ABC- (left-hand panel) discrimination using the equations proposed by Rescorla and Wagner (1972). For both discriminations ABC- represents the ITI in which numerous black squares were presented on a white background, or, in the case of auditory intensity, the loud clicker ITI.

Auditory Intensity

In contrast, the Rescorla-Wagner (1972) theory can readily account for the results from Chapter 2 where rats were required to discriminate between clickers of different intensities. In Experiment 3, where the ITI was of higher intensity than S+ and S-, the medium+/soft- discrimination, AB+ A- ABC-, was acquired more readily than the soft+/medium- discrimination, A+ AB- ABC-. With reference back to the computer simulations presented in Figure 36, it is clear that the Rescorla-Wagner model is able to predict these results. In addition, although the Rescorla-Wagner

model is unable to predict the results obtained in Experiment 8, where removal of the ITI eliminated the asymmetry between 20+/5- and 5+/20- discriminations, it is able to predict the contrasting results obtained from the equivalent experiments with auditory intensity (Experiments 4 and 5). In these cases, the removal of the ITI did not eliminate the asymmetry, a result entirely in keeping with the *feature-positive* analysis described above.

Stimulus Salience

So far, I have considered the magnitude of a stimulus as being determined by the number of its elements. Thus, a stimulus with more elements, such as ABC, is of greater magnitude than a stimulus made up of fewer elements, such as stimulus A. An alternative to this is to characterise magnitude in terms of salience, determined in the equations proposed by Rescorla and Wagner (1972; Equation 6) by the value denoted by alpha (α).

If it is assumed that a stimulus with a large degree of perceptual difference to background cues has more relative salience (in other words, a larger α) than a stimulus with a smaller perceptual difference to background cues, then the Rescorla-Wagner (1972) model can readily account for the results presented in chapters 3 and 4. In Experiment 6, rats were presented with either a 20+/5- or 5+/20- discrimination where the ITI was a white screen containing zero squares. As 20-square stimuli are more dissimilar to the ITI than 5-square stimuli, and thus have more relative salience, it is possible to represent these discriminations as AX+BX-X- and BX+AX-X-, where A (20-squares) has a greater value for α than B (5-squares). X in this instance represents background cues. In this instance, the Rescorla-Wagner

model predicts that the AX+ component of the AX+BX-X-discrimination will acquire associative strength more rapidly than BX+ in the BX+AX-X-discrimination, and thus AX+BX-X- will be acquired more readily than BX+AX-X-. Assuming that the salience of a stimulus is determined by it similarity relative to the background cues, then this account can readily explain all the results from Chapters 3 and 4.

The salience account, however, struggles to account for other experimental findings. First, Experiments 4 and 5 in Chapter 2 demonstrate that presenting a clicker of louder intensity than S+ and S- during the ITI did not reverse the direction of the asymmetry. This occurred despite the absence of a clicker (Experiment 4) being more different to a loud ITI than a soft clicker, and thus having greater relative salience. One possibility is that, unlike stimuli differing in the number of squares, stimuli differing in auditory intensity have some degree of intrinsic salience dependent on their absolute, rather than relative, magnitude. Thus, if this were the case, a high intensity S+ would always acquire associative strength more rapidly than a low intensity S+. Secondly, defining magnitude in terms of α leads to the incorrect prediction that a discrimination between two high magnitude (and thus highly salient) stimuli will be acquired more readily than a discrimination between two lower magnitude stimuli, if the magnitude of the difference between the two stimuli for one discrimination is the same as that for the other discrimination. This prediction contradicts both Weber's law, and the findings of Kosaki, Jones and Pearce (2013) who found that a discrimination between two wall panels of different lengths was solved more readily by rats when the panels were 15 cm and 45 cm rather than 70 cm and 100 cm.

Variations in the value of β

For all of the above simulations, Pearce's (1987) theory and the Rescorla-Wagner (1972) model, in keeping with arguments put forward by Wagner and Rescorla (1972) and Jones and Pearce (2015), the value for the learning rate parameter, β, which plays the same role in both theories, was greater for reinforced than nonreinforced trials. If the same value of β is used for the reinforced and nonreinforced trials then the Rescorla-Wagner equation predicts quite well the results from experiments in which the ITI is made more similar to the high-magnitude stimulus than the low-magnitude stimulus (e.g. Experiment 7), as well as the other experiments. Varying the values of β for the reinforced and nonreinforced trials also influences predictions derived from the theory of Pearce (1987, 1994). The asymmetry that was observed in each experiment continues to be correctly predicted by the theory, but the magnitude of the asymmetry becomes smaller as the value of β for reinforced trials decreases and that for the nonreinforced trials increases. Thus, it appears that the predictions made by Pearce are robust to changes to β.

The role of the ITI

The foregoing analysis has two implications for the rate of responding during the ITI. Consider the basic design where S+ and S- are separated by an ITI, A, with no additional stimulation. First, as the inhibition that is predicted to be associated with the cues present during the ITI gains in strength, so responding during these intervals will become progressively weaker. Second, the rate of responding during the ITI is predicted to be stronger for the groups trained with the ITI cues more similar to S+ than S- (e.g. 5+/20-), than for the groups trained with the opposite

arrangement (e.g. 20+/5-). Unfortunately, the overall pattern of results from Chapters 3 and 4 makes it hard to evaluate these predictions. For Experiments 9 to 13 (Chapter 3; discrimination of number by pigeons) the rate of responding during the pre-CS periods was so slow, with many subjects making zero responses, that it is debatable whether any meaningful theoretical conclusions can be drawn from them. In Chapter 2 (discrimination of number by rats), both predictions were confirmed in Experiment 6, but the small within-group and between-group differences do not present compelling support for the theory from which the predictions were derived. In Experiment 7, the between-group differences were opposite to that predicted, while only one group showed a clear decline in responding as training progressed. Finally, in Experiment 8, the ITI response rates of approximately two responses per minute were again too slow for any meaningful theoretical interpretation. Overall then, the response rates during the ITI from the Experiments in Chapters 3 and 4 are of little interest theoretically.

In summary, it therefore seems that neither theory can adequately explain all of the results presented in this thesis. The results in Chapter 2 can be accounted for in terms of the Rescorla-Wagner (1972) model but not by Pearce's configural theory (1987), while those in Chapters 3 and 4 can be accounted for readily by Pearce, but less well by the Rescorla-Wagner theory. It may well be the case that the asymmetry observed for auditory intensity has different underlying mechanisms than that for quantity. The following section will outline one reason for why this may be the case.

The Representation of Magnitude

Thus far, I have assumed that an increase in auditory intensity, and in number, excites hypothetical elements in the same way with the lowest magnitude

stimulus represented by a single element, A, and increases in magnitude represented by the addition of new elements: AB, ABC, and so on. To what extent is this consistent with what is known about the effects on neurons of changes in magnitude?

According to Ghirlanda and Enquist (2003), when a stimulus is presented it stimulates receptors across the relevant sensory organ. Furthermore, as a stimulus changes along a dimension the activation of receptors can change in two characteristic ways; the pattern of activation across the receptors, and the degree to which these receptors are activated. The identity of a stimulus can thus be determined by the receptors that it activates and by how much it activates them. Take for example the presentation of a tone of a specific frequency and amplitude. When the tone is presented, receptors in the ear tuned to this frequency will be stimulated. It must also be noted here that receptors tuned to other frequencies will also be activated but to a lesser degree - the more dissimilar the tuning, the weaker the activation (see also McLaren & Mackintosh, 2000)[5]. In addition, the level of activation of these receptors will be dependent on the amplitude of the tone; as the intensity of the tone increases, so too will the activation of these receptors. Crucially, as the intensity of the tone increases, the increased activation of neurons will also result in a larger number of neighbouring receptors being activated.

Representation of Auditory Intensity

In Chapter 2, rats were presented with clickers that differed in intensity. Thus, in line with the above description, the soft clicker is expected to result in a weaker activation of fewer receptors, A, than the loud clicker which can be expected

[5] This pattern of activation can be described as a normal distribution.

to cause a stronger activation in more receptors, ABC. This pattern of activation is shown in Figure 37. Importantly, a high-intensity clicker will always activate all the receptors activated by a lower-intensity clicker.

How does this conceptualisation affect the predictions drawn above? If we accept that activated neurons can enter into associations and if we ignore changes on the y-axis (i.e. increases in neuronal activation), then not at all. The Rescorla-Wagner (1972) model is able to account for the data obtained with auditory intensity, while Pearce's (1987, 1994) configural theory, on the other hand, is not.

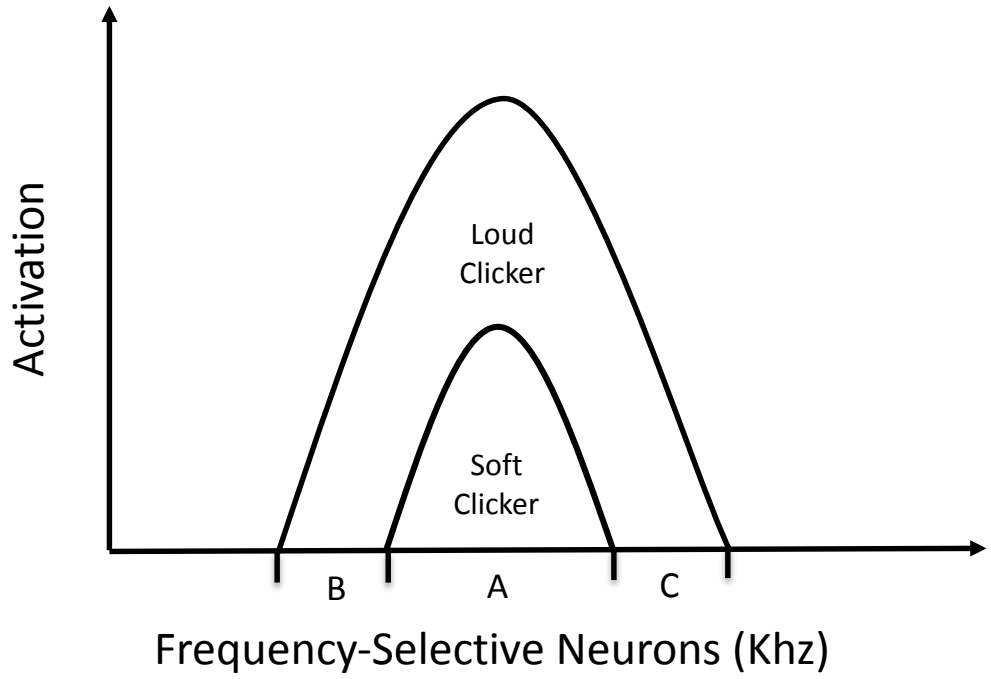

Figure 37. The representation of increased auditory intensity in receptor space. A = receptors activated by the soft clicker. B and C = additional receptors activated by the loud clicker.

Representation of Number

It is possible that an increase in the intensity of an auditory cue, and an increase in the number of objects displayed on a screen have different effects at the neuronal level. As we have just seen, an increase in the intensity of a clicker results in the same neurons being activated as those by a weaker clicker, but to a greater degree, and some additional neurons as well. In contrast, it has been shown that within the visual system there are numerosity-selective neurons that are tuned to the number of items in a visual display. That is, they show maximum sensitivity to a specific quantity, and a progressive drop off in activations as the quantity becomes more remote from the original value (Nieder et al., 2002; Nieder & Miller, 2004). Thus, in the same manner as for a change in the frequency of light, a change in the number of visual objects displayed will result in no change in the overall rate of firing of number-sensitive neurons, but a change in the pattern of neurons that fire and crucially a change in the identity of the neurons that fire maximally. Thus when there are no squares a zero-number receptor (assuming for now that they exist) will be activated. It also follows that when there are more squares there is less `zero-square´ space to be activated.

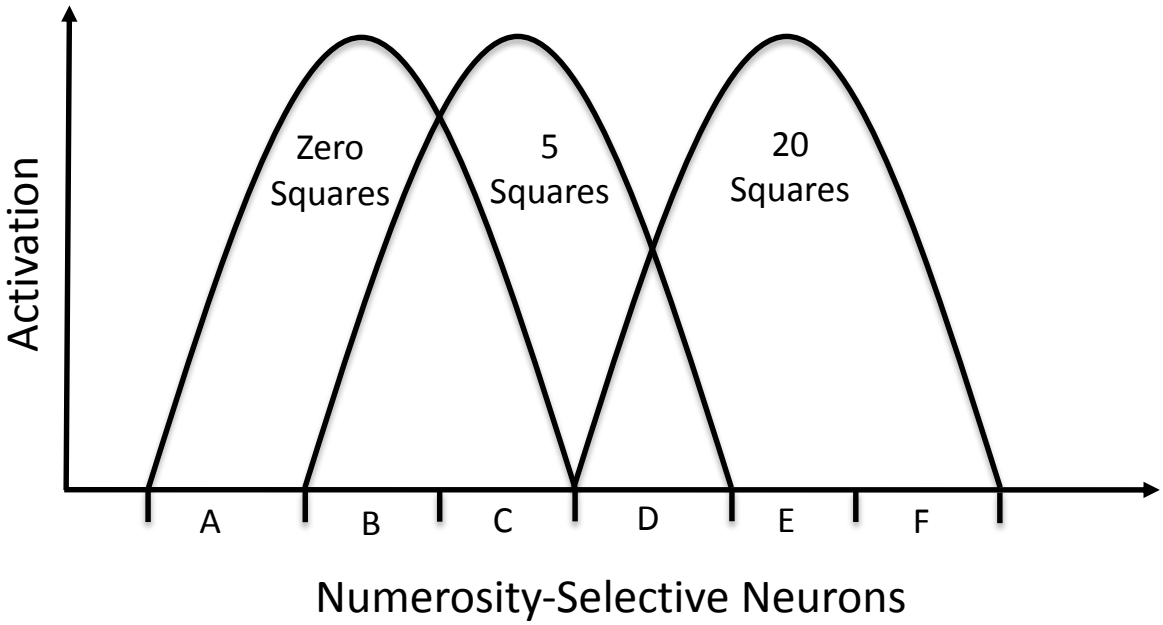

Figure 38. The representation of increased number in receptor space. A = receptors activated uniquely by zero-squares. B & C = receptors common to zero-square and 5-square stimuli. D = receptors common to 5-squares and 20-square stimuli. E & F = receptors activated uniquely by 20-squares.

Figure 38 shows that the presentation of a numerical stimulus activates a set of receptors. Importantly, unlike with an increase in auditory intensity, a change from smaller to larger numbers does not result in a larger number of receptors being active. In other words, a change in magnitude for the dimension of number represents movement along the X-axis of Figure 38, rather than along the Y-axis. What is evident from Figure 38 is that some receptors are common to zero-square and 5-square stimuli, C, and others common to 5-square and 20-square stimuli, E. Thus, a 5+/20- discrimination with an ITI in which no squares are presented can be described as ABC- BCD+ DEF-, and a 20+/5- discrimination with the same ITI as

ABC- BCD- DEF+[6]. The similarity between stimuli is determined by the number of shared elements, which here represents the number of shared receptors. How does this conceptualisation of stimuli affect the predictions made by the Rescorla-Wagner (1972) model and Pearce's (1987) configural theory?

Rescorla- Wagner (1972) Theory

Computer simulations using the equations proposed by Rescorla and Wagner (1972; Equation 6) were applied to ABC- BCD+ DEF- and ABC- BCD- DEF+ discriminations and are presented in the top row of Figure 39. These simulations reveal that the 20+/5- discrimination should be acquired more successfully than the 5+/20- discrimination when the ITI is composed of zero black squares. The lower panels of Figure 39 present simulations for the same discriminations but with an ITI comprising numerous black squares. In this instance, the simulations predict that the 5+/20- discrimination should be acquired more readily than the 20+/5- discrimination, although the predicted difference in acquisition between the two groups is small. This is unsurprising considering that for the zero-square ITI groups there should be a large amount of generalisation between the ITI (ABC) and the 5-square stimulus (BCD). Conversely, the amount of generalisation between the ITI (DEF) and the 20-square stimulus (BCD) for the many-square ITI groups is lower, and thus it follows that the asymmetry should be less pronounced.

[6] Of course, this is only one possible way in which numerical stimuli might be represented. The number of unique and common elements for each stimulus will depend on their relative positions along the dimension.

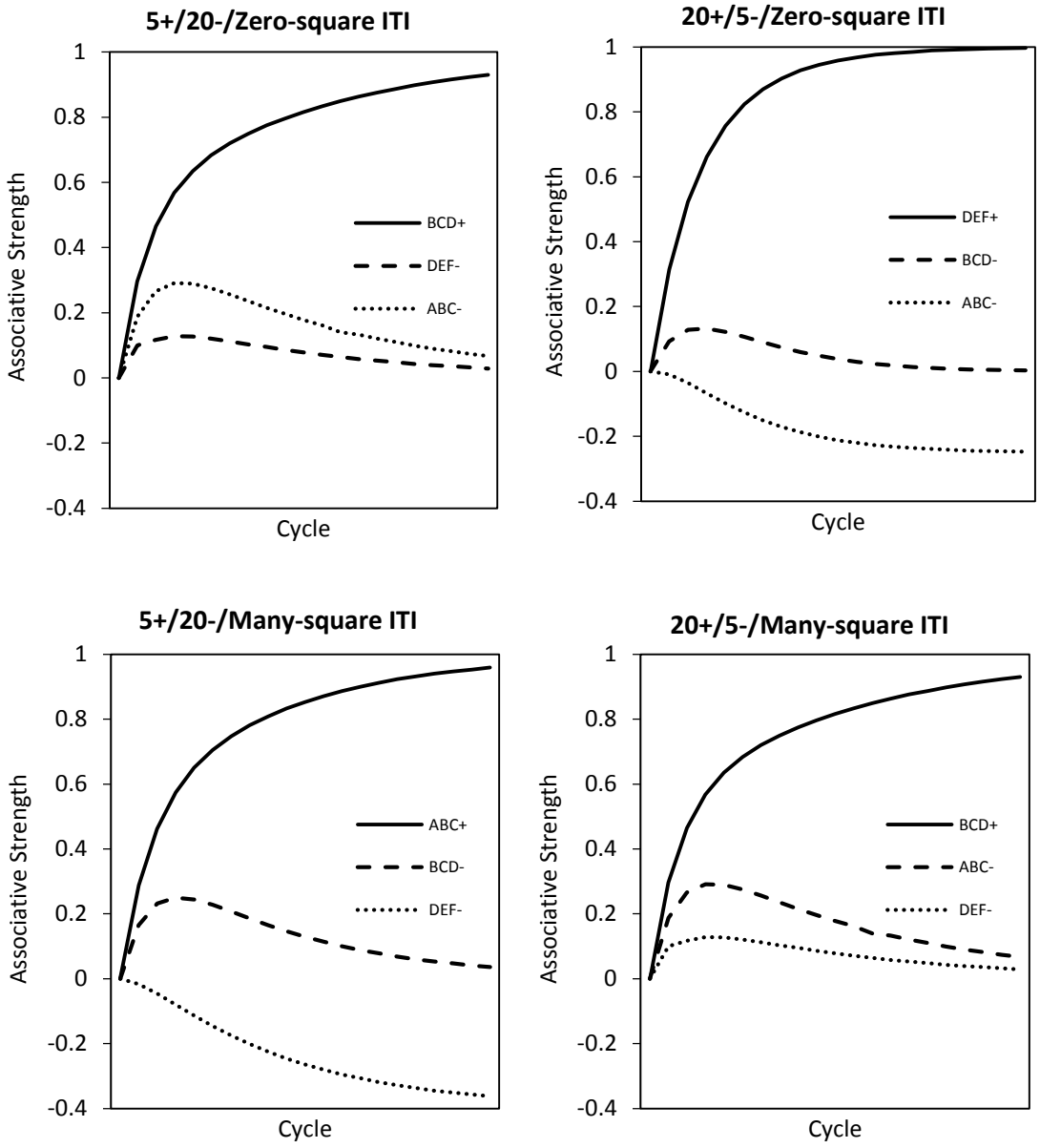

Figure 39. Computer simulations of a BCD+ DEF- ABC- (upper left-hand panel) and DEF+ BCD- ABC- (upper right-hand panel), and ABC+ BCD- DEF- (lower left-hand panel) and BCD+ ABC- DEF- discriminations (lower right-hand panel) based on the Rescorla-Wagner (1972) model. For the upper panels ABC- reflects an ITI in which zero black squares were presented. For the lower panels DEF- represents an ITI in which numerous black squares were presented.

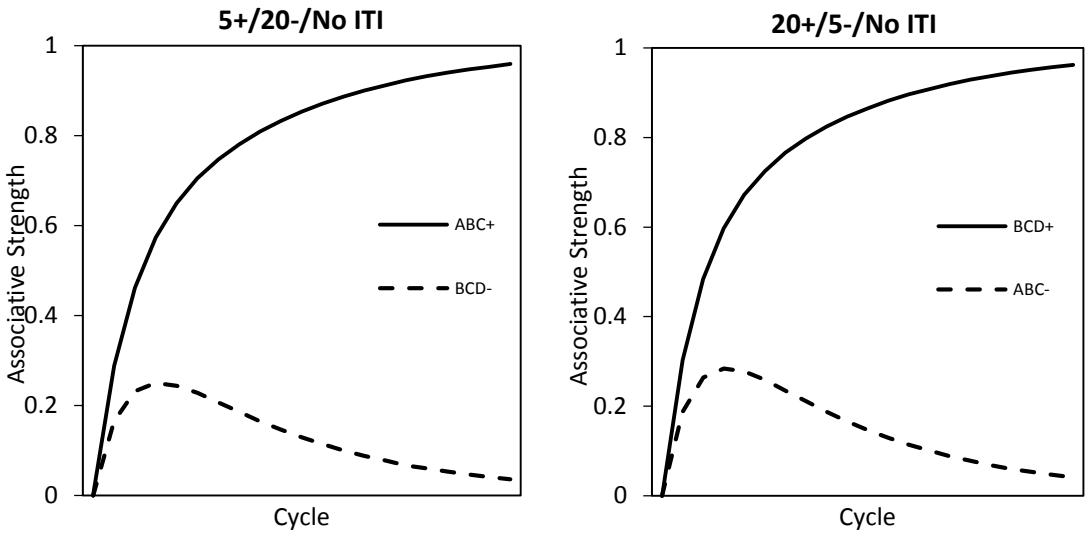

Figure 40. Computer simulations of an ABC+ BCD- (left-hand panel) and BCD+ ABC- (right-hand panel) discrimination based on the equations proposed by Rescorla and Wagner (1972).

Figure 40 presents simulations of the Rescorla-Wagner (1972) model when applied to a 20+/5- and 5+/20- discrimination in which S+ and S- are not separated by an ITI. As can be seen by a comparison of the two panels, the model predicts that these two discriminations should be acquired at the same rate. The Rescorla-Wagner model can therefore comfortably predict the data obtained in Chapters 3 and 4 when increments in number are considered as changes in the pattern of activation, rather than as increasing numbers of activated receptors.

Pearce's (1987) Configural Theory

Computers simulations were also conducted using the equations proposed by Pearce (1987). As can be seen in the top row of Figure 41, these simulations correctly predict that the 20+/5- discrimination should be more readily acquired than the 5+/20- discrimination when the ITI was a white screen with zero black-squares (although it must be noted that the asymmetry is slight). Pearce's model struggles, at least with this conceptualisation of numerical stimuli, to predict the reversed

asymmetry with an ITI comprising many black squares. As can be seen from the lower panels of Figure 41, any asymmetry between the S+ and S- components of the discriminations is so small that it might be considered to be negligible.

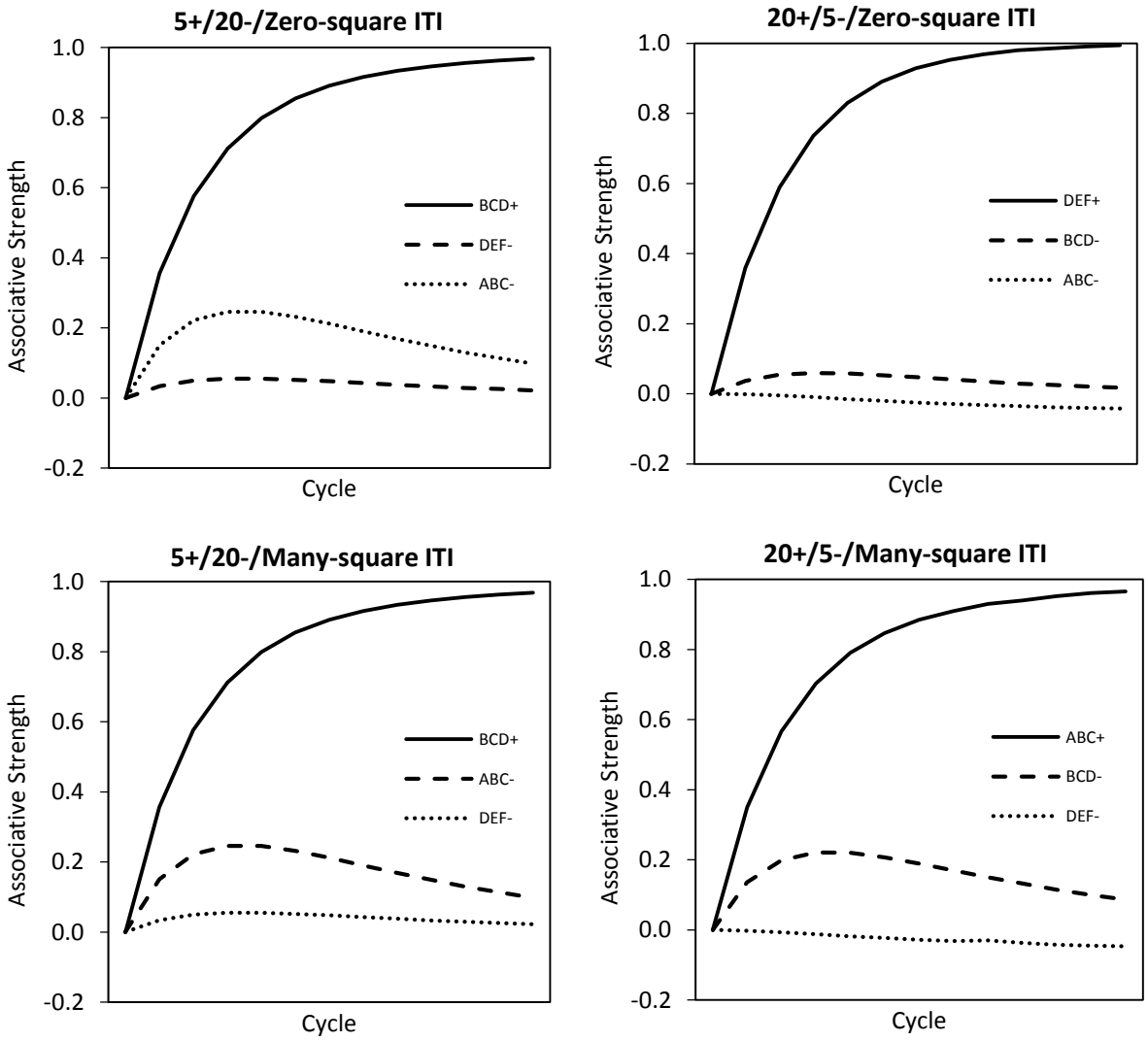

Figure 41. Computer simulations of a BCD+ DEF- ABC- (upper left-hand panel) and DEF+ BCD- ABC- (upper right-hand panel), and ABC+ BCD- DEF- (lower left-hand panel) and BCD+ ABC- DEF- discriminations (lower right-hand panel) based on Pearce's (1987) model. For the upper panels ABC- reflects an ITI in which zero black squares were presented. For the lower panels DEF- represents an ITI in which numerous black squares were presented.

The two simulations presented in Figure 42 predict that when S+ and S- are conditioned without an ITI the 5+/20- (ABC+ BCD-), and 20+/5- (BCD+ABC-)

discriminations are acquired at the same rate. Pearce's (1987) configural theory is therefore also able to account for the results obtained in Experiment 8.

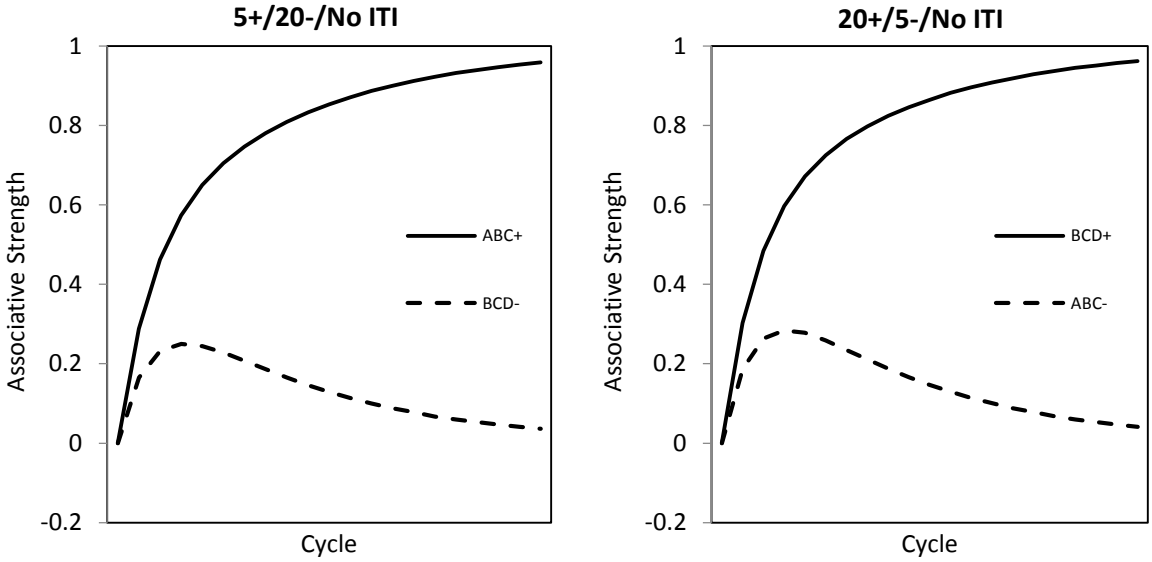

Figure 42. Computer simulations of an ABC+ BCD- (left-hand panel) and BCD+ ABC- (right-hand panel) discrimination based on the equations proposed by Rescorla and Wagner (1972).

Future Directions

One way to test the above proposals concerning the difference between auditory intensity and number is to make a direct comparison between two different changes in dimension using the same stimulus type. An obvious example might be to present two groups of subjects with a discrimination between a low- and high-intensity tone, and a further two groups with a discrimination between a high- and low-pitch tone. Tones of different intensities should be represented in the same manner as clickers of different intensities whereas pitch should be encoded in the same way as quantity due to the presence of pitch sensitive receptors in the ear. If the above proposals are correct, then an asymmetry should be present for the intensity and pitch groups assuming that a tone of lower intensity and lower pitch is presented

during the ITI. If a further four groups were given the same training with the exception that the ITI now contained a higher-intensity and a higher-pitch tone, then it follows that the asymmetry should reverse for the pitch groups, but not the intensity groups. Likewise, removal of the ITI should eliminate the asymmetry for the pitch groups but no the intensity groups.

Summary

When the stimuli are considered in terms of how they are represented at the receptor level, it is clear that the Rescorla-Wagner (1972) model can readily account for the results from Chapters 2 to 4. Pearce´s (1987) configural theory, on the other hand, struggles to make any correct predictions other than the presence of an asymmetry with discriminations between 5 and 20 black squares with an ITI comprising a white screen with zero black squares. It is thus tempting at this point to conclude that in terms of accounting for intensity effects in discrimination learning, the Rescorla-Wagner model is superior to that of Pearce. However, an interesting question is whether Pearce's theory can be modified in order to successfully predict the results presented in Chapters 2 to 4. Indeed, by altering just one assumption made by Pearce's theory, it is possible to offer an alternative account that is superior to the Rescorla-Wagner model.

The assumption in question is that a stimulus excites a set of receptors, and that generalisation from one stimulus to another is determined by the number of receptors they share. This assumption results in symmetrical generalisation gradients when the generalisation gradients are approximately of the same shape and height. However, if the amplitude of one stimulus is greater than the other, such as with clickers of different intensities, then, it follows from the spirit of the theory that

asymmetrical generalisation gradients are predicted. This prediction follows if greater activation is represented by a larger number of hypothetical elements than a weaker activation. Referring back to Figure 36 it is evident that after conditioning with a weak stimulus such as A, presentation of the strong one (e.g. ABC) will activate all the elements activated by the weak stimulus, and lead to a strong CR. Conversely, after conditioning with a strong stimulus, presentation of the weak stimulus will activate only a few of the conditioned elements and lead to a weak CR. Generalisation for intensity dimensions may thus not be determined by the number of shared elements, but by the proportion of common elements shared by both stimuli, to the total number of elements in the conditioned stimulus. Using these principles, I adapt Pearce's configural theory to provide a unified explanation for my results.

A Modification to Pearce's Configural Theory: A Formalisation

The amount of generalisation from two stimuli that differ in intensity, say, stimulus A to stimulus ABC, ${}_{A}S_{ABC}$, shown in Equation 7, determines the generalisation from training with A, when ABC is presented for testing (the stimuli A and ABC have been used here to be consistent with the proposed representation of auditory intensity described above).

$$ {}_{A}S_{ABC} = N_C/(N_C + N_A) \quad (7) $$

In this expression N_C is the number of elements common to A and ABC, and N_A is the number of distinctive elements in A that are not present in AB. Importantly, increases in neuronal stimulation can be regarded as additional elements because of the increase in the number of receptors activated. Similarly, ${}_{ABC}S_{A}$,

represents the generalisation from training with ABC, when A is presented for testing and is calculated using Equation 8.

$$_{ABC}S_A = N_C/(N_C + N_{ABC}) \quad (8)$$

Thus, take the example of a simple discrimination between stimulus A, which activates a pattern of receptors, and a more intense stimulus ABC, which activates all the receptors activated by A, plus two additional receptors. According to Equations 4 and 5, $_{A}S_{ABC}$ has a value of 1, while $_{ABC}S_A$ has a value of 0.33. Importantly, unlike the equations proposed by Pearce (1994), the outcome is that the generalisation between the two stimuli is not symmetrical.

The change in associative strength of A on a conditioning trial, ΔV_A, is then calculated using the same equations proposed by Pearce (1987; see Equation 9). Thus, the change in associative strength is determined by the difference between an asymptote in learning, λ, and what has already been learned, multiplied by the learning rate parameter β.

$$\Delta V_A = \beta \times (\lambda - (V_A + (_{A}S_{ABC} \times V_{ABC}))) \quad (9)$$

The overall associative strength of A, V_A, is therefore given by the expression given in Equation 10, where V_A and V_{ABC} represent the strengths of the associations of A and ABC respectively:

$$V_A = (V_A + (V_{ABC} \times {}_{ABC}S_A)) \quad (10)$$

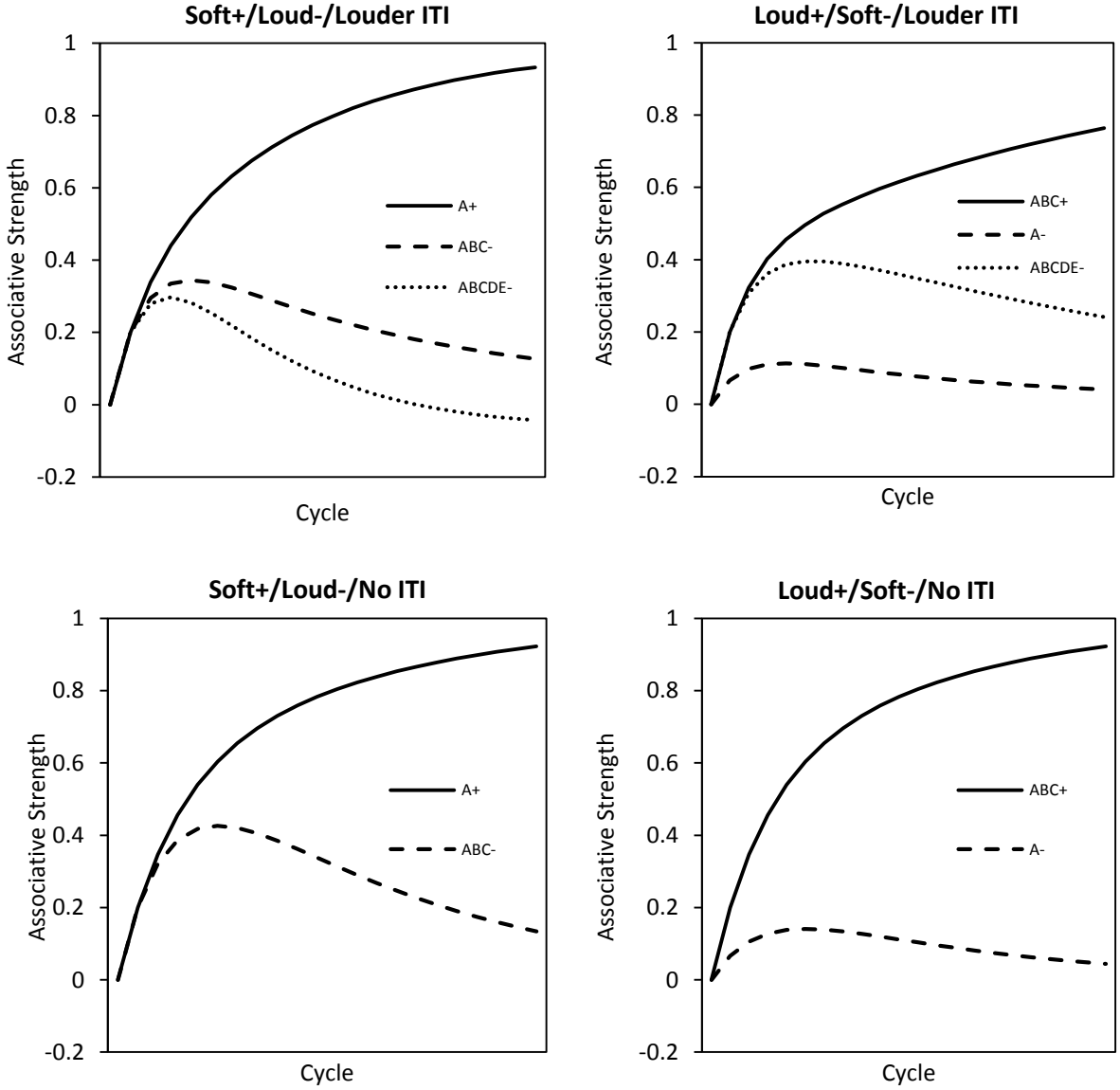

Figure 43. Computer simulations of the A+ ABC- ABCDE- and A- ABC+ ABCDE- discriminations (upper panels) and A+ ABC- and ABC+ A- discriminations (lower panels) using the modified version of Pearce's configural theory. The high-intensity ITI, ABCDE, represents the receptors activated by the loud clicker, ABC, plus elements D and E, which represent the newly activated receptors of higher and lower frequency.

Simulations based on the above equations were applied to discriminations in which the ITI was a clicker of greater magnitude than S+ and S-. In this instance, simulations were run with a third stimulus, ABCDE, that represents the activation of

receptors by ITI: A+ ABC- ABCDE- vs. ABC+ A- ABCDE- (upper panels of Figure 43). In keeping with my findings, a loud+/soft- discrimination is predicted to be acquired more readily than soft+/loud- when a very loud stimulus is presented during the ITI. As can be seen in the lower row of the same Figure, the ABC+ A- component of a loud+/soft- discrimination will be solved more readily than A+ ABC- from a soft+/loud- discrimination. Thus this model readily predicts the results from Experiments 2 and 3 in Chapter 2 and explains why the removal of the ITI in Experiments 4 and 5 did not eliminate the asymmetry.

This model is also able to predict with some success the results obtained with the dimension of number in Chapters 3 and 4. The simulations were applied to a 20+/5- discrimination with an ITI with zero-dots (ABC- BCD- DEF+), and to a 5+/20- discrimination with the same ITI, ABC- BCD+ DEF-. As can be seen in the upper row of Figure 44 the model predicts that the 20+/5- discrimination will be solved more successfully than the 5+/20- discrimination. In the same manner, this model predicts that after extended training a 5+/20- discrimination will be acquired more readily than a 20+/5- discrimination when the ITI comprises more squares than S+ and S- (ABC+ BCD- DEF- vs. ABC- BCD+ DEF-, see the lower panels of Figure 44), although the asymmetry is only slight due to the high amount of generalisation between S+ and S-. This model therefore predicts that the size of the asymmetry depends on the amount of generalisation between S+ and S-. Finally, Figure 45 demonstrates that this model predicts the absence of an asymmetry with a discrimination between 20 and 5 squares when no ITI is present. It must be noted that for this simulation elements A and D represent the receptors unique to the

patterns composed of 5 and 20 squares, respectively, while elements B and C represents receptors common to both patterns: ABC+ BCD- vs. BCD+ ABC-.

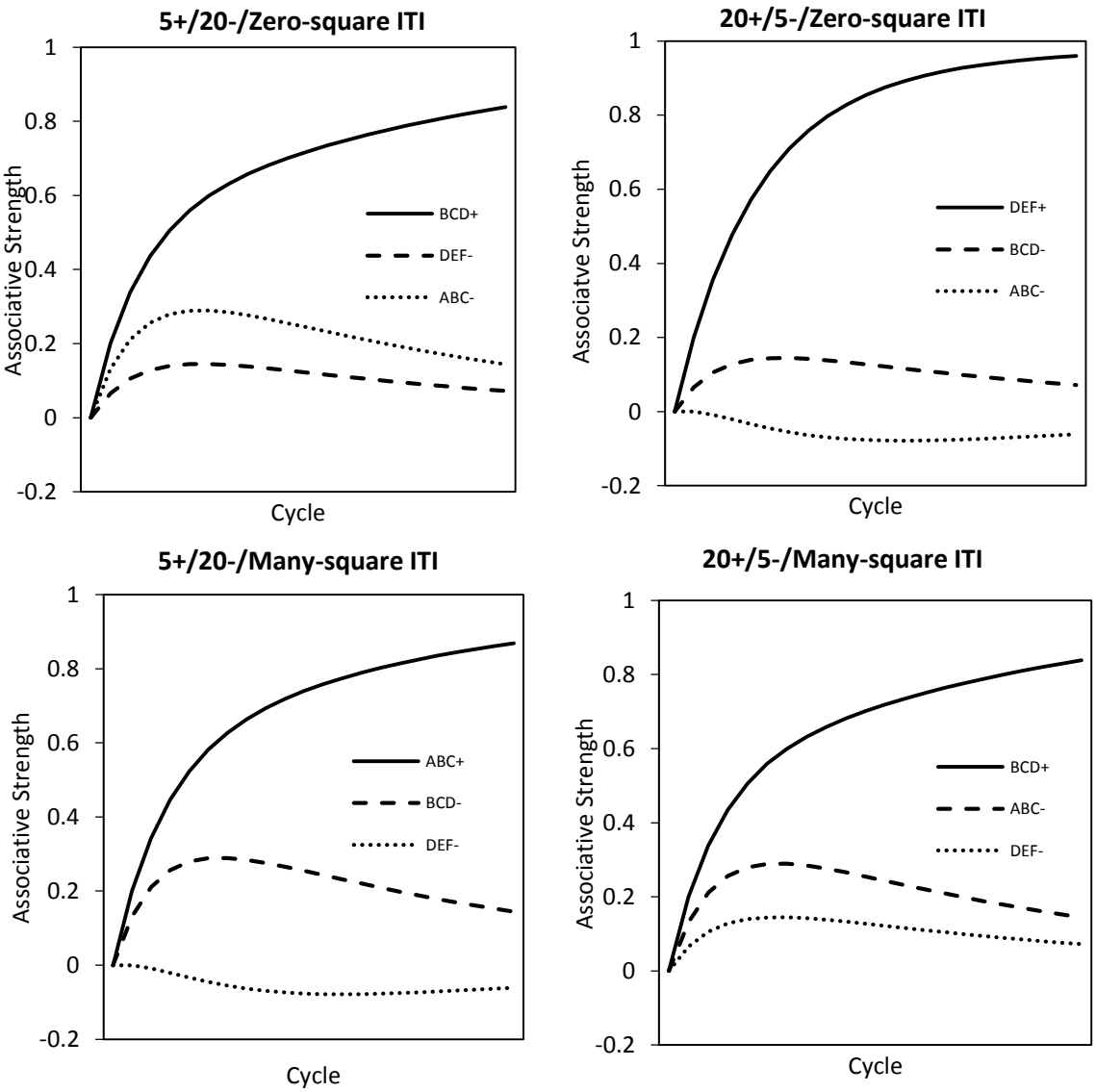

Figure 44. Computer simulations of the ABC- BCD+ DEF- and ABC- BCD- DEF+ discriminations (upper panels) and ABC+ BCD- DEF- and ABC- BCD+ DEF- discriminations (lower panels) using the modified version of Pearce's configural theory. For the upper panels ABC- represents the ITI in which no black squares were presented on a white background. For the lower panels DEF- represents the ITI in which numerous black squares were presented on a white background

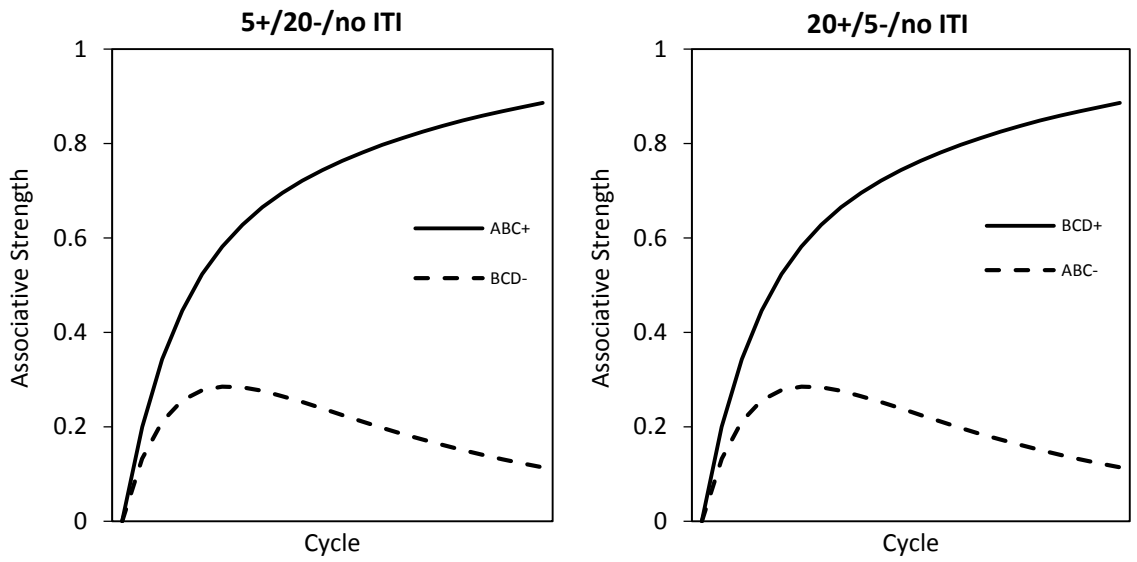

Figure 45. Computer simulations using the modified version of Pearce's configural theory of ABC- BCD+ vs. ABC+ BCD- discriminations.

Beyond the ability to account for the results from all the chapters presented in this thesis, this model has a number of other benefits. First, it correctly predicts that adding common cues will make a discrimination more difficult, in adherence to Weber's law. For example, as described earlier, Kosaki, Jones & Pearce (2013) demonstrated that rats can solve a discrimination between wall panels of length 15 and 45 cm more readily than when they were 70 and 100 cm. To illustrate this consider a 15 cm panel as stimulus A and a 45 cm panel as stimulus AB. According to Equations 4 and 5 the generalisation coefficient from A to AB will be 1, and .5 from AB to A. The addition of common element X to A and AB (thus AX and ABX) in order to account for the additional 55 cm added to each panel, results in a coefficient from ABX to AX of .66, indicating a greater amount of generalisation from the long to the short panel, thus rendering the discrimination more difficult. The generalisation coefficient from AX to ABX will remain 1. This is a particularly

important prediction because it suggests that this model can explain a result that cannot be accounted for by the Rescorla-Wagner (1972) model, which so far is also able to account for the results in this thesis, but incorrectly predicts that the discrimination between panels of 70 and 100 cm should be easier than panels of 15 and 45 cm.

The proposed model can also account for the phenomenon of summation in conditioning where the performance to two cues presented in compound, AB, is superior to that attained in training where the cues are presented independently in separate conditioning trials, A+ and B+ (Kehoe, 1982; Kehoe & Gormezano, 1980; Whitlow & Wagner, 1972). Here, the equations predict that generalisation of associative strength from each component to the compound will be 1; in other words there is complete generalisation from A, and B, to AB. Consequently, the associative strength of the compound is predicted to be the sum of the associative strength of each component. This prediction therefore contradicts Pearce's (1987) model, which is unable to predict summation (although it must be noted that within the summation literature there are numerous examples in which responding to a compound is similar to that of either compound element alone; Aydin & Pearce, 1994, 1995, 1997; Rescorla & Coldwell, 1995).

Finally, this model makes the prediction that there should be greater generalisation to a compound after training with a single component element (training with A+ followed by a test with AB), than to a component element after training with a compound (training with AB+ followed by a test with A). Considering that this prediction does not follow from the original version of Pearce's configural theory (1987, 1994), and yet that there is evidence that indeed there is

greater generalisation from elements to compounds than vice versa (e.g. Bouton, Doyle-Burr, & Vurbic, 2012; Brandon, Vogel, & Wagner, 2000) suggests that the proposed model is superior to that of Pearce.

It must also be noted that this modified version of Pearce's configural (1987, 1994) theory has also been found to make some predictions not consistent with data. One example is that it struggles to explain the greater levels of responding to test stimuli higher in intensity than S+ such as those seen in examples of monotonic generalisation gradients (e.g. Razran, 1949, Scavio & Gormenzano, 1974) and peak shift (e.g. Hanson, 1959). This is because generalisation to a higher magnitude stimulus after conditioning with a lower magnitude stimulus, say A+ to AB, will always be calculated to be complete, using the new generalisation coefficients describes above: ${}_{A}S_{AB} = 1/(1+0) = 1$. Thus responding to a higher magnitude stimulus can never, according to these equations, be greater than to a conditioned stimulus of lower magnitudes stimulus (although I might be added that the observed results pose a problem for all the currently influential theories of learning). Until all the predictions of this new model have been fully tested, it should be viewed with a degree of caution.

A failure to discriminate between S+ and S-

A prediction that is common to all the theories that have been considered is that the 20+/5- discrimination in Experiment 7 (Chapter 3; rats) would be solved, despite the presence of 80 squares shown on the screen during the ITI. Inspection of the left-hand panel of Figure 20 reveals that this prediction was not confirmed as there was no hint of the discrimination being solved, even after 14 sessions of training. Chapter 4 describes a similar failure to solve the equivalent discrimination

with pigeons (Experiment 10). All of the simulations that have just been described for both theories predict this discrimination should have been acquired readily irrespective of the values assigned to β. Is it possible to explain these unexpected results?

Inspection of the results from Experiments 7 and 10 show that the subjects from the 5+/20- groups responded more to S+ and S- than the ITI despite being unsuccessful at the discrimination. Thus the failure to solve the task does not seem to be a result of pigeons being unable to differentiate between the ITI and S+ and S-. One possibility is that pigeons in Experiment 10 differentiated between S+, S- and the ITI based on the average distance between squares.[7] With this in mind it is worth noting that the mean distance for adjacent squares was 10 mm, 6 mm, and 4mm for the patterns containing 5, 20 and 288 squares, respectively. Thus the ITI and 20-square patterns differed in average density by only 2 mm. The close similarity between S+ and the ITI in terms of density may well have made it hard to use this cue for solving the discrimination for the 20+/5- group. The fact that S+ for the 5+/20- group is more dissimilar to the ITI (a 6mm difference) than the S+ for the 20+/5- group may account for why the 5+/20- group was able to successfully solve the discrimination. In order to explain why the 20+/5- group responded more rapidly during the S+ than the ITI, it is worth noting that in Experiment 10, S+ and S- were perceptually different to the ITI because they comprised black squares surrounded by a large area of white background. The ITI, on the other hand, comprised a screen completely covered in black squares. (See the top panel of Figure 46). Thus, the pattern of results can be accounted for if it is considered that the 20+/5- group

[7] See Emmerton (1998) for the first systematic demonstration that numerosity discriminations by pigeons are sensitive to alterations in inter-dot density.

learned to respond in the presence of stimulus composed of black squares surrounded by a large area of white on a partial reinforcement schedule.

The same analysis might also be applied to Experiment 7 (Chapter 3) in which the average distance for adjacent squares was 4.5 cm, 1.8 cm, and 1.8 cm for the patterns containing 5, 20, and 80 squares, respectively. This arrangement made it impossible to discriminate between the ITI and the 20-square stimulus in terms of the density of squares. If this strategy was normally adopted in the other experiments then some other strategy was required for Experiment 7. It seems this strategy did not involve the number of squares, because group 20+/5- should have solved the discrimination between these values. Perhaps it was the area of empty white screen. Inspection of Figure 46 (lower row) shows that the area of white surrounding the black squares is considerably less for the ITI than for S+ and S-. Moreover, this difference for S+ and S- was relatively small, which would explain why the discrimination between the stimuli was not successful for the 20+/5- version of this task.

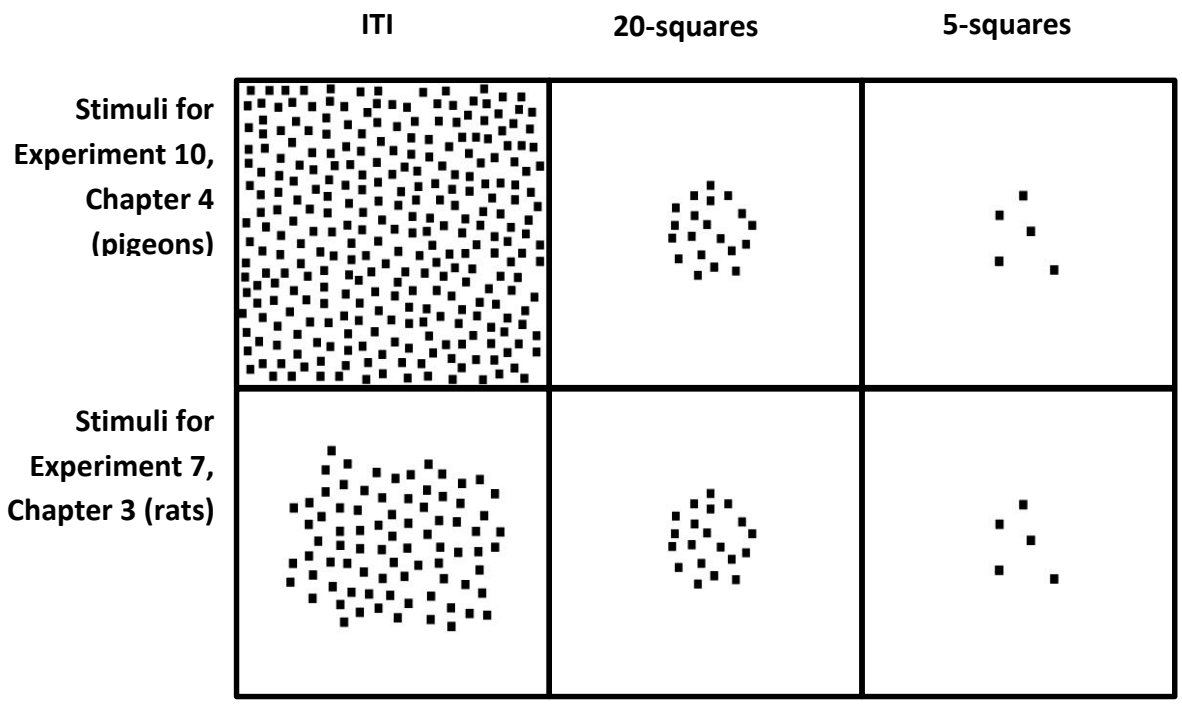

Figure 46. Illustration of the difference in perceptual properties for the ITI in Experiment 10, where 288 squares were covered the entirety of the screen, and Experiment 7, where 80 squares were clustered around the centre. In both experiments responding was slower during the ITI, and at a similar, higher rate for S+ and S-.

Implications for the Representation of Number

The above section poses the interesting question that the discriminations between 20 and 5 black squares may not have been solved based on the dimension of number per se. Inspection of the stimuli used for Experiment 6 (Chapter 3; rats), which are sketched in Figure 18, reveals that they can readily be ordered in ascending order of overall brightness starting with the 20-square pattern (comprising 20 black squares against a white background), then the 5-square pattern (comprising 5 black squares against a white background), and finally the display for the ITI (comprising a white background with no black squares). The results from Experiment 6 alone thus imply that the generalisation of inhibition took place along the basic, physical dimension of brightness. In this case an analysis similar to that applied to auditory intensity may be appropriate. However, the results from

Experiment 11 serve as a forceful counter to this analysis, because changing the stimulation during the ITI from white to black did not reverse the asymmetry, which should have been the case if the magnitude discrimination was based on differences in brightness.

The results from Experiment 11, as well as those from Experiments 12 and 13, thus led us to conclude that any generalisation among the different stimuli used for the experiments involved a less physical, more abstract, dimension than brightness, such as the number of squares displayed within a pattern. The dimension is described as abstract because Experiment 13 revealed substantial transfer of responding between similar numbers of squares, even though they changed in brightness from white to black (see Dumont et al., 2015).

Of course, the brightness of stimuli, the density of squares (as referred to in the above section concerning the failure to discriminate between S+ and S-), and number are not the only possible ways in which the discriminations were solved. Rather than refer to these dimensions, subjects may have relied on the total area occupied by the squares. Since the area occupied by each square was the same, regardless of the pattern to which it belonged, it then follows that the total area occupied by all the squares in a pattern would provide a suitable dimension for generalisation among the patterns. Given the considerable transfer in Experiment 13 between patterns comprising black squares on a white background, and those comprising white squares on a black background, it then follows that the representation of a given area of squares would need to be independent of their colour.

Apart from to say that the transfer effect observed in Experiment 13 suggests that the dimension used for generalisation should be abstract, or, in other words, independent of colour of the squares, it is difficult to make any firm conclusions about which of these dimensions is used based on the data presented in this thesis. Indeed, it is quite plausible that the arrays were represented in terms of multiple different dimensions. This would, nonetheless, remain entirely consistent with the proposal that a change in number equates to a change in the pattern of receptors activated (a qualitative change), rather than the number of receptors activated, and thus has little impact on the conclusions I wish to draw. Nonetheless, it would be interesting for future work to examine closely the extent to which pigeons use the numerical properties of stimuli instead of, or indeed alongside, other covariables such as area (see Emmerton, Lohmann, & Niemann, 1997, for a good demonstration that pigeons are able to solve relative quantity discriminations after controlling for summed area of dots, brightness and shape of dots).

Once it is accepted that the numerical stimuli were represented in an abstract manner then in order to explain the results, it is necessary to assume that when no squares are on the screen (and also when no clicker was presented), the consequent stimulation is represented as zero on the relevant dimension. Taking the Experiments in Chapters 3 and 4 as an example, in order to permit generalisation from trials with zero squares to trials with 5 or 20 squares, it was further necessary to assume that zero is represented by an element that is also activated when one or more squares are portrayed on the screen. The same logic must necessarily be applied to the dimension of auditory intensity. Apart from being able to explain the present results, it is hard to think of any additional justification for these assumptions. In view of the slender

support for them, therefore, it might be prudent not to abandon the search for an alternative to the foregoing explanations for these findings.

Implications for discriminations based on temporal duration

In a series of experiments conducted by Bouton and colleagues (e.g. Bouton & García-Gutíerrez, 2006: Bouton & Hendrix, 2011) rats received food after a tone when successive presentations of this stimulus were separated by a long, but not a short interval (long+/short-), or the opposite treatment (short+/long-). Importantly, cues indicating the trial outcome were present throughout the experimental session, and there was thus no effective ITI. Nevertheless, an asymmetry was still observed in which the long+/short- discriminations were acquired more readily than the short+/long- discriminations. To account for this effect Bouton and Hendrix proposed a '*temporal elements hypothesis'* in which long durations are considered to be composed of more hypothetical temporal elements than short durations. A short interval might be represented as A, say, and a longer interval as ABC. The asymmetry, according to Bouton and Hendix, can thus be considered as an example of the *feature-positive* effect. Although Pearce's original (1987) configural theory struggles to account for the asymmetry in temporal duration, especially in the absence of an ITI, the proposed modification to this theory readily predicts the results.

Summary

At the beginning of this thesis I argued that a number of theories of conditioning are based on the assumption that generalisation gradients between stimuli are symmetrical (e.g. Blough, 1975; Spence, 1936). From this assumption,

the prediction follows that the rate at which a discrimination is acquired should be unaffected by which of two stimuli is S+. In contrast to this prediction, when S+ and S- are different in terms of their relative position on a magnitude continuum this symmetry does not always hold. In many cases, discriminations with a high-magnitude S+ and low-magnitude S- are acquired more readily than when S+ is low-magnitude and S- is high-magnitude. The primary aim of this thesis was therefore to offer an account for this asymmetry. I found evidence for such an asymmetry with clickers differing in intensity, and with arrays differing in the number of black squares on a white background, although it emerged that the mechanisms responsible for this effect were different for these two dimensions. With auditory intensity, the asymmetry was unaffected by cues presented during the ITI, and indeed remained even if the ITI was removed. In contrast, with quantity of dots, removal of the ITI eliminated the asymmetry and changing the properties of this ITI enabled the asymmetry to be reversed. I then offered an explanation for why these modalities present contrasting results in terms of how the stimuli are represented at the receptor level. With clickers, an increase in intensity results in an increased number of receptors being activated, (i.e. a quantitative difference) including all those that had been activated for weaker stimuli, and the activations of the receptors also increases. With the visual arrays, an increase in results in a change in activation that is similar to that seen when there is a change in the qualitative properties of a stimulus such as a change from a red to a green light. In an attempt to account for both sets of data, I proposed that an adaptation to Pearce´s (1987) theory should be made whereby similarity between stimuli is considered as the proportion of common elements, rather than the number of common elements. This theory appears, at least in terms of the cursory first analysis presented above, to be superior to that of Pearce, and,

importantly, also to the Rescorla-Wagner (1972) model, which is also able to account for the results presented in Chapters 2 to 4. Of course, whether or not this new model has any longevity has yet to be fully assessed, and I am certain that there are many predictions that follow from this model that require testing. However, for now, this new model should stand as starting block from which to further our understanding of the effects of stimulus magnitude on learning, and as a reminder that our abstract conceptualisations of how stimuli are represented should have some grounding at the receptor level.

References

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2009). Use of number by fish. *PloS one, 4,* e4786.

Astley, S.L., Aird, T., & Bouton, M.E. (2015). Asymmetries in time-based and feature-based discriminations in humans: linking the long+ and feature-positive effects. *Journal of Experimental Psychology: Animal Learning and Cognition, 41,* 193-205.

Aydin, A., & Pearce, J.M. (1994). Prototype effects in categorization by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 20,* 264-277.

Aydin, A., & Pearce, J. M. (1995). Summation in autoshaping with short- and long-duration stimuli. *Quarterly Journal of Experimental Psychology, 48B,* 215-234.

Aydin, A., & Pearce, J. M. (1997). Some determinants of response summation. *Animal Learning & Behavior, 25,* 108-121.

Blough, D.S. (1975). Steady state data and a quantitative model of operant generalization and discrimination. *Journal of Experimental Psychology: Animal Behavior Processes, 104,* 3-21.

Bouton, M.E., Doyle-Burr C, & Vurbic, D. (2012). Asymmetrical generalization of conditioning and extinction from compound to element and element to compound. *Journal of Experimental Psychology: Animal Behavior Processes, 38,* 381-393.

Bouton, M.E., & García-Gutiérrez, A. (2006). Intertrial interval as a contextual stimulus. *Behavioural Processes, 71*, 307-317.

Bouton, M.E., & Hendrix, M.C. (2011). Intertrial interval as a contextual stimulus: Further analysis of a novel asymmetry in temporal discrimination learning. *Journal of Experimental Psychology: Animal Behavior Processes, 37*, 79-93.

Bouton, M. E., & Sunsay, C. (2003). Importance of trials versus accumulating time across trials in partially-reinforced appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 29*, 62-77.

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*, 67-72.

Dumont, J.R., Jones, P.M., Pearce, J.M., & Kosaki, Y. (2015). Evidence for concrete but not abstract representation of length during spatial learning in rats. *Journal of Experimental Psychology: Animal Learning and Cognition, 41*, 91-104.

Emmerton, J. (1998). Numerosity differences and effects of stimulus density on pigeons' discrimination performance. *Animal Learning & Behavior, 26*, 243-256.

Emmerton, J., Lohmann, A., & Niemann, J. (1997). Pigeons' serial ordering of numerosity with visual arrays. *Animal Learning & Behavior, 25*, 234-244.

Ernst, A.J., Engberg, L., & Thomas, D.R. (1971). On the form of stimulus generalization curves for visual intensity. *Journal of the Experimental Analysis of Behavior, 16*, 177-180.

Friedman, M.P., & Gelfand, H. (1964). Transfer effects in discrimination learning. *Journal of Mathematical Psychology, 1*, 204-214.

Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour, 66*, 15-36.

Gray, J.A. (1965). Stimulus Intensity Dynamism. *Psychological Bulletin, 63*, 180-196.

Grice, R.G., & Saltz, E. (1950). The generalization of an instrumental response to stimuli varying in the size dimension. *Journal of Experimental Psychology, 40*, 702-708.

Gulliksen, H., & Wolfle, D.L. (1938). A theory of learning and transfer: I. *Psychometrika, 3*, 127-149.

Guttman, N., & Kalish, H.I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology, 51*, 79-88.

Hanson, H.M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology, 58*, 321-333.

Hearst, E. (1978). Stimulus relationships and feature selection in learning and behaviour. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), Cognitive processes in animal behaviour. Hillsdale, N.J.: Erlbaum.

Honig, W.K., & Urcuioli, P.J. (1981). The legacy of Guttman and Kalish (1956): 25 years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior, 26,* 405-445.

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

Hull, C.L. (1952). A behaviour system. New Haven: Yale University Press.

Jakubowska, E., & Zielinski, K. (1976). Differentiation learning as a function of stimulus intensity and previous experience with the CS+. *Acta Neurobiologiae Experimentalis, 36,* 427-446.

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

Jones, P. M., & Pearce, J.M. (2014). The fate of redundant cues: Further analysis of the redundancy effect. *Learning & Behavior, 13,* 72-82.

Kamin, L.J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell, & R. M. Church (Eds.), Punishment and Aversive Behaviour (pp. 279-296). New York: Appleton-Century-Crofts.

Kehoe, E. J. (1982). Overshadowing and summation in compound stimulus conditioning of the rabbit's nictitating membrane response. *Journal of Experimental Psychology: Animal Behavior Processes, 8,* 313- 328.

Kehoe, E.G., & Boesenberg, K.G. (2002). Temporal discriminations using different feature-target intervals in classical conditioning of the rabbit's nictitating membrane response. *Animal Learning & Behavior, 30,* 208-216.

Kehoe, E.G., & Gormezano, I. (1980). Configuration and combination laws in conditioning with compound stimuli. *Psychological Bulletin, 87,* 351-378.

Kosaki, Y., Jones, P.M., & Pearce, J.M. (2013). Asymmetry in the discrimination of length during spatial learning. *Journal of Experimental Psychology: Animal Behavior Processes, 39,* 342-356.

Kosaki, Y., & Pearce, J.M. (2015). Asymmetrical generalization of length in the rat. *Journal of Experimental Psychology: Animal Learning and Cognition, 41,* 266-76.

Kyd, R.J., Pearce, J.M., Haselgrove, M., Amin, E., & Aggleton, J.P. (2007). The effects of hippocampal system lesions on a novel temporal discrimination task for rats. *Behavioural Brain Research, 187,* 159-171.

Logan, F.A. (1954). A note on stimulus intensity dynamism (V). *Psychological Review, 61,* 77-80.

Mackintosh, N. (1974). The Psychology of Animal Learning. London: Academic Press.

McLaren, I.P.L., & Mackintosh, N.J. (2002). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning & Behavior, 28,* 211-246.

Minini, L., & Jeffery, K.J. (2006). Do rats use shape to solve “shape discriminations”? *Learning & Memory, 13,* 287-297.

Moore, B.R. (1972). Stimulus control: studies of auditory generalization in rabbits. In A. H. Black & W. F. Prokasy (Eds.), Classical conditioning II: Current research and theory (pp. 206-230). New York: Appleton-Century-Crofts.

Nieder, A., Freedman, D.J., & Miller, E.K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science, 297,* 1708-1711.

Nieder, A., & Miller, E.K. (2004). A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 7457-7462.

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

Pearce, J.M. (1987). A model of stimulus generalization for Pavlovian conditioning. *Psychological Review, 94,* 61-73.

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

Pearce, J.M. (2002). Evaluation and development of a connectionist theory of configural learning. *Animal Learning & Behavior, 30,* 73-95.

Pelz, C., Gerber, B., & Menzel, R. (1997). Odorant intensity as a determinant for olfactory conditioning in honeybees: roles in discrimination, overshadowing and memory consolidation. *The Journal of Experimental Biology, 200,* 837-847.

Perkins, C.C., Jr. (1953). The relation between conditioned stimulus intensity and response strength. *Journal of Experimental Psychology, 46,* 225-231.

Pierrel, J., Gilmour Sheman, J., Blue, S., & Hegge, F.W. (1970). Auditory discrimination: A three-variable analysis of intensity effects. *Journal of the Experimental Analysis of Behavior, 13,* 17-35.

Pierrel, R., & Sherman, J.G. (1960). Generalization of auditory intensity following discrimination training. *Journal of the Experimental Analysis of Behavior, 3,* 313-322.

Razran, G. (1957). The dominance-contiguity theory of the acquisition of classical conditioning. *Psychological Bulletin, 54,* 1-46.

Rescorla, R. A., & Coldwell, S. E. (1995). Summation in autoshaping. *Animal Learning & Behavior*, 23, 314-326.

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

Scavio, M.J. Jr., & Gormezano, I. (1974). CS intensity effects on rabbit nictitating membrane conditioning, extinction and generalization. *The Pavlovian Journal of Biological Science, 9,* 25-34.

Shephard, R.N. (1987). Toward a universal law of generalization for psychological science. *Science, 237,* 1317-1323.

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

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

Spence, K.W. (1952). The nature of response in discrimination learning. *Psychological Review, 59*, 89-93

Sutherland, N.S., & Williams, C. (1969). Discrimination of checkerboard patterns by rats. *Quarterly Journal of Experimental Psychology, 21*, 77-84.

Todd, T.P., Winterbauer, N.E., & Bouton, M.E. (2010). Interstimulus interval as a discriminative cue: Evidence of the generality of a novel asymmetry in temporal discrimination learning. *Behavioural Processes, 84*, 412-420.

Vonk, J., & Beran, M.J. (2012). Bears "count" too: Quantity estimation and comparison in black bears (Ursus Americanus*). Animal Behaviour, 84*, 231-238.

Wagner, A.R., & Rescorla, R.A. (1972). Inhibition in Pavlovian conditioning: application of a theory. In R. A. Boakes & M. S. Halliday (Eds.), Inhibition and Learning (pp. 301-336). London: Academic Press.

Watanabe, S. (1998). Discrimination of "four" and "two" by pigeons. *Psychological Record, 48*, 383-391.

Whitlow, J. W., Jr., & Wagner, A. R. (1972). Negative patterning in classical conditioning: Summation of response tendencies to isolable and configural components. *Psychonomic Science, 27*, 299-301.