Metadata of the chapter that will be visualized online

Chapter Title	Attention and Pavlovian Conditioning	
Copyright Year	2011	
Copyright Holder	Springer Science+Business Media, LLC	
Corresponding Author	Family Name	**Haselgrove**
	Particle	
	Given Name	**Mark**
	Suffix	
	Division/Department	School of Psychology
	Organization/University	The University of Nottingham
	Street	University Park
	Postcode	NG7 2RD
	City	Nottingham
	Country	UK
	Email	Mark.Haselgrove@Nottingham.ac.uk

A

Attention and Pavlovian Conditioning

Mark Haselgrove
School of Psychology, The University of Nottingham, Nottingham, UK

Synonyms

Associability; Classical conditioning; CS processing

Definition

Pavlovian conditioning is a procedure for studying the properties and mechanisms of learning. In this procedure an initially neutral stimulus (the conditioned stimulus, CS) is repeatedly paired with a biologically significant stimulus (the unconditioned stimulus, US). As a consequence of these pairings, the CS comes to evoke a learned response, the conditioned response (CR). The most successful analysis of Pavlovian conditioning is provided by associative theories, which propose that pairings of the CS and US establish an associative connection or link between representations of these stimuli. An issue of continuing theoretical and empirical scrutiny is whether associative connections are determined by (1) variations in the processing of the US or, in contrast, (2) variations in the processing of (or attention to) the CS.

Theoretical Background

Two theories of attention and learning have had a substantial impact upon the Pavlovian conditioning literature. These theories are, at face value, contradictory.

Mackintosh (1975)

According to the theory proposed by Mackintosh (1975), the change in the connection between the CS and the US is determined, on each trial, by the difference between the magnitude of the US and the associative strength of the CS. Importantly, this value is multiplied by a learning rate parameter, α, which reflects the attention paid to a stimulus and changes with experience. More specifically, the value of α increases if the CS is the best available predictor of the US on a conditioning trial, and decreases if it is no better a predictor of the US than any other CS on a conditioning trial. This model builds upon early observations by Krechevsky (1932), which were developed further by Sutherland and Mackintosh (1971). The theory proposed by Mackintosh has been instantiated as a neural network model by Kruschke (2001).

Pearce and Hall (1980)

According to the theory proposed by Pearce and Hall (1980), the change in the connection between the CS and the US is determined, again on each trial, by the product of three parameters, two of which are fixed and are determined by the physical properties of the CS and the US, and the third (again, α), which reflects the attention that will be paid to the CS on the next trial, and again changes with experience. The value of α is equal to the absolute difference between the magnitude of the US and the sum of the associative strengths of all the CS present on that trial. In contrast to Mackintosh's theory, therefore, Pearce and Hall's theory stipulates that learning will progress more to CSs that are followed by surprising, or unpredictable USs. Pearce & Hall's theory was based upon observations first reported by Hall and Pearce (1979), and has been expanded upon by Pearce, Kaye, and Hall (1981). The principles proposed by Pearce and Hall have been incorporated into the neural network model proposed by Schmajuk, Lam, and Gray (1996).

Important Scientific Research and Open Questions

Latent Inhibition

Lubow and Moore (1959, see also Lubow 1973) reported an experiment in which the acquisition of Pavlovian conditioning was retarded if the CS had been pre-exposed in the absence of the US. This effect has been obtained in a variety of Pavlovian conditioning procedures, such as conditioned emotional responding and flavor-aversion learning. In addition, simple pre-exposure to the CS has been shown to attenuate the acquisition of inhibitory conditioning (e.g., Rescorla 1971). Although open to alternative analyses (e.g., Wagner 1981; Bouton 1993)

Norbert Seel (ed.), *Encyclopedia of the Sciences of Learning*, DOI 10.1007/978-1-4419-1428-6,

this latent inhibition effect has been taken as evidence for animals learning to ignore the CS. The effect follows from the theory proposed by Mackintosh (1975) as, during pre-exposure, attention to the CS will fall; this follows because the CS is no better a predictor of the absence of the US than is the background context. Latent inhibition also follows from the theory proposed by Pearce and Hall (1980) as, during pre-exposure, the CS is never followed by a surprising US – attention to the CS will therefore fall. Latent inhibition has been taken as a model of the attentional dysfunction that is observed in acute schizophrenia (e.g., Weiner 2003).

Blocking

Kamin (1968) described a series of experiments in which prior conditioning with CS A prevented, or blocked, conditioning with CS X when CSs A and X were subsequently conditioned in compound (see Table 1). Blocking is a robust property of Pavlovian conditioning and has been demonstrated across a wide variety of conditioning procedures and species. According to the theories proposed by Mackintosh (1975), and Pearce and Hall (1980), prior conditioning with A should result in a loss of attention to X. A number of experiments are consistent with this prediction, for blocked stimuli are resistant at acquiring new associations (Mackintosh 1978). Furthermore, when a surprising upshift or downshift in the magnitude of the US is introduced following AX, blocking is predicted to be attenuated, as attention should be restored. Again, extant evidence is consistent with this prediction (Dickinson et al. 1976). It must be stated, however, that it seems likely that an additional mechanism contributes to blocking. Baxter, Gallagher, and Holland (1999) showed that lesions of the cholinergic inputs of the hippocampus disrupted the attenuation of subsequent learning about a blocked CS but left blocking itself unaffected. These results might be taken to imply that a US processing mechanism (e.g., Rescorla and Wagner 1972) might also contribute to blocking.

Learned Irrelevance

According to Mackintosh's (1975) theory, conditioning will be retarded if the CS has, in the past, been a poor predictor of the US; that is to say, it has acquired irrelevance. Evidence consistent with this prediction has been provided by, for example, Mackintosh (1973) who exposed one group of rats (Group random) to random presentations of the CS and US, before then examining the speed of conditioning in a test stage in which a predictive relationship was established between the CS and the US. Rats in Group random were slower to learn in the test stage than control rats, for whom the initial, random, training was omitted. Random presentations of the CS and US have also been shown to attenuate the subsequent acquisition of inhibitory conditioning (e.g., Baker and Mackintosh 1977). There is, however, an alternative analysis of learned irrelevance which appeals to the summed effects of CS pre-exposure and US pre-exposure – both of which, alone, are known to retard the acquisition of conditioning (e.g., Bonardi and Ong 2003). It remains to be determined if learned irrelevance represents more than the sum of CS and US pre-exposure effects. If it does, it then remains to be determined if these two phenomena can be explained with an attentional mechanism alone.

An alternative method of demonstrating the effect on learning of irrelevance training is exemplified by the superiority of an intradimensional shift (IDS) over an extradimensional shift (EDS). A particularly clear demonstration of the effect was described by George and Pearce (1998) who presented pigeons with different CSs that signaled the presence and absence of the US and which each comprised two features: a color, and lines at a particular orientation. Once learning in stage one was complete the pigeons transferred to a test discrimination, which again involved different CSs that signalled the presence and absence of the US, and which again comprised color and line orientation features. However, the specific colors and orientations were different to those used in stage one. For animals in the IDS group, the dimension that was relevant to the solution of the discrimination in stage1 (e.g., color) was again relevant at test. For the EDS group the dimension that was irrelevant in stage one was relevant at test. The results showed the test discrimination was learned faster in Group IDS than in Group EDS. These results are compatible with Mackintosh's theory as stage 1 training should establish, for example, color, as the best predictor of the US, and thus attention to this stimulus dimension should increase – easing learning in the test discrimination for Group IDS. At the same time line orientation is irrelevant to the solution of the discrimination in stage one, and this dimension should therefore come to be ignored – hardening the learning in the test discrimination for Group EDS. The IDS/EDS effect has, again, been demonstrated in a variety of species using different conditioning procedures. Furthermore, lesions to the medial frontal cortex in rodents (e.g., Birrell and Brown 2000) and the lateral prefrontal cortex in primates (e.g., Dias et al. 1996) have been shown to attenuate the IDS/EDS effect. The Wisconsin card sorting task is a variety of the IDS/EDS task, and is widely used by neuropsychologists to test for attentional dysfunction in patients with frontal lobe injury or mental illness such as schizophrenia. The attenuation of learning by learned irrelevance training is not consistent with the theory proposed by Pearce and Hall (1980).

Continuous and Partial Reinforcement

It follows from the proposals of Pearce and Hall (1980) that if a CS is followed on each trial with a US (continuous reinforcement), attention to the CS will fall. This has been confirmed by Hall and Pearce (1979) who showed that continuous reinforcement of a CS with a weak shock, retarded conditioning of the same CS when it was subsequently paired with a stronger shock. It also follows from the Pearce and Hall theory that if the CS is intermittently paired with a US (partial reinforcement) attention to the CS will be maintained as the presentation of the US – or its omission – will always be surprising. Consequently, partial reinforcement of a CS should facilitate later conditioning. This prediction was confirmed by Kaye and Pearce (1984) who presented a continuously reinforced group of rats with the sequence light-tone-food, and a partially reinforced group the same sequence intermixed among trials in which the light was presented by itself. In a final test stage, the light was paired directly with food: the previously partially reinforced group showed superior conditioning relative to the continuously reinforced group. Kaye and Pearce (1984) provided direct evidence that a partially reinforced CS maintains more attention than a continuously reinforced CS. They examined the extent to which a localized light evoked an orienting response in rats. Their results showed that a CS that was partially reinforced with food maintained an orienting response for longer than a CS that was continuously reinforced with food. Lesion experiments with rodents have identified the amygdala as a crucial structure that mediates the types of attention posited by Pearce and Hall (Holland and Gallagher 1999). The effects of continuous and partial reinforcements that have been investigated by Pearce and his colleagues are not consistent with the theory proposed by Mackintosh (1975).

Hybrid Models of Conditioning and Attention

It should be apparent from the preceding discussion that there exists – at both a theoretical and an empirical level – a contradiction. On the one hand, Mackintosh's (1975) theory stipulates that CSs that are good predictors of the US will come to attract more attention than CSs that are poorer predictors of the US, and a number of studies have supported this stipulation. On the other hand, the theory proposed by Pearce and Hall (1980) stipulates, to the contrary, that CSs that are poor predictors of the US will come to gain more attention than CSs that are good predictors of the US, and again, a number of studies have supported this stipulation. To resolve this contradiction, it has been suggested that the attention paid to a CS is affected by two processes, and it is the net outcome of the interaction between these processes in any conditioning task that determines whether attention to a CS is high or low (Le Pelley 2004; Pearce et al. 1998). A common assumption of these theories, which differ in detail, is that on every conditioning trial, a calculation is made about how well each CS predicts the US (ala Mackintosh) and about the extent to which each CS is followed by an accurately predicted US (ala Pearce and Hall); and evidence which supports this assumption has recently been provided by Haselgrove, Esber, Pearce and Jones (2010). According to Le Pelley's (2004) theory, the product of Mackintosh and Pearce-Hall attention is then used to determine the total attention that is paid to the CS on the subsequent trial, and simulations of this theory have provided a good fit to the existing conditioning data. A crucial goal for future research is to determine the conditions under which attention and conditioning adheres to the proposals of Mackintosh, or the proposals of Pearce and Hall, and indeed whether separate models of attention are required (Esber and Haselgrove 2011).

Cross-References

▶ Animal Learning and Intelligence
▶ Animal Perceptual Learning
▶ Associative Learning
▶ Computational Models of Classical Conditioning
▶ Conditioning
▶ Discrimination Learning Model
▶ Formal Learning Theory
▶ Mathematical Models/Theories of Learning
▶ Pavlov, Ivan P. (1849–1936)

References

Baker, A. G., & Mackintosh, N. J. (1977). Excitatory and inhibitory conditioning following uncorrelated presentations of CS and UCS. *Animal Learning & Behavior, 5*, 315–319.

Baxter, M. G., Gallagher, M., & Holland, P. C. (1999). Blocking can occur without losses in attention in rats with selective removal of hippocampal cholinergic input. *Behavioural Neuroscience, 113*, 881–890.

Birrell, J. M., & Brown, V. J. (2000). Medial frontal cortex mediates perceptual attentional set shifting in the rat. *Journal of Neuroscience, 20*, 4320–4324.

Bonardi, C., & Ong, S. Y. (2003). Learned irrelevance: A contemporary overview. *Quarterly Journal of Experimental Psychology, 56B*, 80–89.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114*, 80–99.

Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380*, 69–72.

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

Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: A model of attention in associative learning. *Proceedings of the Royal Society B*, (in press).

Hall, G., & Pearce, J. M. (1979). Latent inhibition of a CS during CS–US pairings. *Journal of Experimental Psychology: Animal Behavior Processes, 3*, 31–42.

Haselgrove, M., Esber, G. R., Pearce, J. M., & Jones, P. M. (2010). Two kinds of attention in Pavlovian conditioning: Evidence for a hybrid model of learning. *Journal of Experimental Psychology: Animal Behavior Processes, 36*, 456–470.

Holland, P. C., & Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. *Trends in Cognitive Science, 3*, 65–73.

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior, 1967: Aversive stimulation*. Coral Gables, Florida: University of Miami Press.

Kaye, H., & Pearce, J. M. (1984). The strength of the orienting response during Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 10*, 90–109.

Krechevsky, I. (1932). Hypotheses in rats. *Psychological Review, 39*, 516–532.

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45*, 812–863.

Le Pelley, M. E. (2004). The role of associative history in models of associative learning: A selective review and a hybrid model. *Quarterly Journal of Experimental Psychology, 57B*, 193–243.

Lubow, R. E. (1973). Latent inhibition. *Psychological Bulletin, 79*, 398–407.

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

Mackintosh, N. J. (1973). Stimulus selection: Learning to ignore stimuli that predict no change in reinforcement. In R. A. Hinde & J. S. Hinde (Eds.), *Constraints on learning*. London: Academic Press.

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

Mackintosh, N. J. (1978). Cognitive or associative theories of conditioning: Implications of an analysis of blocking. In H. Fowler, W. K. Honig, & S. H. Pulse (Eds.), *Cognitive processes in animal behavior*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Pearce, J. M., Kaye, H., & Hall, G. (1981). Predictive accuracy and stimulus associability. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Acquisition* (Vol. 3). Cambridge, MA: Ballinger.

Pearce, J. M., George, D. N., & Redhead, E. S. (1998). The role of attention in the solution of conditional discriminations. In N. A. Schmajuk & P. C. Holland (Eds.), *Occasion setting: Associative learning and cognition in animals*. Washington, DC: American Psychological Association.

Rescorla, R. A. (1971). Summation and retardation tests of latent inhibitions. *Journal of Comparative and Physiological Psychology, 75*, 77–81.

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*. New York: Appleton-Century-Crofts.

Schmajuk, N. A., Lam, Y. W., & Gray, J. A. (1996). Latent inhibition: A neural network approach. *Journal of Experimental Psychology: Animal Behavior Processes, 22*, 321–349.

Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning*. New York: Academic Press.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spears & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms*. New Jersey: Erlbaum.

Weiner, I. (2003). The "two-headed" latent inhibition model of schizophrenia: modelling positive and negative symptoms and their treatment. *Psychopharmacology, 169*, 257–297.

t1.1 **Attention and Pavlovian Conditioning. Table 1** Design of a blocking experiment

	Group	Stage 1	Stage 2	Test (and result)
t1.2	Blocking	A → US	AX → US	X (weak CR)
t1.3	Control	–	AX → US	X (Stronger CR)
t1.4				