

## Selective Attention in Animal Discrimination Learning

THOMAS R. ZENTALL  
*Department of Psychology*  
*University of Kentucky*

DONALD A. RILEY  
*Department of Psychology*  
*University of California, Berkeley*

---

**ABSTRACT.** The traditional approach to the study of selective attention in animal discrimination learning has been to ask if animals are capable of the central selective processing of stimuli, such that certain aspects of the discriminative stimuli are partially or wholly ignored while their relationships to each other, or other relevant stimuli, are processed. A notable characteristic of this research has been that procedures involve the acquisition of discriminations, and the issue of concern is whether learning is selectively determined by the stimulus dimension defined by the discriminative stimuli. Although there is support for this kind of selective attention, in many cases, simpler nonattentional accounts are sufficient to explain the results. An alternative approach involves procedures more similar to those used in human information-processing research. When selective attention is studied in humans, it generally involves the steady state performance of tasks for which there is limited time allowed for stimulus input and a relatively large amount of relevant information to be processed; thus, attention must be selective or divided. When this approach is applied to animals and alternative accounts have been ruled out, stronger evidence for selective or divided attention in animals has been found. Similar processes are thought to be involved when animals search more natural environments for targets. Finally, an attempt is made to distinguish these top-down attentional processes from more automatic preattentional processes that have been studied in humans and other animals.

---

**THERE HAVE BEEN TWO APPROACHES** to the study of selective attention in animal discrimination learning. The first, more traditional view involves the hypothesis that experience with a discrimination can determine not only an organism's overt approach to and avoidance of the discriminative stimuli, but also its sensitivity to the dimension along which the discriminative stimuli differ. In the more extreme form of this view (e.g., Krechevsky, 1932), animals select one aspect or dimension of the discriminative stimuli at a time (e.g., spatial location or brightness) and test "hypotheses" about whether differences in stimulus value

along that dimension (e.g., left vs. right) control the delivery of reinforcement (e.g., all turns to the left are reinforced). In a more moderate version of this hypothesis, proposed by Mackintosh (1965; see also Krechevsky, 1937), animals gradually learn to attend to the dimension along which the discriminative stimuli differ (e.g., if the discrimination is black positive, white negative, they will learn to attend to the brightness of the stimuli more than to other irrelevant dimensions, such as spatial location).

The second, more contemporary view of selective attention includes the notion of information overload, borrowed from the human literature on selective attention (Broadbent, 1958). According to this view, attentional processes can best be seen when processing time is limited because information from multiple dimensions cannot be processed simultaneously, and attention to one source often means the loss of information from others.

The difference between these two approaches to selective attention involves the distinction between what dimensional information can be acquired under conditions of virtually unlimited access to the discriminative stimuli and what information the animal is able to extract from a display when the time available may be insufficient to process the entire stimulus complex.

The purpose of this article is to summarize and integrate the results of these two lines of research. The first line of research deals with the question of what is learned when access to the discriminative stimuli is not limited by either brief exposure time or by the amount of information that must be processed (e.g., in a complex stimulus array). According to nonattentional views (Hull, 1943; Spence, 1936), animals learn to associate the absolute properties of all discriminable features of the discriminative stimuli with the appropriate response (depending only on the salience of those features). According to attentional theories (e.g., Sutherland & Mackintosh, 1971), however, when animals learn a discrimination, they also learn to attend to the relationship between the discriminative stimuli (e.g., in addition to learning that a particular shade of gray is correct and a different shade of gray is incorrect, they learn that the lighter stimulus is correct) because brightness differences provide a better differential prediction of reinforcement than do differences along other dimensions.

### **Free Access to the Discriminative Stimuli**

The historical roots of the study of selective attention in animals can be traced to Krechevsky (1932; see also Lashley, 1929), who proposed that learning was a discontinuous process because animals could attend to, or test hypotheses about,

---

*Preparation of this article was facilitated by National Science Foundation Grant BNS-9019080 and National Institute of Mental Health Grant MH-45979.*

*Address correspondence to Thomas R. Zentall, Department of Psychology, University of Kentucky, Lexington, KY 40506; Zentall@pop.uky.edu (e-mail).*

only one dimension at a time. Although the bulk of the evidence suggested that learning is continuous (see, e.g., Blum & Blum, 1949), more moderate views in which attention to the relevant dimension is acquired gradually (Sutherland & Mackintosh, 1971) may be quite consistent with continuous learning.

Another controversy in which attentional processes are thought to be implicated is whether discrimination learning involves the absolute properties of the discriminative stimuli (Hull, 1943; Spence, 1936) or the relationship between them (Köhler, 1929). Relational learning implies that organisms attend to those aspects of the discriminative stimuli that distinguish them from each other. According to such a view, what is learned about a positive stimulus consisting of a large, bright circle, for example, depends on the nature of the negative discriminative stimulus. If it is a small, bright circle, the animal should learn to base its response on size and learn to respond to the stimulus that is *larger than* the other, whereas if the negative stimulus is a large, dark circle, the animal should learn to base its response on brightness and learn to respond to the stimulus that is *brighter than* the other. Thus, relational learning implies that attention is directed to the dimension along which the stimuli fall, and phenomena that depend on relational learning, such as transposition, implicate attentional processes even when not explicitly stated (see Lawrence, 1949, 1952).

The notion of perceptual learning (see Epstein, 1967) also implies the involvement of attentional processes. Although the term *perceptual learning* has been used more broadly to include certain changes in discrimination performance produced by mere exposure to discriminative stimuli in the absence of differential reinforcement (e.g., Gibson & Gibson, 1955; Gibson & Walk, 1956), in many cases the terms *attention* and *perceptual learning* are used interchangeably (see Hall, 1991).

### *Stimulus Generalization Gradients*

When animals are trained to respond to a single stimulus and test stimuli are introduced that differ from the training stimulus, generally along a single dimension, the systematic decrement in responding typically found has been called the *gradient of stimulus generalization* (see Guttman & Kalish, 1956). For example, if a pigeon is trained to peck at a yellow hue, the degree to which it will also peck at other hues will be a direct function of the similarity of the test hues to the training hue. However, there has been considerable controversy about the mechanism responsible for the generalization decrement. Is the generalization process automatic or does the slope of the generalization gradient depend on attentional processes?

According to Hull (1943), generalization is an automatic process that reflects the declining spread of habit strength from the training stimulus to decreasingly similar training stimuli. This view is similar to Pavlov's (1927) notion of automatic irradiation of stimulation, but Hull was not interested in speculating about its physiological basis. In Hull's view, the gradient of stimulus generalization did

not require prior experience with stimulus differences along the tested dimension, nor was it directly influenced by such experience. If the strength of the response varied with such experience, it was because of the algebraic interaction of excitatory gradients (around stimuli associated with reinforced responding) and inhibitory gradients (around stimuli associated with the absence of reinforced responding; Spence, 1937).

Lashley and Wade (1946), on the other hand, proposed that the gradient of responding to test stimuli was not automatic, but was an indication of the extent to which the animal failed to discriminate the training stimuli from the test stimulus. According to Lashley and Wade, the failure to discriminate was influenced by the likelihood that the animals would attend to irrelevant dimensions (i.e., dimensions along which the stimulus values did not vary during the test). Furthermore, attention to irrelevant dimensions should depend on the salience of irrelevant dimensions (e.g., the size, the shape, and the brightness of the stimuli) as well as the magnitude of the difference between the training stimulus and the test stimulus. But most important, according to Lashley and Wade, the animal's ability to discriminate between stimuli (and show a declining gradient of stimulus generalization) should depend on the animal's past experience discriminating between training and test stimuli.

Jenkins and Harrison (1958) reported results that appeared to be consistent with Lashley and Wade's (1946) hypothesis. After training pigeons to peck at a lighted response key in the presence of a 1000-Hz tone, they tested the pigeons with tones that varied in frequency between 250 and 4000 Hz, and found no decrement in responding to the novel tones. When responses by a different group of pigeons were reinforced in the presence of the 1000-Hz tone but not in the absence of the tone, a similar generalization test resulted in regularly decreasing gradients. This result is not consistent with Lashley and Wade's position. Although this second group experienced discrimination training, it was not along the frequency dimension, the dimension that defined differences among the test stimuli.

Critically important to Lashley and Wade's position is the notion that prior experience with stimulus differences is necessary for the appearance of generalization gradients. Thus, experiments in which the prior discrimination history of the animal can be controlled should provide critical tests of their theory.

Some evidence has been reported that monkeys (Ganz & Riesen, 1962) and ducks (Peterson, 1962) reared in the dark and trained to respond to a monochromatic light showed relatively flat generalization gradients. However, when Riley and Leuin (1971) raised chickens in monochromatic light (to avoid the possible degenerative effects of light deprivation) and then trained them to peck a key lit with the same wavelength, when the birds were tested with different monochromatic hues, they showed regular gradients of stimulus generalization. Thus, the results do not offer consistent support for the failure-of-discrimination hypothesis. Prior experience discriminating among hues does not appear to be necessary for the establishment of regularly declining gradients of stimulus generalization.

Thus, the mere demonstration of gradients of stimulus generalization is not sufficient to conclude that attention has been directed to the test dimension.

### *The Easy-to-Hard Effect*

Lawrence (1949) introduced a new view of the attentional interpretation by arguing that in learning a visual discrimination, animals learn both to attend to the dimension defined by differences between the positive and negative discriminative stimuli and to learn an appropriate discriminative response. Lawrence (1952) reported that if rats were trained on an easy (e.g., black–white) discrimination, they would learn a difficult (e.g., light-gray–dark-gray) discrimination faster than if they were given, from the start, the same number of trials of training with the difficult discrimination. Lawrence argued that acquisition of the easy discrimination helped the animal to identify or draw attention to the relevant dimension. Although Lawrence (1955) argued that it would be very difficult to account for his data based on the algebraic summation of absolute gradients of excitation and inhibition, Logan (1966) found convincing support for such an absolute-gradient-summation hypothesis. Thus, although attentional processes may be involved in the easy-to-hard effect, the empirical results are well accounted for by an absolute theory of discrimination learning.

### *Acquired Distinctiveness of Cues*

Lawrence's attack on the gradient-summation account of discrimination learning continued in the form of other transfer of training designs (Lawrence, 1949, 1950). After training rats on a simultaneous black–white discrimination in which the rats were reinforced for approaching the black stimulus, Lawrence (1949) transferred them to a successive black–white discrimination in which the presence of a black maze indicated that a turn to the left would be reinforced, whereas the presence of a white maze indicated that a turn to the right would be reinforced.

According to Lawrence (1949), a theory of learning based on the absolute properties of the stimulus should not lead to the prediction that learning to approach black and to avoid white would facilitate the successive discrimination (relative to prior simultaneous discrimination training involving a different relevant dimension), because in transfer, both response alternatives should be associated either with reinforcement (when black) or with the absence of reinforcement (when white). According to an attentional theory of learning, however, the animals would have learned not only to approach black and avoid white, but also to attend to differences in brightness, and it is this attention to differences in brightness that can account for the positive transfer from the simultaneous discrimination to the successive discrimination found by Lawrence.

Siegel (1969), however, has shown that specific response attachments acquired during simultaneous discrimination training may be relevant during acqui-

sition of successive discrimination. Siegel found that most rats acquired the simultaneous discrimination by consistently orienting to one side of the T maze. If that was the correct alternative (e.g., black), they approached it; if not, they entered the other arm. Siegel suggested that, for approximately half of the rats, the prior orienting response would be compatible with successive discrimination tasks by chance. His reasoning was as follows: if in the original discrimination black was correct, and the rat oriented to the right, and in the successive discrimination right was correct in the presence of black and left in the presence of white, positive transfer should result. For the remaining rats, however, right would be correct in the presence of white, left in the presence of black, and the transfer effects would be expected to be negative. According to Siegel, it is these negative transfer rats that are the key to the effect. These rats have two options. First, they can reverse the approach-black-avoid-white response attachments that they acquired during simultaneous discrimination training. Alternatively, they can maintain their response attachments and reverse their orienting response. If they now orient left, they can maintain their approach-black-avoid-white response attachments. According to Siegel, learning to orient to the opposite side is easier to acquire than learning new response attachments. Thus, attentional processes may not be needed to account for Lawrence's (1949, 1950) transfer of training effects.

### *The Overtraining Reversal Effect*

There is considerable evidence that under a variety of conditions, animals that have been trained to criterion (about 90% correct) on a discrimination often require more trials to reverse that discrimination than animals that have been trained for a number of trials beyond that criterion (Reid, 1953).

To account for this overtraining reversal effect, Mackintosh (1965) proposed a two-process, attentional theory in which the two processes occurred at different rates. According to Mackintosh, the first process involved response attachments to the discriminative stimuli, and the second process involved attention to the dimension defined by the discriminative stimuli (e.g., brightness). The function of overtraining, according to this model, was to increase attention to the relevant dimension, and attention to the relevant dimension should facilitate acquisition of the reversal. A complete account of this attentional theory of discrimination learning, as well as support for it, was reported by Sutherland and Mackintosh (1971).

Although there have been other explanations of the overtraining reversal effect (Riley, 1968), Mackintosh's (1965) attentional theory successfully accounts not only for the effect itself, but also for the several failures to obtain it. According to attentional theory, the overtraining reversal effect should not occur when the original discrimination is relatively easy, or more precisely, when the animal is already attending to the relevant dimension at the start of training. If attention is at an asymptotic level prior to overtraining, no benefit of overtraining would be expected. In fact, failures to find the overtraining reversal effect

often occur when the relevant discrimination is spatial (a salient dimension for most animals), and especially when there are few irrelevant cues (e.g., visual or textural cues) present (Mackintosh, 1965, 1969).

Although Mackintosh's theory has been most thoroughly applied to the successes and failures to obtain the overtraining reversal effect, perhaps a similarly thorough analysis may be equally effective in accounting for the data from experiments involving other phenomena (e.g., the easy-to-hard effect and the acquired distinctiveness of cues) that appear to be inconsistent with an attentional account (see Sutherland & Mackintosh, 1971, for such an attempt).

### *Blocking and Overshadowing*

Attentional effects have also been reported in Pavlovian conditioning experiments, and although the procedures used in that research may not be considered to be directly germane to those described in the other sections of this article, the theoretical implications are quite similar. When, in a Pavlovian conditioning experiment, a conditioned stimulus, A, has regularly preceded an unconditioned stimulus, less conditioning appears to occur to a second stimulus, B, presented in compound with the first (AB), than if no prior conditioning had occurred to A (i.e., if only AB conditioning trials were experienced; Kamin, 1968, 1969). This phenomenon, known as *blocking* because it appears that the conditioning of B has been blocked by prior conditioning of A, has been interpreted as evidence that attention to A during original training reduces attention that could be directed to B during training with the AB compound (Pearce & Hall, 1980). The basis of the Pearce-Hall model is that B will be ignored because it fails to add anything to the predictive value of A.

Although blocking has been explained by some in terms of selective attention, it has also been hypothesized by others to occur at the level of association between the conditioned stimulus and the unconditioned stimulus (Rescorla & Wagner, 1972). According to Rescorla and Wagner, each unconditioned stimulus can support only a certain amount of associative strength, and prior conditioning with A will reduce the amount of associative strength left for B.

Overshadowing is a phenomenon related to blocking, in which the conditioning of B is reduced when it is presented in compound with A, relative to conditioning to B when it is presented alone (Pavlov, 1927). The idea is that attention must be shared between A and B when they are presented in compound, but all of the attention can be directed toward B when it is presented alone (Kamin, 1969). Again, however, Rescorla and Wagner (1972) would attribute the deficit in conditioning of B (when presented in compound with A) to a failure of association. When presented alone, all of the associative strength can be acquired by B, but when presented in compound with A, the associative strength must be shared with A.

Both the attentional and the failure-of-association accounts have been challenged by evidence that overshadowing effects may not necessarily reflect a

deficit in learning at all, but rather may reflect a performance deficit (Matzel, Schachtman, & Miller, 1985). For example, there is evidence that if responding is extinguished to the nonovershadowed stimulus after conditioning of the compound, responding to the overshadowed stimulus may show complete recovery.

Overall, there is suggestive evidence that animals selectively attend to certain aspects of the discriminative stimuli (e.g., the dimension defined by variability in the discriminative stimuli; Mackintosh, 1965) when processing time is not restricted (Riley & Leith, 1976). The inconsistency in findings over the range of effects that have been reported may be attributable in part to the fact that the training conditions may not always encourage animals to use attentional processes. When processing time is not restricted, it may be possible for animals to shift attention often enough for them to appear to attend to all aspects of the stimulus display at once, especially when the discriminations are relatively easy. It may be that stronger evidence for selective attention can be obtained only if the task demands placed on the animal prevent the animal from easily shifting attention among elements of the stimulus display (Riley & Leith, 1976; Riley & Roitblat, 1978).

### **Limited Access to the Discriminative Stimuli**

Access to discriminative stimuli can be limited by allowing relatively little time for the stimuli to be processed and by including more than one relevant element in the stimulus display. The effect of shared or divided attention can then be demonstrated by comparing accuracy of performance when there is only one relevant element in the stimulus array with accuracy when there are two (or more) relevant aspects (Riley & Leith, 1976).

#### *The Element Superiority Effect: Matching-to-Sample*

Maki and Leith (1973; see also Maki & Leuin, 1972) trained pigeons on a matching-to-sample task involving hues and line-orientations. When the sample consisted of vertical lines, the vertical-line comparison was correct. When the sample consisted of horizontal lines, the horizontal-line comparison was correct. Similarly, red comparisons were correct when the sample was red, and green comparisons were correct when the sample was green. Matching accuracy on these single-element sample trials constituted the baseline against which to assess compound-sample trial performance. On compound-sample trials, the samples consisted of one element from each dimension (one hue and one set of line orientations), and, over trials, they occurred in all possible combinations. On compound-sample trials, the comparisons were sometimes hues and at other times they were lines. Maki and Leith found that matching accuracy was consistently better on element-sample trials than on compound-sample trials, at each sample duration tested. In addition, sample duration was itself a reliable predictor of matching accuracy, and the two variables did not interact. Thus, for a given dura-



tion of sample, it appears that pigeons are better able to process a single element sample than both elements of a compound sample.

Results of this type, which have been referred to as the *element superiority effect*, provide support for the hypothesis that when there is limited access to the discriminative stimuli, the animals must attend to one dimension at a time or, more likely, they must share or divide their limited attentional capacity between the two dimensions. Alternative accounts of this effect were soon proposed, however.

*The generalization decrement hypothesis.* The first alternative account is based on differential similarity between the samples and comparisons on element-sample and compound-sample trials. On element-sample trials, an exact match exists between the sample and the correct comparison, whereas on compound-sample trials, the correct comparison matches only one element of the sample. Thus, it could be argued that on compound-sample trials the generalization decrement from samples to comparisons (not present on element-sample trials) may lead to a decrement in matching accuracy.

Maki, Riley, and Leith (1976) assessed the validity of this hypothesis by using compound comparisons on both element- and compound-sample trials. In one experiment, on compound-sample trials, the untested sample element was presented on both comparison keys, such that the correct comparison exactly matched the compound sample. Thus, if the sample was vertical lines on a red background, the correct comparison was the same, and the incorrect comparison was either horizontal lines on a red background (to test for the line orientation of the sample) or vertical lines on a green background (to test for the hue of the sample). If the element superiority effect found by Maki and Leith (1973) was attributable to generalization decrement, this manipulation should have eliminated (or even reversed) the element superiority effect, but it did not.

One could argue, however, that interference may have been produced by including one of the elements from the sample on the incorrect comparison key. Thus, in this experiment, only the comparison for which both elements matched the sample was correct.

Maki and Leith (1973, Experiment 3) controlled for this potential interference by including a condition in which there were compound comparisons with redundant cues. For example, if the sample was vertical lines on a red background, the correct comparison was the same and the incorrect comparison was horizontal lines on a green background (i.e., neither element of the incorrect comparison matched the sample). Not only was an element superiority effect found (i.e., matching accuracy on these trials was worse than on standard element-sample, element-comparison trials), but the pigeons did not perform better on these redundant-cue trials than on standard compound-sample, element-comparison trials. Thus, the absence of sample-comparison identity on compound-sample trials does not appear to be responsible for the element superiority effect.

More direct evidence against the generalization decrement account comes from research in which an element superiority effect has been found when symbolic matching tasks have been used (Brown & Morrison, 1990; Langley & Riley, 1993; Zentall, Sherburne, & Zhang, 1997). For example, although the samples may consist of hues and line orientations, the correct comparison following red and green samples are circle and dot, respectively, whereas the correct comparison following vertical- and horizontal-line samples are blue and white hues, respectively (Zentall et al., 1997). Because neither element of the sample ever matched the correct comparison, when this procedure has been used, the generalization decrement between samples and comparisons cannot account for an obtained element superiority effect.

*The degraded compound hypothesis.* A second alternative to the divided attention account is the notion that compounds cannot be processed as well as elements because when in compound, the elements are partly degraded. That is, in the case of lines and hues, the white lines partially obscure the hues, and the hue background reduces the contrast between the lines and the background. The results of an experiment by Leith and Maki (1975) suggest, however, that when the elements appeared in compound, they were not more difficult to detect than when they appeared by themselves. After demonstrating the element superiority effect in the typical way, Leith and Maki showed that the effect could be almost eliminated if, for an extended number of sessions, only one dimension was tested on compound sample trials. Thus, if the pigeons could learn to attend to only one dimension and ignore the other, the element superiority effect could be greatly reduced (see also D. S. Blough, 1969). If the element superiority effect resulted from stimulus degradation on compound-sample trials, such blocking of test trials should not have reduced the magnitude of the effect. Similarly, the stimulus degradation hypothesis cannot account for the finding that if the to-be-tested element was identified by a cue presented prior to the sample, the element superiority effect was virtually eliminated (Leuin, 1976).

*The comparison uncertainty hypothesis.* Another difference between element-sample trials and compound-sample trials is that only one set of comparisons can be presented on element-sample trials, whereas there is uncertainty about which pair of comparisons will be presented on compound-sample trials. There is evidence, in fact, that in similar contexts, comparison predictability can play a role in matching accuracy (Stonebreaker & Rilling, 1984). Pigeons were trained with a one-to-many, successive matching task (each sample was associated with two correct comparisons), but cues were presented that signaled the dimension from which the comparison would be presented (hues vs. lines). The effect of comparison predictability was demonstrated when, on probe trials, the comparisons were miscued and a significant drop in matching accuracy was found.

At least some of the disruption of matching accuracy found by Stonebreaker and Rilling (1984) could be attributed to the miscuing procedure itself because it involved the presentation of novel stimulus configurations. Furthermore, Lang-

ley and Riley (1993) described an unpublished experiment in which—following two-alternative, one-to-many matching training, involving only element samples—the pigeons were given blocks of sessions during which one of the pairs of comparisons was presented on all trials. Although the comparisons were predictable on these sessions, there was no facilitation of matching accuracy relative to blocks of sessions during which either pair of comparisons could have been presented. Thus, the lack of comparison predictability does not appear to play a major role in this task.

*The differential training hypothesis.* Another alternative account of the element superiority effect attributes it to a bias produced by the typical training procedure. Grant and MacDonald (1986) noted that element-sample training generally precedes the introduction of compound-sample trials; thus, if there is more experience with element-sample trials, better matching accuracy on element-sample trials is to be expected. In recent research, however, training with element and compound samples has been equated, and a clear element superiority effect has been found (Langley & Riley, 1993; Zentall et al., 1997).

But even if the number of compound- and element-sample trials is equated, there may be more experience with each of the element-sample-comparison associations to be acquired than compound-sample-comparison associations. Although in the typical procedure there are four element samples (e.g., red, green, vertical, and horizontal) and there are four compound samples (e.g., red-vertical, red-horizontal, green-vertical, and green-horizontal), for each element sample there is only one sample-comparison association to be acquired; for each compound sample, animals must acquire two sample-comparison associations (one with each possible correct comparison). To control for the possibility that the compound-sample-comparison associations received less training than the element-sample-comparison associations, Zentall et al. (1997) gave pigeons twice as much compound-sample training as element-sample training, and a large element superiority effect was still found. Thus, differential experience with the sample-comparison associations cannot account for the element superiority effect.

*The receptor orientation hypothesis.* Kraemer, Mazmanian, and Roberts (1987) have proposed an account of the element superiority effect based on receptor orientation (what Langley & Riley, 1993, have called the gaze-direction hypothesis). They suggested that, when elements appear at different locations on the sample key (e.g., lines as figure and hues as background), the pigeon may not be able to see both elements equally well at the same time. Such a conflict regarding where to look would not occur, however, on element-sample trials. But, when the need for shifts in gaze direction has been eliminated by using stimulus compounds in which the lines themselves were colored, rather than placing the lines on a colored background, element superiority effects have still been found (Cook, Riley, & Brown, 1992; Lamb, 1988; Lamb & Riley, 1981).

*The retrieval deficit hypothesis.* One of the more theoretically interesting alternative accounts of the element superiority effect was proposed by Lamb (1991). According to Lamb, the element superiority effect may be attributable not to an attentional limitation at the time of sample presentation (i.e., divided attention) but to a retrieval deficit at the time of test. Retrieval deficits have been implicated in apparent memory loss in both humans (e.g., Tulving & Pearlstone, 1966) and animals (e.g., Spear, 1973), and their involvement in the element superiority effect is not unreasonable. Lamb (1991) reported that an element superiority effect found with humans could be eliminated by simply requiring that the choice response on both element- and compound-sample trials be delayed for a short time. Presumably, this delay could not have affected the participants' processing of the compound samples, but it may well have affected their ability to retrieve both elements of the compound sample.

The basis for a retrieval deficit may be that on compound-sample trials, at the time of test, there should be two elements in memory rather than one. If those elements must be retrieved successively, and the relevant element is not retrieved first, either the delay in retrieval or the interference from the first-retrieved element may lead to the decrement in performance. Furthermore, delayed retrieval would lead to delay of reinforcement, and in the case of a two-alternative task (with chance correct at 50%), the pigeon may well have a time-based response criterion of this sort: "If the relevant element has not been retrieved within (for example) 1.0 s, choose randomly." Such a criterion may develop because the reduced probability of reinforcement associated with responding without retrieval may be compensated for by the reduction in delay of reinforcement associated with only a somewhat decreased probability of reinforcement. The probability of reinforcement may be decreased only somewhat because on some trials, retrieval may not be possible because there may be an encoding failure, a memory loss, or a total failure of retrieval.

Zentall et al. (1997) attempted to test the retrieval deficit hypothesis in pigeons in a number of ways. First, they reasoned that longer retrieval times on compound sample trials should result either in longer comparison choice latencies on those trials or, more likely, in differences in the magnitude of the element superiority effect at different comparison choice latencies. However, analysis of the data by latency of the comparison response indicated that there was no evidence for either.

On the basis of a suggestion by Lamb (1991), Zentall et al. (1997) also reasoned that if the element superiority effect resulted from a retrieval deficit and if comparison choice latency could be artificially increased, then the element superiority effect may be reduced or even eliminated. After a number of unsuccessful attempts to train pigeons to withhold their responses to the comparison stimuli once they were presented (while maintaining a reasonable level of matching accuracy on element sample trials), Zentall et al. moved the sample stimuli to the back wall of a long operant chamber with the comparison stimuli presented, 60 cm away, on the front wall. Pigeons in this group were required to peck at the sam-

ple on the back wall to produce the comparisons on the front wall and then walk to the comparisons to make their choice. In spite of the added delay in making their comparison responses (but not necessarily in observing the comparisons), these pigeons showed a clear element superiority effect. In fact, the effect was as large as that found for the control group, which had samples and comparisons presented on the same wall, on adjacent response keys.

*Insensitivity of the element superiority effect to sample duration.* As already mentioned, Maki and Leith (1973) found that pigeons matched element samples more accurately than compound samples at all sample durations tested (.04–5.0 s; see also Brown & Morrison, 1990; Cook et al., 1992; Lamb & Riley, 1981). Furthermore, there is no indication that the magnitude of the element superiority effect decreases with increasing sample duration, as would be expected if the effect results from a limitation on the time needed to process both elements of a compound sample. With sufficient time, both elements should be adequately processed. It may be, however, that processing of the sample elements must occur sequentially, and even with extended time, the pigeon must switch back and forth between them. Roberts (1998) has suggested that whichever element is processed last may interfere to some extent with the memory for the earlier processed element.

*Absence of an element superiority effect with cross-modal compounds.* When samples are composed of elements that come from different modalities, such as hue and auditory frequency (Kraemer & Roberts, 1985) or hue and spatial location (Kraemer et al., 1987), an element superiority effect typically has not been found. These results led Kraemer et al. (1987) to propose that the element superiority effect may result from selective peripheral orientation rather than selective central processing. As already mentioned, however, the element superiority effect can be found even when the elements of the compound appear at the same location (e.g., lines of different hues and different orientations; Cook, Riley, & Brown, 1992; Lamb, 1988; Lamb & Riley, 1981).

As Langley and Riley (1993) suggest, the reason that compound samples consisting of cross-modal elements do not produce an element superiority effect may be because the elements of the compounds are sufficiently different from each other that they do not compete at the level of central processing. Rather than taking such findings as evidence against the divided attention hypothesis, one could view them as defining the boundary conditions of the element superiority effect. A better understanding of the mechanisms responsible for divided attention may be gained by determining the conditions under which stimulus processing is limited when channel capacity is exceeded. As suggested by the preceding discussion, the divided attention account of the element superiority effect appears to have survived a proliferation of alternative accounts, and it is reasonable to conclude that no one theory accounts best for the variety of data reported.

### *The Element Superiority Effect: Maintained Generalization Gradients*

D. S. Blough (1969, 1972) has developed a procedure that allows for the assessment of steady state generalization gradients. In a successive discrimination procedure, brief presentations of the positive stimulus are interspersed among brief presentations of negative stimuli that vary in similarity to the positive stimulus along a particular dimension (e.g., hue). The steepness of the resulting generalization gradient can be taken as a measure of the control of responding by hue. When D. S. Blough (1972) presented stimuli that varied along two dimensions (hue and line orientation) with the positive stimulus consisting of one value on each dimension, control by either dimension was poorer (i.e., the gradient was flatter) than when the presented stimuli differed only along one dimension (either hue or line orientation). These results are generally consistent with results of matching-to-sample experiments (see, in particular, Leuin, 1976) and with the notion that the opportunity to base response decisions on a single element allows for more efficient processing of compounds.

### *Search Image*

Limited access to stimuli may also occur under natural conditions when animals are searching for food, but in this case the limited access would have to be considered self-imposed. Because food items (targets) are often cryptic (i.e., difficult to detect), learning to look for targets or to attend to selected aspects of the stimulus display may allow an animal to detect targets more quickly and thus to forage more efficiently.

*Search image* is the notion that the animal forms a representation of the target stimulus (Pietrewicz & Kamil, 1979) or of selected visual features of the stimulus that make it more detectable from the background (Dawkins, 1971; Langley, 1996); the search image makes it easier for the target to be detected. Tinbergen (1960), for example, working with titmice in forests in the Netherlands, counted the number of different prey types of insects and larvae that parent birds brought to the nest. The critical finding was that the birds overselected the most prevalent prey and underselected newly occurring insects and larvae. Tinbergen assumed that the overselection effect occurred because the frequent encounters with the most prevalent prey established a search image, which facilitated the further search for the most prevalent prey at the expense of the others. Thus, the search image acts as a filter, improving detection of the more prevalent target and lowering detection of the less prevalent target. The hypothesis assumes that search images develop as an adaptive response to the crypticity of the prey (i.e., it allows for the faster and more efficient capture of prey).

In time, laboratory analogs to Tinbergen's (1960) naturalistic research appeared, with pigeons searching displays of grain on backgrounds of different levels of crypticity (Bond, 1983), blue jays searching photographs of moths placed

on tree trunks that varied in crypticity (Pietrewicz & Kamil, 1979), and pigeons searching computer displays of many letters for a single target letter (D. S. Blough, 1989; P. M. Blough, 1984).

Bond (1983) proposed a model of search image that maximized the rate at which targets were taken. He simulated the searching image effect as it might be found in nature by using two different grains of equal discriminability to his pigeons. These grains were presented on display cards covered either with fine gravel, which in its natural state rendered the grain quite cryptic, or with the gravel painted gray, rendering the grain highly visible. The two types of grain were presented in different reciprocal proportions. At all proportions, there was reliable overselection of the more prevalent grain. An important feature of Bond's research is that the overselection occurred only in the cryptic condition, an outcome consistent with the notion that search image occurs only when it is useful to facilitate detection of the grain. There is also evidence for a prediction of Bond's attention-threshold model of search image that time spent searching without encountering a target will tend to disrupt the search image (Langley, Riley, Bond, & Goel, 1996).

An alternative account of the search image effect that requires no internal target representation has been proposed by Gendron and Staddon (1983) and Gendron (1986), and extended by Guilford and Dawkins (1987). According to this account, there is no search image. Instead, animals merely adjust their search rate to the detectability of the targets. A relatively cryptic stimulus requires a slower search rate, whereas for the same level of accuracy, a less cryptic stimulus allows for a faster search rate. Imagine an animal engaging in a mixed run involving both targets. If an intermediate search rate is adopted, it should result in a nonoptimal search rate for either target. Some of the more cryptic targets are likely to be missed, and more time will be spent than is required to find the less cryptic targets. Thus, according to this view, animals should overselect because doing so allows them to search at a rate that is optimal for the particular target being selected. The consequent behavior appears to show divided and selective attention, but for reasons that are unrelated to attentional processes.

An important prediction of the search rate hypothesis is that if the two targets are very similar in crypticity, they should be searched at the same rate, and thus they should be taken in proportion to their availability. Reid and Shettleworth (1992; see also Bond, 1983) have found, however, that overselection occurs even when the targets are of equal crypticity. Furthermore, they reported that experience with one target type increased the probability of selecting a target of that type over another, equally detectable, target type. But it is important that the preference for the target that had been recently experienced occurred only when the two target types were relatively cryptic (thus ruling out a simple preference for the more familiar target).

Another question raised in research involving search image is whether search images occur as a *response* to crypticity or whether they occur under noncryptic conditions but are only *revealed* when the targets are cryptic (Langley et al., 1996). In an experiment in which the search-image-establishing conditions could

be separated from the test conditions, Langley et al. were able to show that the establishment of the search image could occur equally well under either conspicuous or cryptic conditions but that once it was established, the search image was revealed only under the cryptic conditions.

There has also been some discussion in the literature about the nature of the search image. The issue is whether the search image is a representation of the entire stimulus (e.g., a template-like image of the target, Endler, 1988; or learning what the prey looks like, Pietrewicz & Kamil, 1979) or perhaps only of those components of the stimulus that most effectively allow the observer to discriminate the target from the background (Dawkins, 1971). Reid and Shettleworth (1992), arguing from evidence that preferential cuing of one colored grain also improved detection of a grain of a different color, concluded that the animals must be attending to selected features rather than to the image of the total object. Langley (1996) directly tested the feature-versus-image hypotheses by presenting runs of either of two grains on a computer screen. The grains were shown either in runs of either type of grain or in mixed runs, and were followed by single-grain tests of either grain and relatively infrequent tests of grains that were altered so that either the color of the grain or its shape, or both, were changed. If, for example, a change in color produced no increase in reaction time to detect the grain but a change in shape did, then one might conclude that the search image was for some property of the shape. In fact, for one of the grains (wheat), only the color change made a difference. For the other grain (bean), however, both change in the color and change in the shape reduced detection, and a change in both further reduced detection. Langley suggested that for the wheat, the notion of a filter for the value of a specific feature could account for the data, but that for bean detection the hypothesis of a retained image remained tenable. Apparently, under certain conditions, the establishment of a search image will not facilitate target detection if the basis for detection during training (e.g., differential color) is not relevant on test trials (e.g., shape is the feature that distinguishes targets from the background).

A phenomenon that may be related to search image has been reported by P. M. Blough (1989, 1991). Blough found that visual search for a target letter embedded in a display of alphanumeric characters showed priming effects like those found in human visual search experiments. Priming occurs whenever a cue forecasts the identity of the target stimulus and can be distinguished from differential reinforcement that requires selective discrimination for differential outcomes. Blough obtained priming effects as measured by reaction times either by an unbroken series of positive stimulus with no interspersed negative displays (i.e., a runs procedure) or by visual cues that predicted the occurrence of a display containing a positive stimulus. The runs procedure was developed by Pietrewicz and Kamil (1979) and had been found effective in their search image work with blue jays. Blough found strong evidence for divided attention in experiments that showed cuing of one target to be more efficient than cuing of two targets, that is, there was a cost of searching for two designated targets



at the same time. Selective attention was also demonstrated by showing negative effects of target miscuing. Three visual cues were used: Cue A predicting Target A, Cue B predicting Target B, and the control, Cue C, predicting either Target A or Target B. When either Cue A or B correctly predicted the target, reaction time was faster than identifications following Cue C. However, when miscuing occurred (e.g., Cue A followed by Target B), the reaction time was slower than the control reaction times. Blough also found evidence for heightened effects under cryptic conditions, as would be expected if search images occur as an adaptive response to crypticity.

The notion of search image as an attentional process has been challenged recently by evidence reported by Plaisted (1997). Plaisted noted that the facilitated detection that occurs when runs of the same target are encountered is generally confounded with the time between successive presentations of target stimuli. Thus, repeated presentations of the target generally occur with a shorter interstimulus interval in the runs condition than in the mixed-target condition. Plaisted found that when she controlled for the time between repeated presentations of the same target, the interpolation of different targets had little effect on accuracy or latency of target detection. Thus, interpolation of different targets does not result in an attentional shift but rather allows for greater decay of the target memory trace.

Under similar conditions, however, attentional processes have been found to play a role in studies of search image (P. M. Blough & Lacourse, 1994). When Blough and Lacourse presented pigeons with sequences of targets consisting of randomly varying runs or nonruns of the same target for three trials, there was little difference in their reaction time either between trials within a run or between three-trial blocks of runs versus nonruns. However, when the three-trial runs of the same target occurred predictably (each block of three trials consisted of a run, but not the same run), a significant decrease in reaction time was found. Thus, it appears that attentional processes do play a role in search image. However, if the various effects that have been reported using search image procedures are to be reconciled, the parameters involved in those effects must be studied further.

### *Preattentive Search Processes: Texture Discriminations*

Finally, using the feature-discrimination and feature-conjunction-discrimination paradigm developed by Treisman and Gelade (1980) with humans, Cook and his associates (Cook, 1992; Cook, Cavoto, & Cavoto, 1996) attacked the problem of divided attention in animals in a recent series of experiments. In this context, the target might be a green circle or square surrounded by a set of stimuli differing from the target in one feature only (e.g., color: red circles and squares). In a conjunction discrimination, the target differs from the surround in two nonredundant features (e.g., a red circle or a green square surrounded by a mixture of green circles and red squares; in this case no single feature distinguishes the target from the surround). Treisman and Gelade reported that for feature discriminations, the time to

find the target stimulus was the same, regardless of the number of distractors (the number of items in the surround), whereas reaction times for conjunction discriminations increased linearly with increases in the number of distractors. In their account of this effect, they assumed that feature differences are processed preattentively and in parallel, whereas the search for a conjunction requires that the subject attend to each pair until the correct one is found—a search that must be carried out serially. With changes to make the task appropriate for pigeons, Cook demonstrated an analogous effect on accuracy of pecking to the target.

Preattentive processes have also been reported in pigeons during the acquisition of an element discrimination, an oddity task involving a simple red–green hue discrimination (Zentall, Hogan, Edwards, & Hearst, 1980). In this case, however, acquisition was actually facilitated by the increase in the number of distractors (i.e., the number of matching stimuli among which the odd stimulus was placed). Similarly, in human visual search, increasing the number of distractors can sometimes increase the speed with which a target is found (Chastain & Cheal, 1998). Apparently, with these simple displays, not only can the stimuli be processed in parallel, but increasing the number of distractors increases the degree to which the figure stands out from the ground and allows the target to be found more easily.

Unlike the attentional processes examined earlier and those that occur when the target consists of conjunctions of elements from the surround, it appears that these preattentional processes are relatively automatic and do not have to be learned. In the case of conjunctions, in which attentional search processes do seem to be involved, for both humans (Wolfe, 1992) and pigeons (Cook et al., 1996), it appears that it is not the number of features from the surround that are conjoined to produce the target that determines the difficulty of the discrimination, as predicted by feature integration theory (Treisman & Gelade, 1980), but the overall similarity of the distractors to the target (Duncan & Humphreys, 1989). This finding is quite compatible with an analysis in terms of figure–ground relations.

### Summary

In this article, we have attempted to review the literature on attention in animal discrimination learning. Early research focused on whether animals learned not only about the absolute properties of stimuli but also about the relations between the discriminative stimuli (i.e., the dimension defined by the differences between the discriminative stimuli). Attentional processes are relevant to relational learning because it has been proposed that the distinguishing characteristics of the discriminative stimuli cause the animal to attend to the dimension or dimensions that are defined by those differences (e.g., it is assumed that if the discrimination involves learning to respond to a light gray stimulus but not to a dark gray stimulus, then the animal also learns to attend to the dimension of brightness as well).

A different approach to selective attention in animals resulted from the view

of attention as a process necessitated by the inherent limited processing capacity of the organism. In this case, attention refers to the limited ability of the animal to process more than one element of a compound when constraints are placed on processing time, which requires the animal to divide its attention between them. We have also tried to show that similar processes may be involved in more natural tasks involving target search. In the case of more natural tasks, the constraints result more from the value of efficient foraging than from limited access to the discriminative stimuli. Finally, we have tried to distinguish these attentional processes from more fundamental and automatic preattentive processes that do not require learning or search.

### REFERENCES

- Blough, D. S. (1969). Attention shifts in a maintained discrimination. *Science*, *166*, 125–126.
- Blough, D. S. (1972). Recognition by the pigeon of stimuli varying in two dimensions. *Journal of the Experimental Analysis of Behavior*, *18*, 345–367.
- Blough, D. S. (1989). Form similarity and categorization in pigeons' visual search. In M. L. Commons, R. J. Herrnstein, & S. Kosslyn (Eds.), *Quantitative analyses of behavior: Pattern recognition and concepts in animals, people, and machines* (pp. 129–143). Hillsdale, NJ: Erlbaum.
- Blough, P. M. (1984). Visual search in pigeons: Effects of memory set size and display variables. *Perception & Psychophysics*, *35*, 344–352.
- Blough, P. M. (1989). Attentional priming and search images in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *15*, 211–223.
- Blough, P. M. (1991). Selective attention and search images in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 292–298.
- Blough, P. M., & Lacourse, D. M. (1994). Sequential priming in visual search: Contributions of stimulus-driven facilitation and learned expectancies. *Animal Learning & Behavior*, *22*, 275–281.
- Blum, R., & Blum, J. (1949). Factual issues in the "continuity" controversy. *Psychological Review*, *56*, 33–50.
- Bond, A. B. (1983). Visual search and selection of natural stimuli in the pigeon: The attention threshold hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 292–306.
- Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.
- Brown, M. F., & Morrison, S. K. (1990). Element and compound matching-to-sample performance in pigeons: The roles of information load and training history. *Journal of Experimental Psychology: Animal Behavior Processes*, *16*, 185–192.
- Chastain, G., & Cheal, M. L. (1998). Automatic versus directed attention with single-element and multiple-element precues. *Visual Cognition*, *5*, 339–364.
- Cook, R. G. (1992). Dimensional organization and texture discrimination in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*, 354–363.
- Cook, R. G., Cavoto, K. K., & Cavoto, B. R. (1996). Mechanisms of multidimensional grouping, fusion, and search in avian texture discrimination. *Animal Learning & Behavior*, *24*, 150–167.
- Cook, R. G., Riley, D. A., & Brown, M. F. (1992). Spatial and configural factors in compound stimulus processing by pigeons. *Animal Learning & Behavior*, *20*, 41–55.
- Dawkins, M. (1971). Shifts of "attention" in chicks during feeding. *Animal Behaviour*, *19*, 575–582.

- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.
- Endler, J. A. (1988). Frequency-dependent predation, crypsis, and aposematic coloration. *Philosophical Transactions of the Royal Society of London*, *319*, 505–523.
- Epstein, W. (1967). *Varieties of perceptual learning*. New York: McGraw-Hill.
- Ganz, L., & Riesen, A. H. (1962). Stimulus generalization to hue in the dark-reared macaque. *Journal of Comparative and Physiological Psychology*, *55*, 92–99.
- Gendron, R. P. (1986). Searching for cryptic prey: Evidence for optimal search rates and the formation of search images in quail. *Animal Behaviour*, *34*, 898–912.
- Gendron, R. P., & Staddon, J. E. R. (1983). Searching for cryptic prey: The effect of search rate. *American Naturalist*, *121*, 172–186.
- Gibson, E. J., & Walk, R. D. (1956). The effect of prolonged exposure to visually presented patterns on learning to discriminate them. *Journal of Comparative and Physiological Psychology*, *49*, 239–242.
- Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment. *Psychological Review*, *62*, 32–41.
- Grant, D. S., & MacDonald, S. E. (1986). Matching to element and compound samples in pigeons: The roles of sample coding. *Journal of Experimental Psychology: Animal Behavior Processes*, *12*, 160–171.
- Guilford, T., & Dawkins, M. (1987). Search images not proven: A reappraisal of recent evidence. *Animal Behaviour*, *35*, 1838–1845.
- Guttman, N., & Kalish, H. I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology*, *51*, 79–88.
- Hall, G. (1991). *Perceptual and associative learning*. Oxford, UK: Oxford University Press.
- Hull, C. L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.
- Jenkins, H. M., & Harrison, R. H. (1958). Effect of discrimination training on auditory generalization. *Journal of Experimental Psychology*, *59*, 246–253.
- Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9–31). Miami, FL: University of Miami Press.
- Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). New York: Appleton-Century-Crofts.
- Köhler, W. (1929). *GeStalt psychology*. New York: Liveright.
- Kraemer, P. J., Mazmanian, D. S., & Roberts, W. A. (1987). Simultaneous processing of visual and spatial stimuli in pigeons. *Animal Learning & Behavior*, *15*, 417–422.
- Kraemer, P. J., & Roberts, W. A. (1985). Short-term memory for simultaneously presented visual and auditory signals in the pigeon. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 137–152.
- Krechevsky, I. (1932). "Hypotheses" in rats. *Psychological Review*, *39*, 516–532.
- Krechevsky, I. (1937). A note concerning "The nature of discrimination learning in animals." *Psychological Review*, *44*, 97–103.
- Lamb, M. R. (1988). Selective attention: Effects of cuing on the processing of different types of compound stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 96–104.
- Lamb, M. R. (1991). Attention in humans and animals: Is their capacity limitation at the time of encoding? *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 45–54.
- Lamb, M. R., & Riley, D. A. (1981). Effects of element arrangement on the processing of compound stimuli in pigeons (*Columba livia*). *Journal of Experimental Psychology:*

- Animal Behavior Processes*, 7, 45–58.
- Langley, C. M. (1996). Search images: Selective attention to specific visual features of prey. *Journal of Experimental Psychology: Animal Behavior Processes*, 22, 152–163.
- Langley, C. M., & Riley, D. A. (1993). Limited capacity information processing and pigeon matching-to-sample: Testing alternative hypotheses. *Animal Learning & Behavior*, 21, 226–232.
- Langley, C. M., Riley, D. A., Bond, A. B., & Goel, N. (1996). Visual search for natural grains in pigeons (*Columba livia*): Search images and selective attention. *Journal of Experimental Psychology: Animal Behavior Processes*, 22, 139–151.
- Lashley, K. S. (1929). *Brain mechanisms and intelligence*. Chicago: University of Chicago Press.
- Lashley, K. S., & Wade, M. (1946). The Pavlovian theory of generalization. *Psychological Review*, 53, 72–87.
- Lawrence, D. H. (1949). Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology*, 39, 770–784.
- Lawrence, D. H. (1950). Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation. *Journal of Experimental Psychology*, 40, 175–188.
- Lawrence, D. H. (1952). The transfer of a discrimination along a continuum. *Journal of Comparative and Physiological Psychology*, 45, 511–516.
- Lawrence, D. H. (1955). The applicability of generalization gradients to the transfer of a discrimination. *The Journal of General Psychology*, 52, 37–48.
- Leith, C. R., & Maki, W. S., Jr. (1975). Attention shifts during matching-to-sample performance in pigeons. *Animal Learning & Behavior*, 3, 85–89.
- Leuin, T. C. (1976). *Selective information processing by pigeons*. Unpublished doctoral dissertation, University of California, Berkeley.
- Logan, F. A. (1966). Transfer of a discrimination. *Journal of Experimental Psychology*, 71, 616–618.
- Mackintosh, N. J. (1965). Selective attention in animal discrimination learning. *Psychological Bulletin*, 64, 124–150.
- Mackintosh, N. J. (1969). Further analysis of the overtraining reversal effect. *Journal of Comparative and Physiological Psychology, Monograph Supplement*, 67, 1–18.
- Maki, W. S., Jr., & Leith, C. R. (1973). Shared attention in pigeons. *Journal of the Experimental Analysis of Behavior*, 19, 345–349.
- Maki, W. S., Jr., & Leuin, T. C. (1972). Information processing by pigeons. *Science*, 176, 535–536.
- Maki, W. S., Riley, D. A., & Leith, C. R. (1976). The role of test stimuli in matching to compound samples by pigeons. *Animal Learning & Behavior*, 4, 13–21.
- Matzel, L. D., Schachtman, T. R., & Miller, R. R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learning and Motivation*, 16, 398–412.
- Pavlov, I. P. (1927). *Conditioned reflexes*. London: Oxford University Press.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87, 532–552.
- Peterson, N. (1962). Effect of monochromatic rearing on the control of responding by wavelength. *Science*, 136, 774–775.
- Pietrewicz, A. T., & Kamil, A. C. (1979). Search image formation in the blue jay (*Cyanocitta cristata*). *Science*, 22, 1332–1333.
- Plaisted, K. (1997). The effect of interstimulus interval on the discrimination of cryptic targets. *Journal of Experimental Psychology: Animal Behavior Processes*, 23, 248–259.

- Reid, L. S. (1953). Development of non-continuity behavior through continuity learning. *Journal of Experimental Psychology*, 46, 107–112.
- Reid, P. J., & Shettleworth, S. J. (1992). Detection of cryptic prey: Search image or search rate? *Journal of Experimental Psychology: Animal Behavior Processes*, 18, 273–286.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. A. Prokasy (Eds.), *Classical conditioning: II. Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Riley, D. A. (1968). *Discrimination learning*. Boston: Allyn and Bacon.
- Riley, D. A., & Leith, C. R. (1976). Multidimensional psychophysics and selective attention in animals. *Psychological Bulletin*, 83, 138–160.
- Riley, D. A., & Leuin, T. C. (1971). Stimulus generalization gradients in chickens reared in monochromatic light and tested with a single wavelength value. *Journal of Comparative and Physiological Psychology*, 75, 389–402.
- Riley, D. A., & Roitblat, H. L. (1978). Selective attention and related cognitive processes in pigeons. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 249–276). Hillsdale, NJ: Erlbaum.
- Roberts, W. A. (1998). *Principles of animal cognition*. Boston: McGraw-Hill.
- Siegel, S. (1969). Discrimination overtraining and shift behavior. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning* (pp. 187–213). San Diego, CA: Academic Press.
- Spear, N. E. (1973). Forgetting as retrieval failure. *Psychological Review*, 80, 163–175.
- Spence, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review*, 43, 427–449.
- Spence, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, 44, 430–444.
- Stonebreaker, T. B., & Rilling, M. (1984). Retrospective versus prospective processes in delayed matching to sample. *Bulletin of the Psychonomic Society*, 22, 372–375.
- Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning*. San Diego: Academic Press.
- Tinbergen, L. (1960). The natural control of insects in pine woods: I. Factors influencing the intensity of predation by songbirds. *Archives Neerlandaises de Zoologie*, 13, 265–343.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381–391.
- Wolfe, J. M. (1992). “Effortless” texture segmentation and “parallel” visual search are not the same thing. *Vision Research*, 32, 757–763.
- Zentall, T. R., Hogan, D. E., Edwards, C. A., & Hearst, E. (1980). Oddity learning in the pigeon as a function of the number of incorrect alternatives. *Journal of Experimental Psychology: Animal Behavior Processes*, 6, 278–299.
- Zentall, T. R., Sherburne, L. M., & Zhang, Z. (1997). Shared attention in pigeons: Retrieval failure does not account for the element superiority effect. *Learning and Motivation*, 28, 280–308.

*Manuscript received October 16, 1998*

*Revision accepted for publication February 1, 1999*