Inman, Richard A. and Pearce, John M. 2018. The discrimination of magnitude: A review and theoretical analysis. Neurobiology of Learning and Memory 153 (PB) , pp. 118-130. 10.1016/j.nlm.2018.03.020

Publishers page: http://dx.doi.org/10.1016/j.nlm.2018.03.020

The Discrimination of Magnitude: A Review and Theoretical Analysis

Richard A. Inman[a], John M. Pearce[b]

[a] Centro de Investigação em Psicologia para o Desenvolvimento, Universidade Lusíada do Porto, Porto, Portugal

[b] School of Psychology, Cardiff University, Cardiff, UK and School of Psychology, University of Sydney, Australia

Short Title: Discrimination of Magnitude

Address for Correspondence:
Richard A. Inman
Centro de Investigação em Psicologia para o Desenvolvimento
Universidade Lusíada
Rua Dr. Lopo de Carvalho
4369-006 Porto
Portugal
inmanr@por.ulusiada.pt

Abstract

In a discrimination based on magnitude, the same stimulus is presented at two different magnitudes and an outcome, such as food, is signalled by one magnitude but not the other. The review presented in the first part of the article shows that, in general, such a discrimination is acquired more readily when the outcome is signalled by the larger rather than the smaller of the two magnitudes. This asymmetry is observed with magnitudes based on sound, odour, temporal duration, quantity, and physical length. The second part of the article, explores the implications of this pattern of results for the theory of discrimination learning presented by Pearce (1994). The asymmetry found with discriminations based on magnitude contradicts predictions derived from the original version of the theory, but it can be explained by a modified version. The asymmetry also has important implications for understanding how animals represent magnitudes.

Keywords: Magnitude Discrimination, Intensity, Quantity, Duration

The ease of solving a discrimination, where one cue, S+, but not another, S-, signals a biologically significant outcome, has long been said to be determined by their similarity. The greater the similarity between the two cues, the more difficult will be the discrimination between them (e.g. Mackintosh, 1974; Pearce, 1994; Shepard, 1987; Spence, 1936). Furthermore, the similarity between two cues is often seen as being symmetrical, with the similarity of cue A to cue B, being the same as the similarity of cue B to cue A. From these proposals it then follows that a discrimination between any pair of cues should progress at the same rate, regardless of which of them signals the outcome. Given this fundamental prediction, the purpose of the present article is twofold. First, a set of experiments will be reviewed whose findings do not conform to the pattern just predicted. These experiments involve discriminations where S+ and S- are the same stimulus presented at two different magnitudes[1]. The experiments show that the discrimination is acquired more readily when the cue of greater magnitude serves as S+ than when the cue of lesser magnitude serves as S+. This outcome is important, not only because it challenges a widely held view about the role played by similarity in discrimination learning, but also because it is hard to explain with a theory that was developed specifically to account for the relationship between similarity and the ease with which a discrimination is solved (Pearce, 1994). The second purpose of the article, therefore, is to explore in some detail how the theory of Pearce might be modified to explain the asymmetry in the discrimination of magnitudes.

Experimental investigations of discriminations based on stimulus magnitude

---

[1] Magnitude will be defined as the position of a stimulus along a dimension (e.g. loudness), where greater distance from ‘zero’ on that dimension corresponds to greater magnitude. Intensity, which will also be referred to, especially in relation to auditory stimuli, relates to the perception of magnitude. With some dimensions, such as loudness, there is a direct relationship between the magnitude of a stimulus and its intensity. However, with other magnitudes, such as the size of a rectangle, this relationship is less obvious.

Despite being relatively small in number, demonstrations that animals are able to discriminate between different magnitudes of the same stimulus have used a wide variety of stimuli and species. The overall conclusion to be drawn from these experiments is that the asymmetry in the ease with which such discriminations are acquired is both robust and of considerable generality.

Auditory Intensity

Experiments described by Zielinski and Jakubowska (1977; see also Jakubowska & Zielinski, 1976) revealed an asymmetry in magnitude discriminations when two groups of rats were trained to differentiate between a 70-dB and a 50-dB white noise, using conditioned suppression of lever pressing for food. The duration of each trial, of which there were 8 in each 2-hr session, was 3 min. In the intervals between trials no white noise was presented. For the 70+/50- group, each presentation of the high-intensity white noise was followed by a foot shock, whereas the low-intensity white noise was never paired with shock. These contingencies were reversed for the 50+/70- group. Both groups came to suppress the rate of lever pressing during the presentation of the stimulus associated with shock, whilst maintaining a high rate of lever pressing during the stimulus not associated with shock. However, the difference between the effects of S+ and S- developed more rapidly for the 70+/50- than the 50+/70- group, thus revealing an asymmetry in the acquisition of the discrimination between the groups.

Pierrel, Sherman, Blue and Hegge (1970) presented groups of rats with an instrumental appetitive discrimination involving a pulsed tone that was of either high or low intensity. The two intensities were presented in an alternating sequence, without an intertrial interval (ITI), with each presentation lasting 1, 2, or 3 min. Lever presses during only one of the two intensities resulted in food, which was delivered according to a VI 2-min schedule.

The training just described continued for 8 hours, followed by a 4-hour period when the rats were left in the chambers with no presentations of the stimuli and the lever was withdrawn. At this point the entire 12-hr cycle was repeated until 15 days of training had been completed, with the consequence that every rat remained continuously in the same chamber for 15 days. The experiment contained seven pairs of groups. The two groups for each pair were trained with the same two stimuli, but for one group food was signalled by the louder of the two stimuli, while for the other group food was signalled by the quieter stimulus. The pairs of groups differed in both the size of the difference between the two intensities, which ranged from 10 to 40 dB, and the position on the intensity scale from where the values were selected, which varied from 60 to 100 dB. In general, the discrimination was acquired more readily when S+ was louder than S- for each pair of groups, although this difference was not evident when S+ and S- were relatively quiet. In all cases the difference between the members of each pair disappeared with extended training. A similar asymmetry has been reported in experiments using similar training to that just described, but with more conventional periods of exposure to the training apparatus (Blue, 1967; Sadowski, 1966). In both cases presentations of S+ and S- alternated throughout every session and there was thus no effective ITI.

For theoretical reasons that will be made evident later, it is of interest to know whether the asymmetry observed with magnitude discriminations depends upon there being periods of nonreinforced exposure to the experimental context in the absence of S+ and S-. Some of the experiments above indicate that an asymmetry in the acquisition of a discrimination based on auditory intensity can be found when there is a period of nonreinforced exposure to the apparatus between successive trials (Zielinski & Jakubowska, 1977), or between successive blocks of trials (Pierrel et al., 1970). On the other hand, the experiments by Sadowski (1966) and Blue (1967) indicate that a similar outcome can be found when there is no exposure to the apparatus in the absence of the experimental stimuli. The conclusion pointed to by these studies is that the asymmetry, at least as far as discriminations based on the intensity of an auditory cue are concerned, dioes not depend upon the presence of an ITI.

Additional support for the foregoing conclusion can be found in an experiment by Inman, Honey and Pearce (2016, Experiment 3). Rats received Pavlovian conditioning with a clicker that was either 57 dB or 82 dB, and which was presented for 73 s at a time in an alternating sequence of the two intensities. Food was delivered into a magazine occasionally during one intensity of the clicker, but not the other. As training progressed, the amount of magazine activity during the intensity that signalled food, S+, became greater than during the intensity that signalled the absence of food, S-, and this occurred more rapidly when S+ was louder than S-, than when S+ was softer than S-. This difference was observed with and without an interval of 73 s between each trial, during which neither the clicker nor food was presented.

The clear conclusion to be drawn from the above experiments is that a discrimination between two intensities of the same auditory cue will be acquired more readily when the reinforcer is signalled by the more intense of the two cues. Moreover, this outcome can be found with, or without an interval between successive trials. Normally, when an ITI is present the auditory cue used for S+ and S- is absent. Experiment 4 by Inman, Honey, and Pearce (2016) has shown that the asymmetry can be found even when the auditory cue is present during the ITI. In this experiment, the experimental cues consisted of a clicker with an intensity of 49 dB and 64 dB. They were each presented for 15 s and sucrose solution was presented at the end of S+, but not S-. The trials were separated by a mean ITI of 6 min, throughout which the clicker was presented at an intensity of 82 dB. Despite the changes to the methodology, the discrimination was acquired more readily by a group for which the intensity of S+ was 64 dB and the intensity of S- was 48 dB, than by a group given the opposite arrangement.

Finally, in this section we can note that an asymmetry in discriminations based on two different intensities of a tone has been found with humans using eye-blink conditioning, with a puff of air to the eye as the unconditioned stimulus (Moore, 1964).

Odour Intensity

The asymmetry found with a discrimination based on sound intensity can also be found with stimuli that differ in intensity from at least one other modality. 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 S+ than S-. 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.

Quantity

If the asymmetry described above is a general characteristic of discriminations between stimuli that differ in magnitude, then it should be observed with magnitudes other than those based on intensity. One obvious magnitude is that of quantity. A discrimination could consist of the occurrence of an important outcome being signalled by a certain number of items in an array, but not by a different number of items. If the asymmetry that has just been described has any generality, then the discrimination will be easier to solve when a large number of items signals the outcome, and a small number its absence, than when the opposite is true. To our knowledge, there are only two reports of experiments that were designed to test this prediction. Before considering these studies, it is worth mentioning two experiments whose findings are relevant to the present discussion, even though they were not designed with the purpose of seeking an asymmetry in the acquisition of a discrimination based on quantity. Vonk and Beran (2012), testing bears, and Watanabe (1998), testing pigeons, found that a discrimination based on quantity is acquired more readily when S+ is of larger magnitude than S-, than when S- is larger than S+. By way of example, the bears demonstrated this asymmetry when required to discriminate between two different amounts of circles presented on a screen, even when total area occupied by circles was unrelated to quantity. However, the number of subjects receiving each condition were so small, one or two in each study, that it is impossible to draw any strong conclusions from them.

Inman, Honey and Pearce (2015) gave two groups of pigeons autoshaping using patterns on a monitor that comprised either 20 or 5 black squares that were randomly distributed in a region in the centre of a white background. Stimuli were presented for 10 s and presentations were separated by a mean ITI of 1 min, when the screen was entirely white. Food was signalled by 20, but not 5 squares for the 20+/5- group while this relationship was reversed for the 5+/20- group. As training progressed, the rate of pecking at the monitor during S+ increased, while during S- the rate decreased. The rate at which this discrimination developed, however, was significantly faster for the 20+/5- than the 5+/20- group. A similar pattern of results has been found with rats (Inman, Honey, Eccles, & Pearce, 2016).

It is tempting to conclude from these results that the asymmetry found with discriminations based on intensity can also be found with discriminations based on quantity. However, it is possible that responding in the experiments just described was not controlled by differences in the number of squares on the screen, but by differences in the amount, or intensity, of blackness on the screen. Patterns with either 5 or 20 black squares on a white background not only differ in the quantity of squares that are displayed but also in their overall level of darkness. In order to determine which of these dimensions pigeons relied on to solve the discriminations, Inman et al. (2015) trained pigeons with a 20+/5- and a 5+/20- discrimination with black squares on a white background. Once both discriminations had been mastered, the cues were changed to being white squares on a black background. If the birds relied on differences in quantity for their original discrimination, then there should be considerable positive transfer to the new task. In contrast, if the solution to the original discrimination was based on differences in the amount of blackness on the screen, then the change to the stimuli should disrupt performance considerably. In fact, the discrimination was hardly affected by the change, which led Inman et al. (2015) to conclude that the pigeons relied on the number of squares on display to solve the magnitude discrimination they were given. This conclusion has been echoed by others, such as Emmerton, Lohmann and Niemann (1997) who have shown that pigeons were able to discriminate between stimuli differing in quantity when overall brightness, summed contour in outline elements, and total area of elements was held constant for S+ and S-. It should be noted that these experiments did not investigate whether there is an asymmetry in discriminations based on quantity.

It thus appears that the asymmetry found with discriminations based on the magnitude of stimulus intensity can also be found with the magnitude of quantity. There are, however, two findings that indicate there may be a difference in the way in which discriminations based on intensity and quantity are solved. First, recall that Inman, Honey and Pearce (2016) found that the presence of a loud, 82-dB, clicker during the ITI did not prevent a discrimination between a 48- and 64-dB clicker being easier to solve when the loud rather than the quiet cue signalled food. A related experiment by Inman, Honey and Pearce (2015) revealed a different outcome with a discrimination based on quantity. The design was similar to the experiment with pigeons described above, except that the screen during the ITI displayed 288 squares, rather than being white. This manipulation is similar to that of increasing the intensity of the clicker during the ITI by Inman, Honey and Pearce (2016) but, on this occasion, it was found to have a marked effect by reversing the asymmetry. The performance of the group receiving the 5+/20- discrimination was superior to the group receiving the 20+/5- discrimination.

The second finding that reveals a difference between the ways in which discrimination based on loudness and quantity are solved, concerns the effect of removing the ITI between successive training trials. As mentioned earlier, the asymmetry found with discriminations based on loudness can be found when there is no interval between successive presentations of the experimental stimuli (Blue, 1966; Inman et al., 2016; Sadowski, 1966). On the other hand, Inman, Honey, Eccles and Pearce (2016), found with rats that the asymmetry observed with a discrimination based on two different numbers of black squares presented on a white screen was eliminated when the interval between successive trials was removed. The implications of both sets of contrasting results are explored in a subsequent section.

Temporal duration

The duration for which a stimulus is presented can be regarded as a magnitude – a tone that lasts for 20 sec can be said to be of greater magnitude than a tone that lasts for 10 sec. In keeping with the results already described, experiments have revealed an asymmetry when two different durations of the same cue signal different outcomes. Kyd, Pearce, Haselgrove, Amin and Aggleton (2007), examined if rats could discriminate between two different durations of a tone. For two groups these durations were 12 sec and 3 sec. Each presentation of the tone was followed by a 10-sec trace interval which, after S+ but not S-, was followed by a delivery of food into a magazine. When the long stimulus served as S+, and the short as S-, then the amount of magazine activity during the trace interval after the long S+ was considerably greater than after the short S-; but when the short duration stimulus served as S+ and the long duration stimulus as S-, this difference was diminished substantially and developed relatively slowly. A similar asymmetry was found when the duration of the tone was 0.5 and 1.5 sec.

Todd, Winterbauer, and Bouton (2010, Experiment 1) report a similar asymmetry when presentations of white noise for either 70 or 250 sec were accompanied by a tone for the final 10 sec of each trial. The tone was followed by food when it was accompanied by the long noise, but not the short noise, for the long+/short- group, whereas for the short+/long- group the tone was followed by food when it was accompanied by the short but not the long noise. Both groups displayed an increase in magazine activity during the tone as training progressed, but only the long+/short- group revealed a mastery of the discrimination by responding significantly more frequently on reinforced than nonreinforced trials.

A discrimination between a long and a short cue is not always acquired more readily when the outcome is signalled by the former rather than the latter. In fact, an experiment by Todd et al. (2010, Experiment 2) revealed the opposite outcome. The design of the experiment was much the same as for their study above, but the noise lasted for either 15 sec or 60 sec before the tone was presented. In contrast to the above results, the short+/long- discrimination was acquired more readily than long+/short-. This effect was also observed when responding was recorded during the 10 sec before each presentation of the tone, and during the tone itself. Todd et al. (2010) explained this anomalous outcome by referring to the observation that a short interstimulus interval, ISI, supports stronger conditioned responding than a long ISI. Thus the unusual pattern of results was said to occur because there was more scope for stronger responding during the noise in the short+/long- than the long+/short- condition. This argument was tested by repeating the experiment (Todd et al., 2010, Experiment 3), but with the addition of trials where the noise was presented without the tone or food in order to reduce the amount of responding elicited shortly after its onset. The manipulation was successful and resulted in the more conventional asymmetry of the long+/short- discrimination with the tone being mastered more readily than short+/long-.

The anomalous asymmetry reported by Todd et al. (2010, Experiment 2) is similar to an effect described by Kehoe and Boesenberg (2002). Rabbits received Pavlovian nictitating membrane conditioning with a 400-msec light embedded in a 66-sec tone. The light was presented either 5 or 45 sec after the onset of the tone. For Group 5+/45-, shock was presented after the light when it occurred 5 sec, but not 45 sec, after tone onset, while for Group 45+/5- shock was presented when the light occurred 45 sec but not 5 sec after tone onset. After 11 days of training the probability of a CR was more likely on reinforced than nonreinforced trials, but this discrimination was acquired more readily by Group 5+/45- than Group 45+/5-. As noted by Todd et al. (2010), this outcome may well have been a consequence of the same factors that they proposed were responsible for their similar finding in Experiment 2.

For all of the experiments described in this section, there was an interval between successive training trials, and this interval provided no information that was relevant to the solution of the discrimination. Such a relationship was not true for a series of experiments by Bouton and his colleagues where the length of the ITI was the cue for a temporal discrimination (Bouton & García-Gutiérrez, 2006; Bouton & Hendrix, 2011). Despite this change, an asymmetry was again observed. Rats received repeated presentations of a 10-s tone separated by intervals of either 4 or 16 min in one experiment, or separated by an intervals of 1 or 4 min in another experiment. In both experiments, when the tone was followed by food after the longer but not the shorter ITI, then there was considerably more magazine activity during the tone on reinforced trials than the nonreinforced trials. This discrimination did not develop, however, when food was presented after the shorter rather than the longer of the two ITIs. A related experiment with humans, but with ITI durations of 1 and 4 sec, yielded a similar outcome (Astley, Aird & Bouton, 2015).

The acquisition of a temporal discrimination thus shows the same asymmetry that has been found with quantity and auditory intensity. Moreover, this asymmetry has been found with a variety of tasks, and with a wide range of durations. For the experiments in which the duration of the ITI provided the temporal cue for solving the discrimination, there was necessarily no interval of nonreinforced exposure to the apparatus between successive trials. This finding is interesting because it implies that magnitude discriminations based on temporal duration have more in common with those based on loudness, where the asymmetry does not depend upon nonreinforced exposure to the apparatus between trials, than those based on quantity, where such exposure is necessary if the asymmetry is to be seen.

Length

The final magnitude to consider is length. If an animal must choose between a long and a short object in order to find a goal then, on the basis of the results considered thus far, it would be expected that the long+/short- discrimination will be easier to solve than short+/long-. This prediction has been confirmed by Kosaki, Jones and Pearce (2013) who trained rats to find one of two hidden platforms in a square swimming pool, the walls of which were grey and 141 cm in length. 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). The hidden platforms were in the middle of the long, but not the short panels (long+/short-) for half of the subjects, while for the remainder they were in the middle of the short but not the long panels (short+/long-). This training was followed by test trials in which the hidden platforms were removed from the pool. 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 also found when the short and long panels were, respectively, 25 or 100 cm in length. In a variation of this design, Kosaki, et al. (2013) required rats to find an escape platform in the middle of a wall in a rectangular swimming pool (180 cm x 90 cm). During test trials with the platforms removed from the pool, the group trained to find the platform in the middle of a long wall spent considerably more time searching near the middle of the long rather than the short walls. In stark contrast, rats trained to find the platform in the middle of a short wall failed to express any preference at all for searching in the middle of the short rather than the long wall. It would seem that the asymmetry brought about with magnitude discriminations extends to the magnitude of length, and that it can severely interfere with the ability to locate a hidden goal in a simple environment.

Theoretical Analysis

Despite the relative lack of systematic investigation, the foregoing review has revealed a strong and diverse body of evidence demonstrating that discriminations based on magnitude are easier to solve when the outcome is signalled by the larger rather than the smaller of the two magnitudes. We now consider two theoretical accounts that were proposed more than fifty years ago as explanations for the results just described. There are shortcomings with each of them. We therefore turn to a more recent theory, which was developed specifically to account for the role of similarity in discrimination learning. Although the original version of the theory is unable to explain the asymmetry found with the discrimination of magnitude, we offer a modification to the theory which enables it to accommodate the results described above.

Stimulus Intensity Dynamism

Hull (1949) proposed that the magnitude of the conditioned response shares a linear relationship with the intensity of the conditioned stimulus. This positive correlation is referred to as stimulus intensity dynamism, and it leads naturally to the prediction that a discrimination between, say, a strong and a weak tone will be acquired more readily when the strong tone is S+, and stimulus intensity dynamism will work in its favour, than when the strong tone is S- and stimulus intensity dynamism will work against the discrimination.

It might thus appear that stimulus intensity dynamism provides an attractively simple explanation for the asymmetry that is found with magnitude discriminations, but there are good reasons for questioning this conclusion. One problem for such an explanation is posed by experiments with magnitudes that do not necessarily reflect the intensity of a cue. Consider, for example, the temporal discrimination experiments by Bouton & García-Gutiérrez (2006) and Bouton & Hendrix (2011), in which the duration of an ITI served as the cue for responding during a tone. If the asymmetry that was observed is to be explained by stimulus intensity dynamism, then it must be assumed that a long ITI provides a greater intensity of stimulation during the subsequent tone than a short ITI. It is hard to find a justification for this assumption that is convincing. A similar problem is posed by the results of Kyd et al. (2007), where responding was recorded during a trace interval of fixed duration after either a long or a short CS.

Turning now to the magnitudes of quantity and length, the same argument applies. It is not easy to see why a long black panel on a white wall is more intense than a short black panel on the same wall (see Kosaki et al., 2013), nor is it easy to see why numerous black squares on a white background is more intense than a few black squares on the same background (see Inman et al., 2015). Both of the above claims would have to be true if stimulus intensity dynamism is to explain the results from these experiments. Furthermore, in the study by Kosaki et al. (2013), the asymmetry was made evident by a stronger preference for one place over another, which does not fit comfortably with Hull's (1949) claim that stimulus intensity is effective by enhancing the vigour of a response.

Finally, in an extremely thorough review devoted to stimulus intensity dynamism, Gray (1965) points out that although there is a good deal of evidence showing that the response to an intense CS is stronger than to a weak CS, this effect is normally found only when a gap separates repeated presentations of the CS. According to Hull's proposals, the the effects of stimuls intensity dynamism will be observe with, or without an ITI. Gray argues that such results are entirely consistent with an explanation for stimulus intensity dynamism put forward by Logan (1954) and Perkins (1953), and which we shall now consider as an account for the results reviewed above.

At first sight stimulus intensity dynamism might appear to offer a plausible account for the asymmetry found with magnitude discriminations. Closer inspection, however, reveals no good reason for regarding it as an explanation for the majority, if not all of the results reviewed above.

Excitatory-inhibitory gradient interaction

In order to explain why conditioning is more effective with a strong rather than a weak stimulus, as an alternative to stimulus intensity dynamism, Logan (1954) and Perkins (1953) both appealed to the influence of the interval between successive conditioning trials (see also Mackintosh, 1974, p. 532-3). They argued that nonreinforced exposure to these intervals will enable the cues present during the ITI to acquire inhibition. This inhibition was then assumed to generalise more readily to the weaker than the stronger of the two CSs, and result in the weak CS eliciting a weaker CR than the strong CS. Clear support for these proposals can be found in Gray's (1965) observation that the effects attributed to stimulus intensity dynamism are generally obtained only when the training trials are separated by an ITI.

It is simple to see how these ideas account for the asymmetry observed with magnitude discriminations, when there is an interval between successive trials (e.g. Moore, 1964). In the case of a large+/small- discrimination there will more generalisation of inhibition from the cues present during the ITI to the small S- than the large S+, and the discrimination will be easy to solve. The opposite will be true for a small+/large- discrimination, where the generalisation of inhibition from the ITI cues will disrupt conditioning with S+, and have little impact on conditioning with S-. According to this account, therefore, the asymmetry depends upon cues being present during the ITI, and on these cues being more similar to S+ than S- for the small+/large- than the large+/small- discrimination.

Although the foregoing analysis can explain many of the results described above, it runs into a problem with experiments where an asymmetry in a magnitude discrimination has been found, and there is no interval between the trials with S+ and S- (Bouton & García-Gutiérrez, 2006; Bouton & Hendrix, 201; Blue, 1966; Inman, Honey & Pearce, 2016; Sadowski, 1966). In these circumstances, there will be no source of inhibition to influence,through generalisation. responding to S+ and S-, and thus no theoretical reason for a large+/small- discrimination to progress more rapidly than small+/large-. As they stand, therefore, the proposals of Logan (1954) and Perkins (1953) offer, at best, an incomplete account for the asymmetry found with discriminations of magnitude.

A configural analysis

In an attempt to find a more satisfactory explanation for the results reviewed above than the two accounts just considered, we shall turn to the theory of Pearce (1994; see also Pearce 1987; Pearce, 2002). One reason for opting for this theory is that the explanation it offers for the present results has its roots in the principles on which the theoretical accounts developed by Logan (1954) and Perkins (1953) were based. That is, the asymmetry is explained by appealing to standard principles of stimulus generalization. In fact, like its predecessors, the theory as it was originally formulated is unable to account for all the results reviewed above but, with a simple modification, it is possible for it to provide a reasonably comprehensive explanation for the asymmetry found with magnitude discriminations. After describing the original theory, we shall then describe and evaluate the modification.

Pearce (1994) proposed that a discrimination between two patterns will result in a representation of each pattern in its entirety entering into a single association with the outcome with which it is paired. In keeping with the accounts just considered, it was further assumed that the strength of the response to one pattern was determined by its own associative strength, together with any associative strength that generalizes to it from the other pattern. These ideas were presented formally in the following way (Pearce, 1994). The amount of generalisation from a training pattern of stimulation to a test pattern, is determined by their similarity, ${}_{Train}S_{Test}$, which is given by Equation 1.

$$ {}_{Train}S_{Test} = (N_C / N_{Train}) \times (N_c / N_{Test}) \qquad (1) $$

According to this equation, the similarity of the two patterns, ${}_{Train}S_{Test}$, is determined by the proportion of elements common to both patterns that are present in the training patern, N_C/N_{Train}, multiplied by the proportion of the same common elements that are present in the test pattern, N_C/N_{Test}. Equation 1 was formulated with the intention of capturing the idea that similarity is symmetrical and in doing so, was consistent with at least one other theory of discrimination learning (Atkinson & Estes, 1963). It thus follows from this equation that the similarity between two patterns will be the same no matter which is the training pattern and which is the test pattern. Equation 2 shows how the overall associative strength of pattern A, V_A, which determines the strength of the response it elicits, is made up of its own associative strength, E_A, together with the associative strength that generalises to it from a similar pattern, $E_{A'}$.

$$V_A = E_A + {}_{A}S_{A'} \times E_{A'} \qquad (2)$$

Finally, Equation 3 shows how the associative strength of A will change on a single training trial.

$$\Delta E_A = \beta \times (\lambda - V_A) \qquad (3)$$

The value of λ is determined by the magnitude of the outcome, and β is a learning rate parameter with a value between 0 and 1. On trials when the US is absent, λ will be equal to 0 which may result in ΔE_A being negative and A might then enter into an inhibitory association with a negative associative strength.

As they stand, it is not possible to apply these proposals to magnitude discriminations because Pearce (1994) did not specify how a weak and a strong stimulus might be represented, which then makes it impossible to determine the similarity between the two stimuli. However, as far as loudness is concerned, there is good evidence that a loud version of a given stimulus not only excites the same neurons as a quieter version, but also excites additional neurons (e.g. Relkin & Doucet, 1997). Thus, it would seem reasonable to characterise a quiet stimulus as A, a moderately loud stimulus as AB, and an even louder stimulus as ABC. A quiet+/loud- discrimination can thus be characterised as A+/AB-, and loud+/quiet- as AB+/A-. While this characterisation means that a discrimination based on loudness, or any other magnitude, is amenable to analysis in terms of configural theory, given the symmetrical relationship of similarity that is embodied in Equation 1, it then follows from the theory that these discriminations will be acquired at the same rate.

How might the theory be modified to account for the asymmetry in magnitude discriminations. One avenue worthy of exploration is suggested by the results from experiments with rabbits (Brandon, Vogel, & Wagner, 2000) and rats (Bouton, Doyle-Burr, & Vurbic, 2012). In these experiments, groups received Pavlovian conditioning with either a single stimulus, A, or a compound, AB, before being tested, respectively, with either AB, or A. Given the symmetrical nature of Equation 1, the theory of Pearce (1994) then predicts, provided conditioning with A and AB has reached the same level, that the response to AB, after training with A, will be of the same strength as the response to A, after training with AB. In other words, the disruptive effect of removing a stimulus from a compound of two CSs is predicted to be the same as the disruptive effect of adding a stimulus to a single CS. In contrast to this prediction, both Brandon et al. (2000) and Bouton et al. (2012) found that the transition from training with AB to testing with A was considerably more disruptive than the transition from training with A and testing with AB. That is, the generalisation decrement induced by removing a cue was greater than when it was induced by the addition of a cue. Given this conclusion, we suggest Equation 1 should be replaced by Equation 4.

$$ {}_{Train}S_{Test} = (N_C \,/\, N_{Train}) \;\text{x}\; (N_C/N_{Test})^y \qquad (4)$$

The equation is the same as Equation 1, except that the expression, (N_C/N_{Test}) , is raised to the power y, where y can have a value ranging from 0 to 1. If y is equal to 1, then Equation 4 is the same as Equation 1. On the other hand, if $y = 0$ then then the expression

(Nc/N_{Test}) will be equal to 1, and the value of ${}_{Train}S_{Test}$ will be given entirely by (N_C / N_{Train}). When $y = 0$, therefore, the degree of generalisation from a training pattern to a test pattern is determined solely by the number of elements that are removed from the training pattern. The greater the number of elements removed, the greater wlll be the generalisation decrement to the test pattern. Adding elements to a training pattern will have no impact at all on the response to the test pattern. Inman, Honey and Pearce (2016) show that it is possible to explain the asymmetry found with magnitude discriminations if ${}_{Train}S_{Test}$ is determined in this manner. It is also possible to predict, in keeping with findings by Brandon et al. (2000) and Bouton et al. (2012), that after training with AB, a test with A will result in a weaker reponse than when training takes place with A and testing with AB. These successes, however, come at a cost. The theory is now unable to explain how certain discriminations are solved, such as negative patterning[2] . The modified theory also incorrectly predicts that after training with A, no generalisation decrement will be seen when a second cue is presented with A, for example, AB. Clear evidence that contradicts this prediction has been reported by Brandon et al, (2000) and can also be found in demonstrations of external inhibition (e.g. Pavlov, 1927).

We propose, therefore, that y is neither 0 nor 1 in Equation 4, but has a value between these limits and, for the purposes of discussion, we shall pursue the implications of its value being set at .5. Suppose training is conducted with AB, and a test trial takes place with A. The number of elements in the training pattern, N_{Train}, and the test pattern, N_{Test}, will be 2 and 1, respectively, while the number of elements common to both patterns, N_C, will be 1. Substituting these values into Equation 4, with y equal to .5, results in ${}_{Train}S_{Test}$ having a

---

[2] We are indebted to Justin Harris for pointing out this shortcoming of allowing the value of ${}_{Train}S_{Test}$ to be determined solely by (N_C / N_{Train}).

value of .5, which is the same as that given by Equation 1. The strength of the response to A is thus predicted to be half that of AB.

Now suppose that after training with stimulus A, a test is given with AB. The value of (N_C / N_{Train}) will be 1, and the value of $(Nc/N_{Test})^y$ will be the square root of ½ (if $y = .5$). Thus according to Equation 4 the value of ${}_{Train}S_{Test}$ will be .7071,which is greater than the value of .5 predicted by this equation when training with AB is followed by a test with A. Equation 4, with $y = .5$ (or any other value between 0 and 1), thus correctly predicts that testing with an element, after training with a compound, will result in a greater decrement in responding than testing with a compound, after training with an element (see Bouton et al., 2012; Brandon et al., 2000). Moreover, provided the value of y is not too close to zero, Equation 4 is also able to predict the successful solution of a negative patterning discrimination; it is also able to explain the solution of other nonlinear discriminations such as biconditional discriminations. Finally, with values of y between 0 ands 1, Equation 4 is able to predict the decrement in responding that is revealed when a novel cue is added to one that has already been used for conditioning, as revealed by demonstrations of external inhibition (e.g. Pavlov, 1927). See also Brandon et al, (2000).

Given these successes, we can now explore the implications of the modification to Pearce's (1994) theory for the discrimination of magnitude. According to Equation 4, with both a large+/small- discrimination, AB+/A-, and a small+/large- discrimination, A+/AB-, there will be generalisation from the reinforced to the nonreinforced pattern. However, because the addition of a cue to a pattern, A+/AB-, results in a weaker generalisation decrement than the removal of a cue, AB+/A-, it then follows that large+/small- will be acquired more readily than small+/large-.

Computer Simulations

In order to determine how well the proposed modification to the theory of Pearce (1994) is able to account for the asymmetry found with magnitude discriminations, a series of computer simulations was conducted. For all the simulations, which were based on Equations 2, 3, and 4, the value of β was .2, the value of λ was 1 on reinforced trials, and 0 on nonreinforced trials and the value of y was .5.

Loudness. One set of simulations explored the predictions made by the amended theory concerning discriminations based on loudness. The first simulation, whose results are shown in Figure 1, determined the predicted course of acquisition of a discrimination between a loud and a quiet stimulus, when the trials were not separated by an ITI. The results for a loud+/quiet- discrimination, which was represented as AB+/A-, can be seen in the upper left-hand panel of Figure 1, while the results for a quiet+/loud- discrimination, A+/AB-, can be seen in the upper right-hand panel. In order to facilitate the comparison of the two conditions, discrimination ratios were calculated using the predicted associative strength for S+, divided by the sum of the predicted associative strengths of S+ and S-. The ratios are presented in the lower panel of Figure 1, which confirms that the loud+/quiet- discrimination is predicted to be acquired more readily than quiet+/loud-. These predictions are in keeping with results reported by Sadowski (1966), Blue (1967), and Inman, Honey, and Pearce (2016, Experiment 3).

FIGURE 1 ABOUT HERE

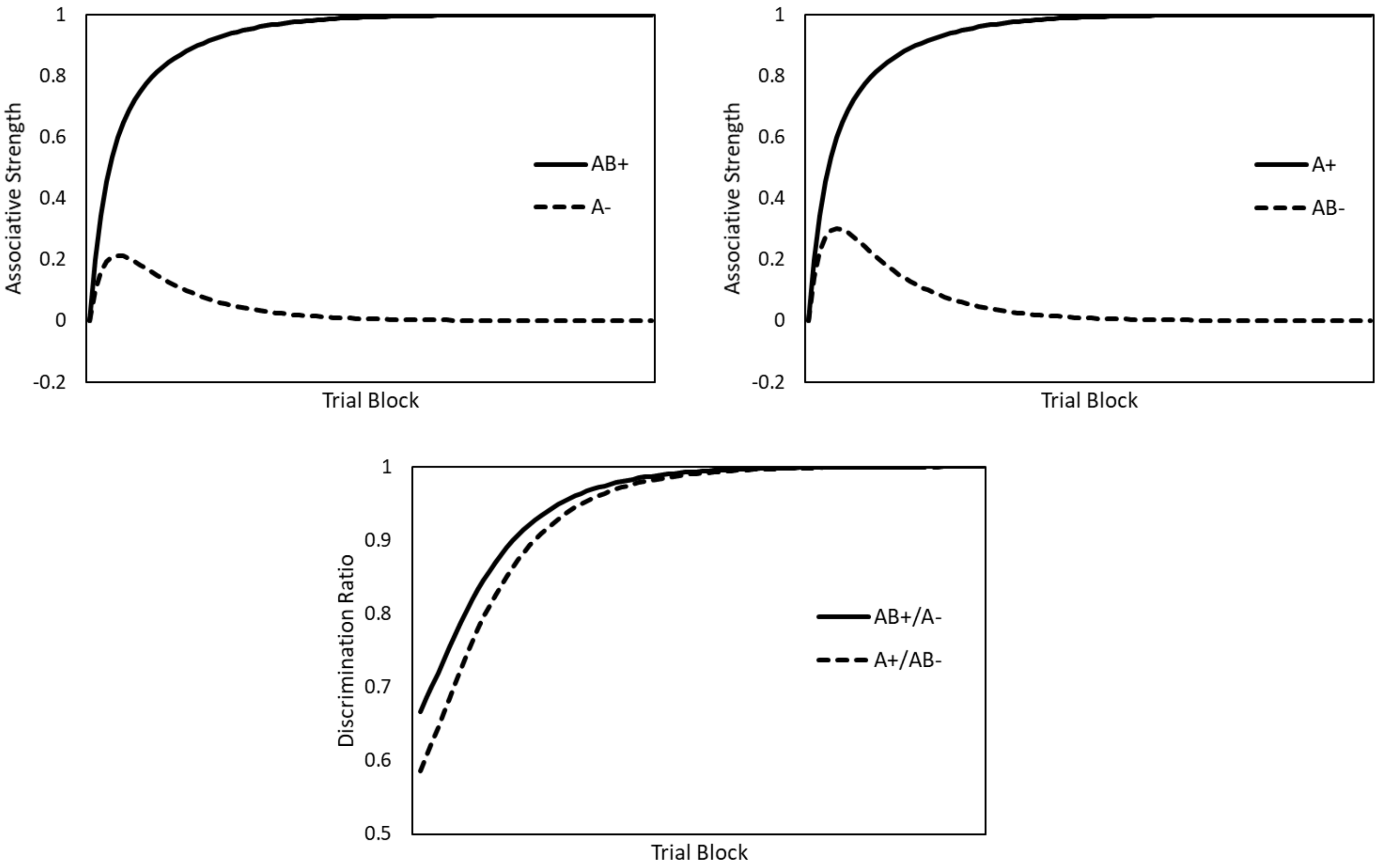

Figure 1. Computer simulation of AB+/A- (upper left-hand panel) and A+/AB- (upper right-hand panel) discriminations based on the modified Pearce equations. The lower panel displays discrimination ratios for both discriminations.

The next simulation examined the predictions for a discrimination with loud and quiet stimuli when there was an ITI devoid of experimental stimuli between successive trials. Evidence that this training will result in an asymmetry in the acquisition of discriminations based on loudness has been provided by, for example, Zielinski and Jakubowska (1977). The same characterisation of the experimental stimuli was used as for the previous simulation, except that a third letter, C, was introduced to represent the contextual cues present throughout each experimental session. Thus, the loud+/quiet- discrimination was characterised as C-/ABC+/AC-, and quiet+/loud- as C-/AC+/ABC-, where C- represents the nonreinforced exposure to the context during the ITI. In keeping with experimental findings, the results in the upper left-hand panel of Figure 2 show that the loud+/quiet- discrimination was acquired more readily than the quiet+/loud- discrimination, whose results are shown in the upper right-hand. The ratios depicted in the lower panel were based on the predicted values for the trials with ABC and AC.

FIGURE 2 ABOUT HERE

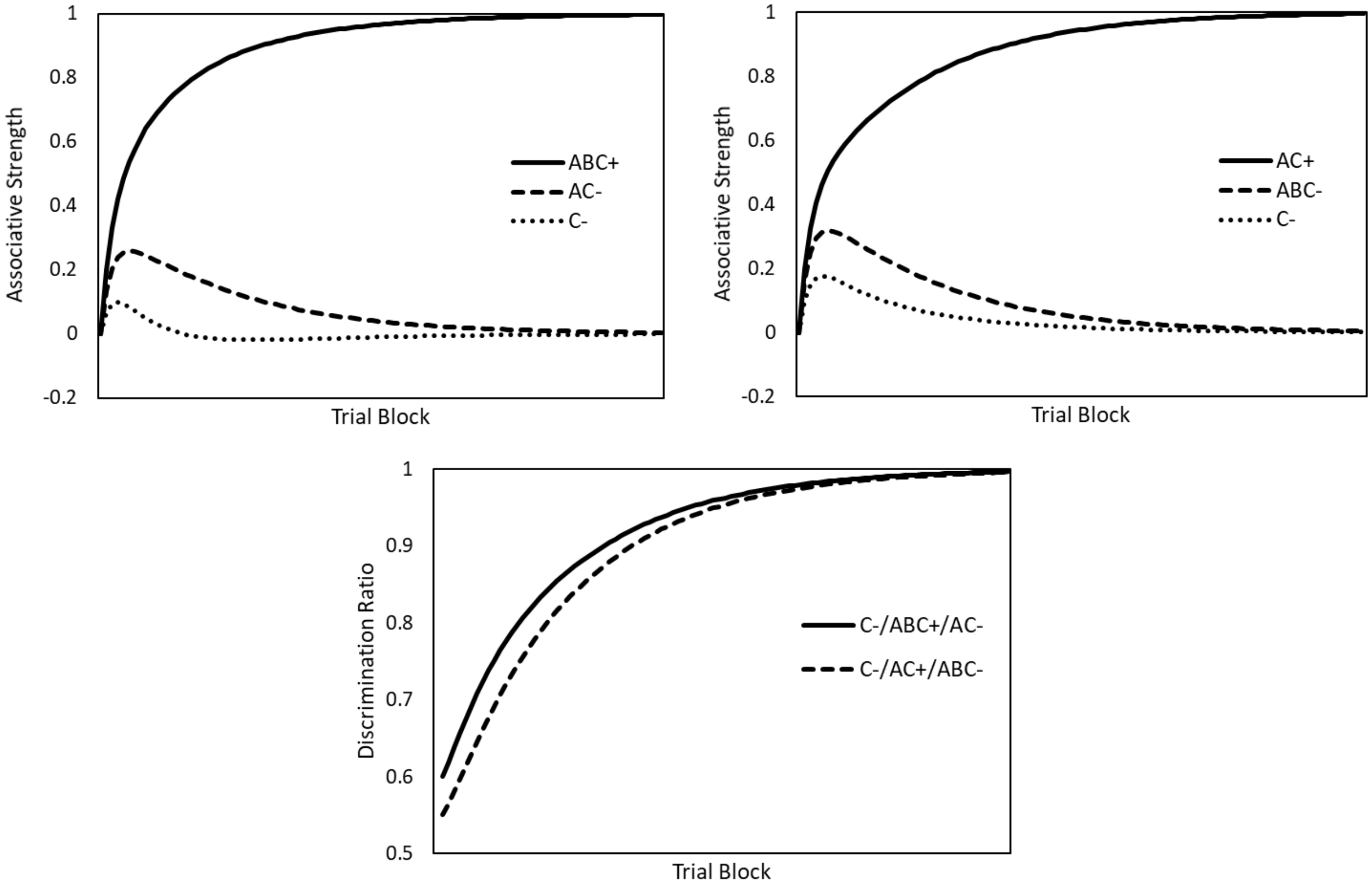

Figure 2. Computer simulation of C-/ABC+/AC- (upper left-hand panel) and C-/AC+/ABC- (upper right-hand panel) discriminations based on the modified Pearce equations. The lower panel displays discrimination ratios for both discriminations.

A further finding reported by Inman, Honey, and Pearce (2016, Experiment 4), is that the asymmetry just described can also be found if a cue that is louder than S+ or S- is played throughout the ITI. In order to discover the prediction made by the modified theory concerning this experiment, a simulation was performed with the quiet CS represented as A, the moderate CS as AB, and the loud ITI as ABC. As can be seen in Figure 3, the modified theory correctly predicts that the moderate+/quiet- discrimination is acquired more readily than the quiet+/moderate- discrimination.

FIGURE 3 ABOUT HERE

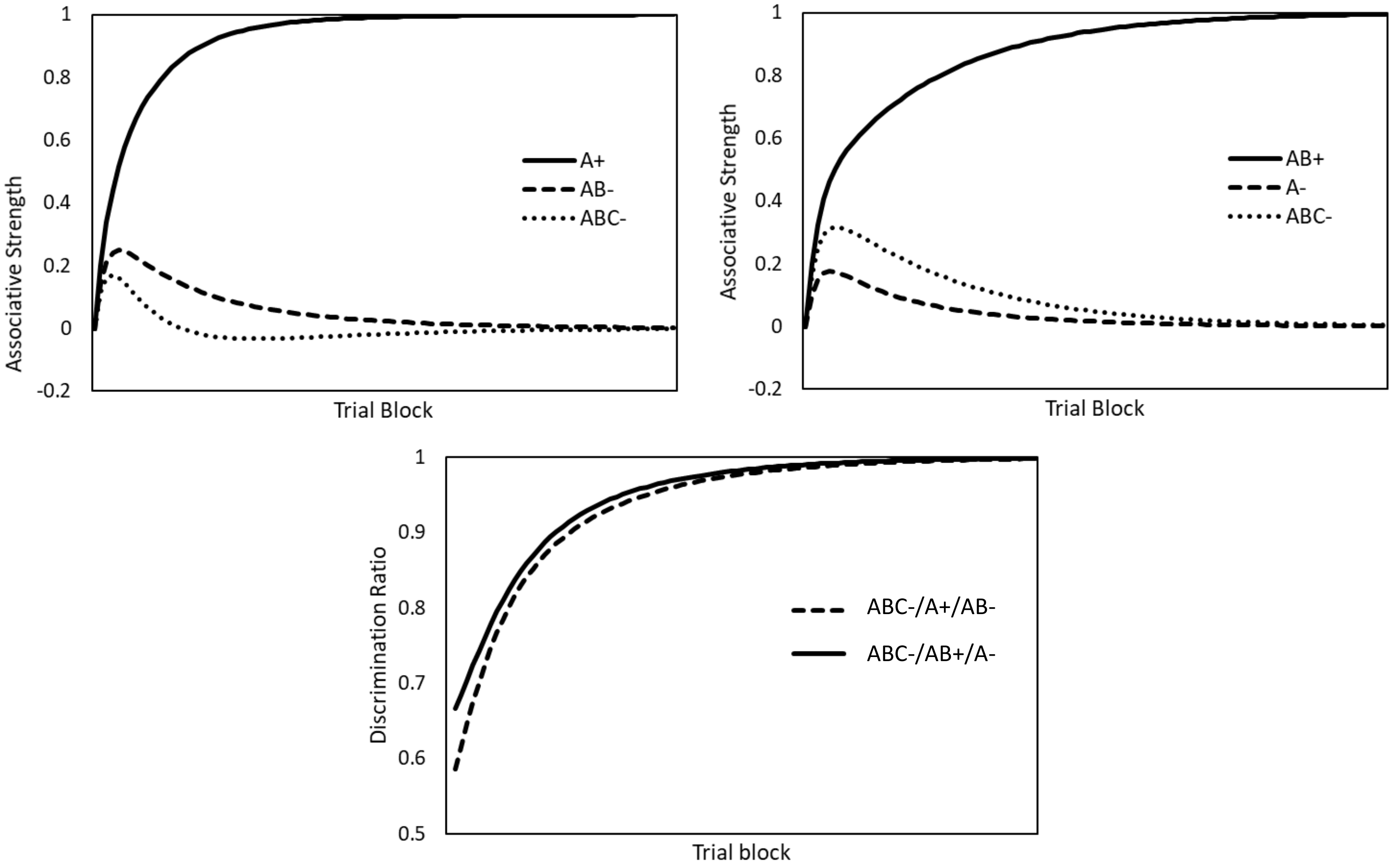

Figure 3. Computer simulation of ABC-/A+/AB- (left hand panel) and ABC-/ABC+/AC- (centre panel) discriminations based on the modified Pearce equations. The right-hand panel displays discrimination ratios for both discriminations.

Substituting Equation 4 for Equation 1 thus enables the configural theory of Pearce (1994) to account for all the results that have been found with discriminations based on the loudness of a stimulus. It was assumed for these simulations that a loud stimulus contained more elements than a quiet stimulus, and that the former contained all the elements present in the latter. A similar assumption was made by Bouton and Hendrix (2011) and Todd et al. (2010), when considering how an animal might represent different durations of the same stimulus. For example, a 20-sec tone might be conceptualised as A, and a 40-sec tone as AB. Once this assumption is made, then it follows that the asymmetry found with discriminations based on temporal duration can be explained in the same way by the modified theory of Pearce (1994) as those based on loudness.

The same conclusion may not apply for the magnitude of quantity. If we assume that a small quantity is represented as A, and a large quantity as AB, then it follows from the foregoing analysis that a discrimination between these quantities will be asymmetrical, irrespective of whether or not the stimuli are separated by an ITI. We have seen that for both rats and pigeons, discriminations based on two quantities are asymmetrical when the trials are separated by an ITI. In an experiment with rats, however, there was no evidence of this asymmetry when the ITI was removed and the trials with the different quantities were presented one immediately after the other (Inman, Honey, Eccles & Pearce, 2016). A further difference between magnitude discriminations based on quantity and loudness, is that presenting an intense version of the stimulus throughout the ITI does not reverse the asymmetry found with the discrimination based on the loud and quiet versions of the same stimulus (Inman, Honey and Pearce, 2016). In contrast, the equivalent manipulation for a quantity discrimination was found to reverse the discrimination based on the two quantities used for the experimental stimuli. That is when 288 squares were presented during the ITI to pigeons, a 20+/5- discrimination was acquired more slowly than 5+/20- (Inman et al., 2015). For a similar result with rats see Inman, Honey, Eccles, and Pearce (2016).

One implication of the different results obtained with discriminations based on quantity and loudness, is that these dimensions are represented in different ways. As noted above, there are grounds for believing that a loud stimulus excites the same elements as a quiet stimulus but also elements that are not excited by the quiet stimulus. The same may not be true as far as quantity is concerned. Neurons have been identified in monkeys (Beran, 2016; Nieder & Miller, 2003; Okuyama, Kuki, & Mushiake, 2015), and in birds (Ditz & Nieder, 2015) that fire maximally to specific quantities of objects, including zero objects, but they will also fire, at progressively decreasing rates, to quantities that are further and further removed from the original quantity. Given this evidence, it is tempting to suggest that a discrimination between two quantities will progress in the same way as between two auditory cues that differ in frequency, but not loudness. The representations of two quantities, therefore, might contain the same number of elements, which are ordered along the dimension of quantity. Some elements may be unique to a given quantity, and others may be shared with adjacent quantities. In a simple form such a proposal leads to a small quantity being represented by ABCD, a moderate quantity as BCDE, and a high quantity as CDEF, where the letter A through F are organised linearly along the dimension of quantity. A small – medium discrimination will then involve the two patterns ABCD and BCDE. Given that both patterns contain the same number of elements, it follows from Equation 4, with y set at .5, that the similarity between them will be symmetrical, with both $_{ABCD}S_{BCDE}$ and $_{BCDE}S_{ABCD}$ having a value of .65 and an asymmetry in the discrimination of magnitude will not be observed with these quantities. Of course, this prediction takes no account of any learning that might take place during the ITI and is thus restricted to procedures without an interval between trials. In support of the foregoing analysis, Inman, Honey, Eccles and Pearce (2016) found that in these circumstances there is, indeed, no asymmetry with discriminations based on quantity.

Once an ITI is added, if it is assumed that the quantity of objects during this interval is zero, then there will be more generalisation from the ITI to the small than the large quantity and, for the reasons put forward by Logan (1954) and Perkins (1953), the large+/small- discrimination will be acquired more readily than small+/large-. Of course, if the number of objects present during the ITI is greater than during the large stimulus, then the opposite result will be found. Both of these predictions have been confirmed (Inman, Honey and Pearce, 2015; Inman, Honey, Eccles and Pearce, 2016).

In order to confirm that the foregoing predictions do indeed follow from the modified theory of Pearce, further computer simulations were conducted. Four quantities were represented in the following way: 0 as ABCDE, small as BCDEF, moderate as CDEFG, and large as DEFGH. The top row in Figure 4 shows the predicted course of acquisition of a small+/ moderate- and a moderate+/small- discrimination, in the absence of an ITI. The middle row shows the same discrimination but with an ITI during which no objects were present, ABCDE. The bottom row of Figure 4 displays the predictions again for the same discriminations, but with a large number of objects present during the ITI. The predictions from each of the simulations are consistent with the experimental findings.

FIGURE 4 ABOUT HERE

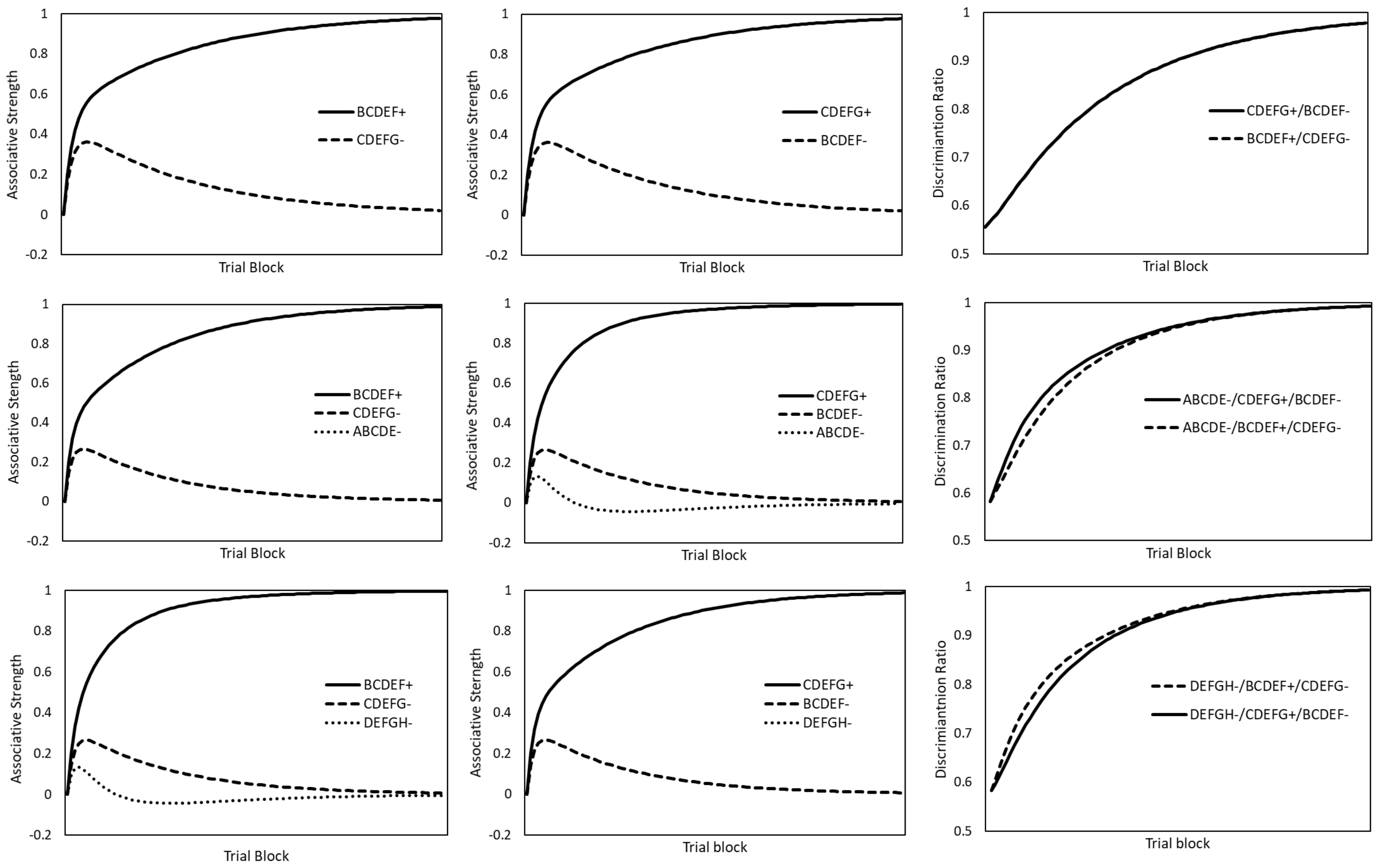
Associative Strength
1
0.8
0.6
0.4
0.2
0
-0.2
Trial Block
BCDEF+
CDEFG-
Associative Strength
Trial Block
CDEFG+
BCDEF-
Discrimiantion Ratio
1
0.9
0.8
0.7
0.6
0.5
Trial Block
CDEFG+/BCDEF-
BCDEF+/CDEFG-
Associative Strength
Trial Block
BCDEF+
CDEFG-
ABCDE-
Associative Strength
Trial Block
CDEFG+
BCDEF-
ABCDE-
Discrimination Ratio
Trial Block
ABCDE-/CDEFG+/BCDEF-
ABCDE-/BCDEF+/CDEFG-
Associative Strength
Trial Block
BCDEF+
CDEFG-
DEFGH-
Associative Sterngth
Trial block
CDEFG+
BCDEF-
DEFGH-
Discrimiantnion Ratio
Trial block
DEFGH-/BCDEF+/CDEFG-
DEFGH-/CDEFG+/BCDEF-

Figure 4. Computer simulations of discriminations involving number stimuli based on the modified Pearce equations. The top row presents BCDEF+/CDEFG- and CDEFG+/BCDEF- discriminations (experiment with no ITI). The middle row presents the same discriminations with zero (ABCDE-) presented during the ITI. The lower row presents the same discriminations with a larger number than S+ and S- (DEFGH-) presented during the ITI.

The modification to the theory of Pearce (1994), has been evaluated thus far by exploring its implications for tasks where variations in magnitude are assumed to alter the number of elements excited by a single stimulus. A different, and more direct, way to explore the merits of the proposed modification, would be to assess how well it accounts for the outcome of experiments in which the number of stimuli, rather than the number of hypothetical elements activated by a single stimulus, is manipulated. On the basis of the available evidence, the proposed amendment to Pearce (1994) is more successful than the original theory.

Summation. There is good evidence that a compound of two stimuli, which have both been paired with the same US for Pavlovian conditioning, will elicit a stronger response than either stimulus alone. Using Pavlovian conditioning, this effect has been found, for example, with rats (Rescorla, 1997), rabbits (Kehoe, 1982; Whitlow & Wagner, 1972) and dogs (Konorski, 1948). It has also been found with instrumental conditioning (e.g. Meltzer & Hamm, 1976). Summation is not easy to explain with the original theory of Pearce (1994) because, according to Equation 1, each stimulus will contribute half of its associative strength to the compound, with the result that the strength of the response to the compound will be the same as to either element alone. In contrast, it follows from Equation 4, with $y = .5$, that if the stimuli are of equal salience they will each contribute the square root of half of their associative strength associative strength to the compound, so that the response to the compound will be greater than to either stimulus alone. It must be said that the factors that determine whether or not summation will be seen when two CSs are presented together are more complex than just implied (see Pearce, 2002) but, at first sight, the proposed amendment to configural theory provide a straightforward explanation for this phenomenon.

Feature positive effect. This effect refers to the finding that a discrimination between a single stimulus, A, and a compound of A with another stimulus, B, is easier to solve when the discrimination is of the feature-positive form, AB+ A-, than of the feature-negative form, A+ AB-. This effect is extremely robust and has been demonstrated with a variety of species including rats (Reberg &LeClerc, 1977), pigeons (Jenkins & Sainsbury, 1969) and humans (Lotz, Uengoer, Koenig, Pearce, & Lachnit, 2012). Given its reliance on Equation 1, the theory of Pearce (1994) incorrectly predicts that the two discriminations will be acquired at the same rate. However, once Equation 1 is replaced by Equation 4, then the asymmetry of the feature-positive effect can be explained in the same way that the amended theory can explain the asymmetry found with a loudness discrimination when the trials are not separated by an ITI. To explain this asymmetry, it was suggested that a loud stimulus is treated as AB and a soft stimulus as A. It was then shown that, by relying on Equation 4, the theory of Pearce predicts that an AB+ A- discrimination will be acquired more readily than A+ AB-. In other words, the asymmetry observed with a loudness discrimination can be regarded as a further example of the feature-positive effect (see also Todd et al., 2010).

An alternative explanation for the asymmetry in the discrimination of magnitude

The amended theory of Pearce (1994) is not unique in being able to explain the asymmetry found with magnitude discriminations. A rather different explanation for the effects that were considered can be obtained from the Rescorla-Wagner (1972) theory. According to this theory, if a pattern of stimulation is presented for a conditioning trial, changes in associative strength do not take place to a representation of the overall pattern of stimulation, as assumed by configural theory. Rather, each element of the pattern is assumed to enter into an association with the US according to Equation 5.

$$\Delta V_A = \alpha.\beta.(\lambda - V_T) \qquad 5$$

In this equation, the increment in associative strength of a single stimulus, A, on a given trial, ΔV_A, is determined by the discrepancy between the asymptote for conditioning and the sum of the associative strengths of all the cues present at the same time, V_T. A further assumption of the theory is that the strength of the response to a compound of two or more stimuli is determined by V_T. The values of the learning rate parameters, α and β, which can vary between 0 and 1, are determined by properties of the CS and the US, respectively. The theory is able to explain readily such effects as summation, the feature positive effect, and the asymmetrical generalisation between elements and compounds that were considered in the previous section. Furthermore, if the above elemental characterisation of the cues used for magnitude discriminations is adopted, and it is assumed that the associative strength of each element changes according to Equation 5, then the Rescorla-Wagner (1972) theory can explain all the results that were successfully accounted for by the amended theory of Pearce (1994).

It would be a mistake to conclude, however, that the two theories make the same predictions concerning all findings relating to the discrimination of magnitude. A very important difference between the theories becomes apparent when considering the results of an experiment by Kosaki, Jones and Pearce (2013). Two groups of rats were trained to escape from a square, grey, pool by swimming to a submerged platform. The walls of the pool were 141cm in length. For one group, a short black panel of width 15 cm was attached to each of two opposite walls, while a long black panel of 45 cm was attached to each of the remaining two walls. An escape platform was situated near the middle of both short panels. Thus rats received a short+/long- discrimination. Subsequent test trials, in which the rats were released into the pool in the absence of the platforms revealed a preference, albeit modest, for the short over the long panels. The second group received similar treatment, except that the lengths of each of the four panels was increased by 55 cm to 70 cm and 100 cm. The rats were unable to solve this discrimination and during testing spent an equal amount of time in front of the long and the short panels. This pattern of results may not be surprising, as it is entirely in keeping with Weber's Law. It follows from the Law that for a given difference between a pair of lengths, in this case 30 cm, a discrimination will be easier to solve when the two lengths are of short rather than large magnitude. In other words, a 15+/45- discrimination will be easier to solve than 70+/100-.

In order to determine the prediction made by the Rescorla-Wagner (1972) theory concerning the experiment by Kosaki et al. (2013), we once again assume that different magnitudes of the same stimulus will excite different sets of common and unique elements. To keep matters simple, black panels of lengths 15, 45, 70, and 100 cm were represented, respectively, by A, AB, ABC, and ABCD. With this characterisation, A represents a panel of 15 cm, B represents the difference between a 15-cm and a 45-cm panel, C represents the difference between a 45-cm and a 70-cm panel and D represents the difference between a 70-cm and a 100-cm panel. For the purpose of analysis, the two discriminations then become A+/AB- and ABC+/ABCD-. Figure 5 shows the results of a computer simulation for these discriminations, which was based on Equation 5, with a value of .2 assigned to the two learning rate parameters, and values of 1 and 0 assigned to λ for reinforced and nonreinforced trials, respectively. The figure shows that the Rescorla-Wagner (1972) theory predicts that the ABC+/ABCD- discrimination will be acquired more readily than A+/AB-, which is opposite to the outcome reported by Kosaki et al. (2013). In essence, the rationale behind this prediction is that on each reinforced trial, there will be an opportunity for three elements to gain associative strength with ABC+/ABCD-, and only one element with A+/AB-. As a consequence, on the nonreinforced trials, D can be expected to gain negative associative strength quite rapidly and lead to a swift solution of the discrimination. In contrast, the acquisition of negative associative strength by its counterpart, B, for the A+/AB-discrimination will be relatively slow and lead to this problem being solved with more difficulty than ABC+/ABCD-. An interesting feature of the prediction by the Rescorla-Wagner (1972) theory is that it implies that a discrimination between two patterns, which might be regarded as being quite different, A+/AB-, is predicted to be acquired more slowly than between two patterns, which might be regarded as quite similar, ABC+ ABCD-. Pearce (1994) identifies a number of additional experimental designs where the Rescorla-Wagner (1972) theory makes a similar, counter-intuitive prediction and, for each of them, the prediction turns out to be incorrect.

FIGURE 5 ABOUT HERE

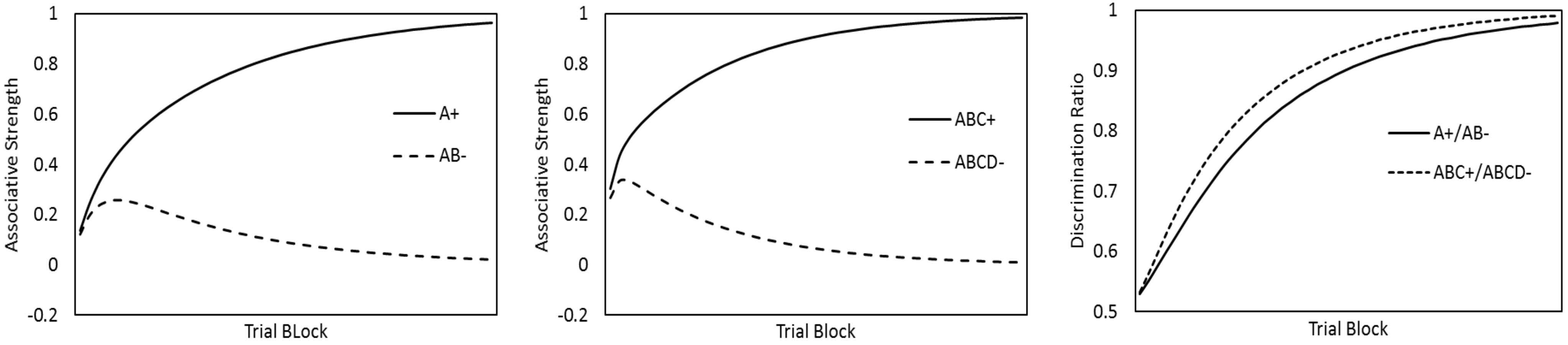

Figure 5. Computer simulation of A+/AB- (left-hand panel) and ABC+/ABCD- (centre panel) discriminations based on the Rescorla-Wagner model. The right hand panel displays discrimination ratios for both discriminations.

A further simulation was conducted in order to determine the prediction made by the modified theory of Pearce (1994) for the experiment by Kosaki et al. (2013). The simulation was based on the same characterisation of the experimental stimuli that was used for the Rescorla-Wagner (1972) simulation. It is evident from the results depicted in Figure 6 that the modified theory of Pearce (1994) predicts correctly that the 15+/45- discrimination will be acquired more readily than the 70+/100- discrimination. According to the theory, A+/AB- training will result in A gaining positive associative strength and the configuration of AB gaining negative associative strength. Likewise, for ABC+/ABCD-, the pattern of ABC will gain positive associative strength, and the pattern of ABCD will gain negative associative strength. After a few trials with A+/AB-, when A is presented it will receive through generalisation, according to Equation 4, half of the negative associative strength of AB. For similar reasons, after a few trials with ABC+/ABCD-, when ABC is presented it will receive three-quarters of the negative associative strength of ABCD. These differences in the generalisation of negative associative strength will result in a greater disruption of responding on trials with ABC than A, and enable the A+/AB- discrimination to be solved more swiftly than ABC+/ABCD-.

FIGURE 6 ABOUT HERE

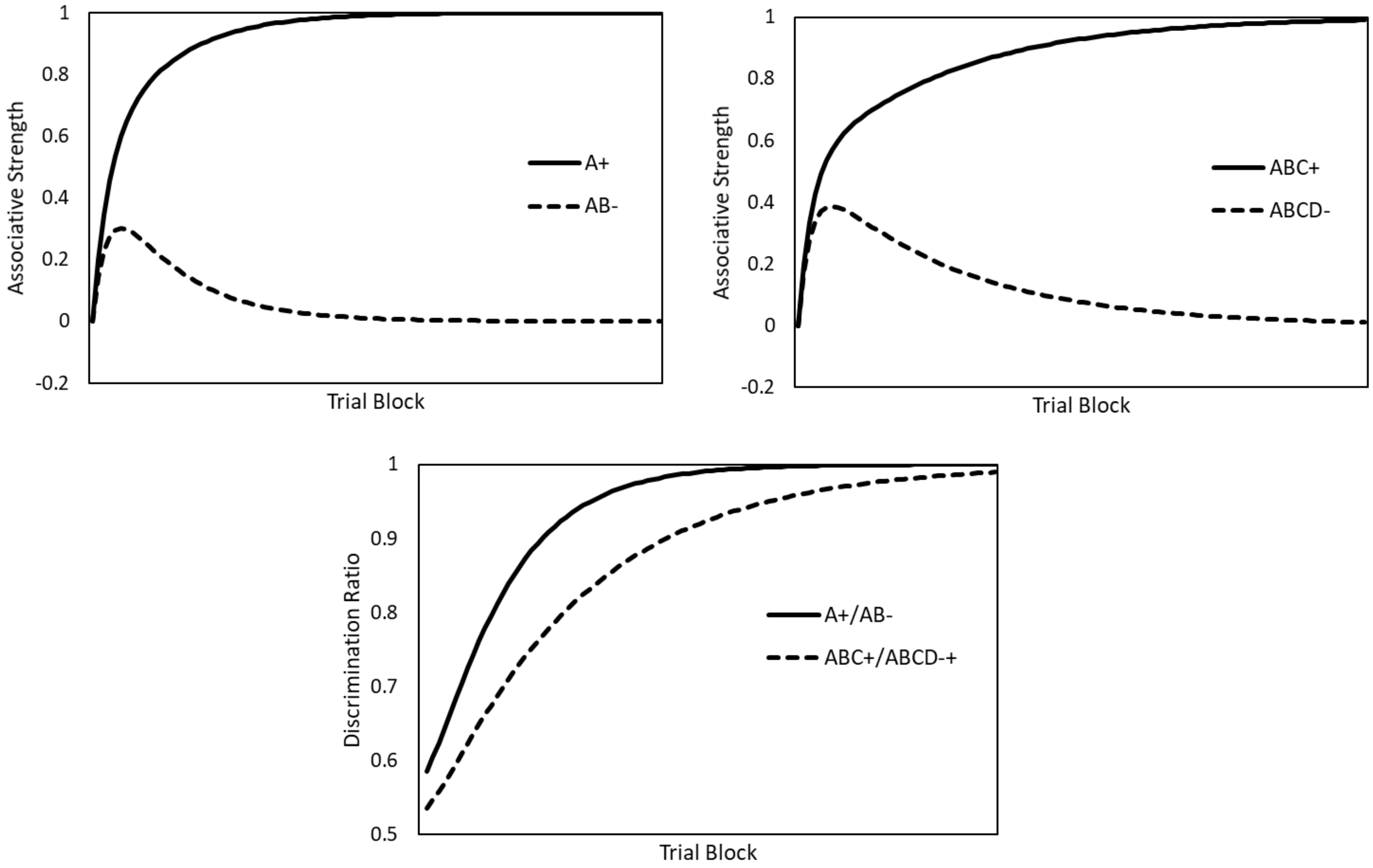

Figure 6. Computer simulation of A+/AB- (left-hand panel) and ABC+/ABCD- (centre panel) discriminations based on the modified Pearce equations. The right-hand panel displays discrimination ratios for both discriminations.

It might be argued that the problem posed by the results of Kosaki and Pearce (2015) for the Rescorla-Wagner (1972) theory depends upon a particular elemental characterisation of the experimental stimuli. This argument does not appear to be correct. For example, it is possible to adopt a different characterisation, where A is again 15 cm of black panel, B is 30 cm, and C is 55 cm. The discrimination 15+/45- then becomes A+/AB-, and 70+/100- becomes AC+/ABC-. In these circumstances Pearce (1994) has shown that that the Rescorla-Wagner (1972) theory again wrongly predicts that 70+/100- will be solved more readily than 15+/45-. For a further way of characterising the stimuli, which again leads the Rescorla-Wagner theory into making the wrong prediction, see Kosaki and Pearce (2015, p. 352). At the same time, these additional characterisations, when analysed with the amended theory of Pearce, lead to it making the correct predictions.

Concluding Comments

A major concern of the present article has been to review the results from discriminations in which an outcome was signalled by either the larger, but not the smaller magnitude of a given stimulus, or by the smaller, but not the larger magnitude, of the same stimulus. In general, such magnitude discriminations are easier to solve when the outcome is signalled by the larger rather than the smaller version of the stimulus. This asymmetry has been observed with the intensity of an auditory cue or an odour, temporal duration, quantity, and physical length.

One reason why this pattern of results is of interest is because it challenges the view that for a given pair of stimuli, it should not matter which of them is used to signal the outcome of a discrimination. That is, the results challenge the prediction that the acquisition of a discrimination between two stimuli will be symmetrical. This prediction follows from a variety of theories of discrimination learning (e.g. Atkinson & Estes, 1963; Pearce, 1994), which are based upon the supposition that the ease of a discrimination between two patterns of stimulation is determined by their similarity, and that the similarity between two stimuli is symmetrical. In view of the results reviewed above, it was therefore deemed necessary to modify one of the theories based on this set of assumptions (Pearce, 1994). The modification was relatively small, yet it enabled the theory to explain the various demonstrations of the asymmetry in the discrimination of magnitude reviewed above. Moreover, as far as we can tell, the modification does not lead the theory of Pearce (1994) into making predictions that are at odds with established experimental findings.

A second reason why the results from the experiments reviewed above are of interest is because of their implications for our understanding of how animals represent different magnitudes of the same stimulus. With a few exceptions, researchers have said rather little about how animals represent magnitudes. One possibility explored in the present article is that a small version of a stimulus excites a set of elements, while a larger version of the same stimulus excites the same set of elements as well as additional, different elements. This idea has been presented in an extremely simple way. It would be possible to elaborate on this presentation by assuming that an intense version of a stimulus not only excites additional elements to a weaker version, but also excites the common elements to a greater extent, and perhaps in greater number. While these possibilities may seem plausible, we have not pursued them here for two reasons. First, the acceptance of these possibilities would require a more sophisticated theoretical model than the one currently under consideration in order to determine the similarity between an intense and a weak version of the same stimulus. Second, we do not know of any behavioural evidence that would justify a more complex development of the theory of Pearce (1994) than the one proposed above.

One, possibly surprising, conclusion that can be drawn from our review of experiments involving magnitude discriminations is that quantity is represented in a different way to stimulus intensity and temporal duration. The justification for this conclusion is that the asymmetry found with discrimination based on quantity, depends upon an interval between successive trials, while this does not appear to be the case for stimulus intensity and stimulus duration. A plausible explanation for why the representation of quantity is different to that of intensity and duration, is that an increase, say, in the number of objects displayed involves a change to the pattern of stimulation, while an increase in the intensity or duration involves the same pattern of stimulation being presented more intensely or for a longer period of time. Thus, even though discriminations based on quantity reveal an asymmetry akin to that found with the magnitudes of loudness and temporal duration, it appears this common outcome occurs for different reasons.

The dominant theoretical perspective throughout this article has been that animals solve discriminations by acquiring associations between representations of stimuli and representations of outcomes. As far as the representation of stimulus magnitude is concerned, it is assumed that the stimulus in question activates a particular pattern of elements, which varies for changes in magnitude. A radically different characterisation of how animals represent different magnitudes is through symbols that stand for different quantitative properties of the stimuli to which they are exposed (Gallistel & Gibbon, 2002). Such properties include intensity, amount, duration, and length. By way of example, the duration of a stimulus is said to be measured by an internal clock, and then stored symbolically as a value. Without going into detail, a discrimination between a short and a long stimulus would then be based on the elapsed time since the current trial commenced being compared with memories of the durations of previous reinforced and nonreinforced trials. At first sight, it is not obvious from this account why a long+/short- discrimination should be easier to master than short+/ long-. Thus, at least at first glance, the asymmetry found with magnitude discriminations based on duration poses a challenge to symbolic accounts of conditioned behaviour. A similar conclusion can be drawn for similar reasons from the asymmetry found with discriminations based on quantity. If different quantities are represented by different symbolic variables, then it is hard to see why a many+/few- discrimination will be easier to solve than few+/many-.

Turning now to the experiment by Kosaki and Pearce (2013), in which rats were required to locate an escape platform in the middle of either the short or the long walls of a rectangular pool, Gallistel (1993) argued that the ability of a rat to located a platform in this kind of environment depends upon the formation of an abstract representation of the overall shape of the pool. In this case it should not matter whether the platform is situated in the middle of the long or the short walls. Provided the representation is accurate, it should be used to find the platform readily, no matter where the platform is situated. In fact, as already noted, the rats found it easier to locate the platform when it is situated in the middle of the long rather than the short walls, which poses a serious problem for this kind of account of spatial learning.

The discrimination of magnitude has received relatively little attention, both empirically and theoretically. Even so, the experiments on this topic may be of considerable importance. On the one hand, the asymmetry in the discrimination of magnitude revealed in these experiments poses a challenge to several different theories of discrimination learning. On the other hand, the asymmetry may have profound implications for our understanding of how animals represent magnitudes, and whether or not this representation is symbolic in nature

References

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. http://dx.doi.org/10.1037/xan0000062

Atkinson, R. C., & Estes, W. K. (1963). Stimulus sampling theory. In R. D. Luce, R. B. Bush, & E. Galanter (Eds.), Handbook of mathematical psychology (Vol. 3, pp. 121-268). *New York: Wiley*.

Beran, M. J. (2016). “Zeroing” in on mathematics in the monkey brain. *Learning & behavior, 44*(1), 4-6.

Blue, S. (1967). Stimulus intensity dynamism during two-valued auditory discrimination acquisition: a systematic replication. *Psychonomic Science*, *8*(1), 29–30. https://doi.org/10.3758/BF03330651

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*(4), 381–393. http://doi.org/10.1037/a0029726

Bouton, M.E., & García-Gutiérrez, A. (2006). Intertrial interval as a contextual stimulus. *Behavioural Processes*, *71*, 307-317. https://doi.org/10.1016/j.beproc.2005.12.003

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 Behaviour Processes, 37*, 79-93. http://dx.doi.org/10.1037/a0021214

Brandon, S. E., Vogel, E. H., & Wagner, A. R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behavioural Brain Research*, *110*(1–2), 67–72. https://doi.org/10.1016/S0166-4328(99)00185-0

Ditz, H. M., & Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. *Proceedings of the National Academy of Sciences*, *112*(25), 7827-7832. https://doi.org/10.1073/pnas.1504245112

Emmerton, J., Lohmann, A., & Niemann, J. (1997). Pigeons' serial ordering of numerosity with visual arrays. *Animal Learning & Behavior*, *25*(2), 234–244. https://doi.org/10.3758/BF03199062

Gallistel, C. R. (1993) A conceptual framework for the study of numerical estimation and arithmetic reasoning in animals. In S. T. Boysen & E. J. Capaldi (Eds.) *The development of numerical abilities: Animal and human models.* Hillsdale, NJ: Lawrence Erlbaum Associates. pp. 211-224.

Gallistel, C. R. and J. Gibbon (2001). Computational versus associative models of simple conditioning. *Current Directions in Psychological Science, 10*(4), 146-150. https://doi.org/10.1111/1467-8721.00136

Gray, J.A. (1965). Stimulus Intensity Dynamism. *Psychological Bulletin, 63*, 180-196. http://dx.doi.org/10.1037/h0021700

Hull, C.L. (1949). Stimulus intensity dynamism (V) and stimulus generalization. *Psychological Review, 56*, 67-76. http://dx.doi.org/10.1037/h0058051

Inman, R. A., Honey, R. C., & Pearce, J. M. (2015). Asymmetry in the Discrimination of Quantity: The Role of Stimulus Generalization. *Journal of Experimental Psychology. Animal Learning and Cognition*, *41*(4), 309–321. http://doi.org/10.1037/xan0000073

Inman, R. A., Honey, R. C., Eccles, G. L., & Pearce, J. M. (2016). Asymmetry in the discrimination of quantity by rats: The role of the intertrial interval. *Learning & Behavior*, (2013). https://doi.org/10.3758/s13420-015-0191-0

Inman, R.A., Honey, R.C., & Pearce, J.M. (2016). Asymmetry in the discrimination of auditory intensity: Implications for theories of stimulus generalisation. In J.B. Trobalon, & V.D. Chamizo (Eds.), Associative Learning and Cognition: Homage to Professor N.J. Mackintosh. Edicions de la Universitat de Barcelona.

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.

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. http://dx.doi.org/10.1037/0097-7403.8.4.313

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 and Behaviour, 30*, 208-216. https://doi.org/10.3758/BF03192830

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

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*(4), 342–356. http://doi.org/10.1037/a0032570

Kosaki, Y., & Pearce, J. M. (2015). Asymmetrical generalization of length in the rat. Journal of Experimental Psychology. Animal Learning and Cognition, 41(3), 266–276. https://doi.org/10.1037/xan0000056

Kyd, R. J., Pearce, J. M., Haselgrove, M., Amin, E., & Aggleton, J. P. (2008). The effects of hippocampal system lesions on a novel temporal discrimination task for rats. *Behavioural Brain Research*, *187*(1), 159–171. https://doi.org/10.1016/j.bbr.2007.09.010

Logan, F.A. (1954). A note on stimulus intensity dynamism (V). Psychological Review, 61, 77-80. http://dx.doi.org/10.1037/h0057534

Lotz, A., Uengoer, M., Koenig, S., Pearce, J. M., & Lachnit, H. (2012). An exploration of the feature-positive effect in adult humans. *Learning {&} Behavior*, *40*(2), 222–230. https://doi.org/10.3758/s13420-011-0057-z

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

Meltzer, D., & Hamm, R. J. (1976). Response summation in the pigeon. *Bulletin of the Psychonomic Society*, *7*(6), 515–518. https://doi.org/10.3758/BF03337268

Moore, J. W. (1964). Differential eyelid conditioning as a function of the frequency and intensity of auditory CSs. *Journal of Experimental Psychology, 68*, 250-259.

Nieder, A., & Miller, E. K. (2003). Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. *Neuron, 37*(1), 149–157. https://doi.org/10.1016/S0896-6273(02)01144-3

Okuyama, S., Kuki, T., & Mushiake, H. (2015). Representation of the Numerosity “zero” in the Parietal Cortex of the Monkey. *Scientific Reports*, *5*, 10059. https://doi.org/10.1038/srep10059

Pavlov, I.P. (1927). Conditioned reflexes. London: Oxford University Press Pearce 1987

Pearce, J.M. (1994). Similarity and discrimination: a selective review and a connectionist model. *Psychological Review, 101,* 587-607. http://dx.doi.org/10.1037/0033-295X.101.4.587

Pearce, J.M. (2002). Evaluation and development of a connectionist theory of configural learning. *Animal Learning and Behaviour, 30,* 73-95. https://doi.org/10.3758/BF03192911

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. http://dx.doi.org/10.1037/h0063567

Pierrel, J., Sherman, J.G., Blue, S., & Hegge, F.W. (1970). Auditory discrimination: A three-variable analysis of intensity effects. *Journal of the Experimental Analysis of Behaviour, 13,* 17-35. http://dx.doi.org/10.1901/jeab.1970.13-17

Relkin, E. M., & Doucet, J. R. (1997). Is loudness simply proportional to the auditory nerve spike count?. *The Journal of the Acoustical Society of America, 101*(5), 2735-2740. http://dx.doi.org/10.1121/1.418561

Reberg, D., & Leclerc, R. (1977). A feature positive effect in conditioned suppression. *Animal Learning & Behavior, 5*(2), 143–147. https://doi.org/10.3758/BF03214069

Rescorla, R. A. (1997). Response inhibition in extinction. The Quarterly Journal of Experimental Psychology: Section B, 50(3), 238-252.

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

Sadowsky, S. (1966). Discrimination learning as a function of stimulus location along an auditory intensity continuum1. *Journal of the Experimental Analysis of Behavior, 9*(3), 219–225. https://doi.org/10.1901/jeab.1966.9-219

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. http://dx.doi.org/10.1037/h0056975

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. https://doi.org/10.1016/j.beproc.2010.01.002

Vonk, J., & Beran, M. J. (2012). Bears "count" too: Quantity estimation and comparison in black bears, Ursus americanus. *Animal Behaviour, 84*(1), 231–238. https://doi.org/10.1016/j.anbehav.2012.05.001

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. https://doi.org/10.3758/BF03328970

Zielinski, K., & Jakubowska, E. (1977). Auditory intensity generalization after CER differentiation training. *Acta Neurobiologiae Experimentalis, 37*(3), 191–205.