THE ROLE OF OBSERVING AND ATTENTION IN ESTABLISHING STIMULUS CONTROL

JAMES A. DINSMOOR

INDIANA UNIVERSITY

Early theorists (Skinner, Spence) interpreted discrimination learning in terms of the strengthening of the response to one stimulus and its weakening to the other. But this analysis does not account for the increasing independence of the two performances as training continues or for increases in control by dimensions of a stimulus other than the one used in training. Correlation of stimuli with different densities of reinforcement produces an increase in the behavior necessary to observe them, and greater observing of and attending to the relevant stimuli may account for the increase in control by these stimuli. The observing analysis also encompasses errorless training, and the selective nature of observing explains the feature-positive effect and the relatively shallow gradients of generalization generated by negative discriminative stimuli. The effectiveness of the observing analysis in handling these special cases adds to the converging lines of evidence supporting its integrative power and thus its validity.

Key words: observing, discrimination, stimulus control, orthogonal training, interdimensional training, feature-positive effect, negative gradients, errorless training, generalization tests

To put the matter as simply and as broadly as possible, the thesis of the present article is that the conformity of behavior to the consequences encountered in the presence of each of two or more alternative stimuli (stimulus control) depends on how much contact the organism has with those stimuli. By contact, I mean first the impingement of the stimulus energy on the receptor cells of the relevant sensory apparatus, which typically requires or is modulated by auxiliary behavior known as observing (e.g., looking at and focusing on the stimulus object, touching it, tasting it, etc.). Second, to complete the picture I think we are obliged to consider analogous processes occurring further along in the sequence of events, presumably in the neural tissue, and commonly known as attention. The processes involved in attention are not as readily accessible to observation as the more peripheral adjustments, but it is my hope and my working hypothesis that they obey similar principles. Otherwise, the study of attention may prove extremely difficult. From the behavioral level it can only be approached indirectly, and we will face the arduous task of distinguishing its behavioral effects in each instance from those to be attributed to changes in observing.

Contemporary analyses of the development of stimulus control continue to be dominated by two very similar accounts, formulated some fifty years ago by Skinner (1933, 1934, 1938) and by Spence (1936, 1937). According to these accounts, the gradual separation of the subject's responding to the alternative stimuli (formation of a discrimination) can be attributed to the growing strength of the behavior produced by continued reinforcement in the presence of one stimulus (S^D or S+) and the waning strength produced by nonreinforcement in the presence of the other stimulus (S^Δ or S−). Yet this account does not deal with the central problem. At the beginning of discrimination training, there is usually a substantial transfer between the two stimuli of the effects of their respective schedules. "The reinforcement of $S_1 - R$ affects the strength of $S_2 - R$ also, and there is a converse sharing of the ex-

tinction" (Skinner, 1933, p. 303). When discrimination training began with the first reinforcement, so that very little strength had been built up in the presence of the positive stimulus, the rate of responding in the presence of the negative stimulus remained quite low (Skinner, 1934, 1938). But when the initiation of differential training had been preceded by several sessions of intermittent reinforcement in the presence of the first stimulus, a substantial number of responses occurred in the presence of the second stimulus (Skinner, 1933, 1938). Sometimes, as Ferster (1951) later observed, the performance in the presence of the negative stimulus could not be distinquished from that in the presence of the positive stimulus. Yet it is clear that by the end of discrimination training this transfer, or induction as Skinner (1934, 1938) called it, has been greatly reduced. Very little responding occurs in the presence of the negative stimulus, despite continued reinforcement in the presence of the positive stimulus.

It may be useful to think of the overall pattern of behavior at any point during discrimination training (acquisition under a multiple schedule) as the product of an interaction between two major factors or processes. The first is the characteristic effect of the schedule of reinforcement in the presence of each stimulus on the rate and the temporal distribution of responding in the presence of that stimulus. This factor is most clearly revealed when the induction between the two components is minimal, so that neither schedule has much effect on the behavior in the presence of the other stimulus. But such independence cannot be taken for granted. The second major factor is the rate or proportion of transfer from one stimulus to the other (induction), which presumably is affected by such variables as the magnitude of the physical difference between the two stimuli, the sensitivity of the species and of the individual subject to that stimulus dimension, and how accurately and how much of the time the subject observes each of the stimuli. It is this second factor, or its converse—the independence of the two performances—that is most appropriately characterized as stimulus control. Although it is

difficult to separate the effects of induction from the more immediate effects of reinforcement and nonreinforcement when both are concurrently producing changes in behavior, it is the decrease in induction as training continues that is the essence of discrimination learning and that sets it apart from other behavioral processes. Both Skinner and Spence were familiar with the basic concept of stimulus generalization and certain of its empirical manifestations, but neither offered an explanation for this change.

According to the conventional account, the final discriminative performance results from the accumulation of small differences. Reinforcement of the response in the presence of S+ increases its strength more in the presence of that stimulus than in the presence of S−; nonreinforcement in the presence of S− reduces the strength more in the presence of that stimulus than in the presence of S+. Contemporary textbook accounts often suggest that these processes take care of the problem, but such an analysis does not work. If the rate of transfer remained as high as it was at the beginning of training, so would the proportion of responses that occurred in S−. The traditional account deals only with what happens in the presence of each stimulus independently, not with the change in the interaction between them.

As mathematical treatments of learning developed, the rigor of their logic revealed much the same problem from a somewhat different perspective. Although Spence (1937) seems to have been aware of the difficulty, Hull (1950) was the first to spell it out. The stimuli controlling the discriminative performance constituted only a small fraction of the total experimental situation. Furthermore, the vast majority of the stimuli that were present when the response was reinforced were also present during periods of nonreinforcement. Consequently, if it was assumed that reinforcement increased the tendency to respond to all of the stimuli present at the time and nonreinforcement decreased the tendency to respond to all of the stimuli, the effects of S+ and S− would be greatly diluted by the control exerted by these "incidental" or "static" stimuli. Even after

prolonged training, the subject should respond to the common stimuli plus S− in a fashion very similar to the way in which it responded to the common stimuli plus S+. It was clear that the effect of these common stimuli had to be reduced before a substantial difference could develop. For some reason, Hull evidently believed that the effect of the nonreinforcement should exactly balance the effect of the reinforcement, resulting in the "neutralization" of these common stimuli, but subsequent writers have not accepted this solution. In one of the earliest mathematical models, Bush and Mosteller (1951) found it necessary to apply a "discrimination operator" (*D*) to reduce the influence of elements common to both sets of stimuli. Similarly, Restle (1955) found it necessary to add an operator expressing the "adaptation" of those cues that were not correlated with or predictive of reward. Another way of saying the same thing, of course, is that the effects of reinforcement and nonreinforcement had eventually to be restricted to those stimuli that were correlated with the two consequences, namely S+ and S−. In the mathematical models, the operators were hypothetical processes necessary to account for the data; but a promising real-world candidate to account for the focusing of the effects of discrimination training on the appropriate stimuli is the observing response, which increases the subject's contact with the relevant stimuli and decreases its contact with those that are not relevant.

Spence (1936, 1937, 1940) recognized the need for what he called "orienting responses" under certain circumstances, before his theory could be applied, but these were treated as a supplementary mechanism rather than as a constituent part of his basic analysis. It remained for Wyckoff (1952, p. 433) to suggest that an initial "crossover" between the effects of reinforcement and nonreinforcement in the presence of the positive and negative stimuli, respectively, might be attributable to the subject's failure to observe the stimuli and that the reduction in this crossover as training continued might not be a purely arbitrary phenomenon but might reflect the acquisition of appropriate observing behavior. Working

within a hypothetico-deductive framework, Sutherland and Mackintosh (1971) proposed a set of rules for changing the strengths of different "stimulus analyzers" in order to account for a variety of special effects in animal discrimination learning, such as transfer, blocking, overshadowing, additivity of cues, and the overlearning reversal effect. Most of their data came from experiments employing simultaneous discriminations (e.g., jumping stands), rather than from the successive discriminations used in the present work. Their rule for the strengthening of stimulus analyzers (p. 44) was very similar to the rule initially proposed by Wyckoff for the strengthening of observing responses. Neither rule, however, is in accord with the findings of more recent research. (For a discussion of this issue, see Dinsmoor, Bowe, Dout, Martin, Mueller, & Workman, 1983, pp. 262-263.) Moreover, Sutherland and Mackintosh assumed that their analyzers were directed toward entire dimensions, rather than toward specific stimuli, and that as the strength of an analyzer for one dimension increased, the strengths of the analyzers for all other dimensions necessarily decreased. Neither of these assumptions is implied by the present account (Dinsmoor, 1973a; for relevant data, see DePaulo, DeWald, & Yarczower, 1977), and both appear subsequently to have been abandoned by at least one of the original authors (Mackintosh, 1975). The approach taken by Honig (1970) and by Honig & Urcuioli (1981) also bears some relation to the present account but does not make use of the research on observing. Sidman (e.g., 1969) has presented an analysis of gradients of generalization that seems quite compatible, attributing flat gradients to control of the response by stimuli other than the one being tested.

ORTHOGONAL DISCRIMINATION TRAINING

One body of data that seems to pose an insurmountable obstacle to interpretation in terms of the traditional strengthening and weakening formulation comes from studies in which the subject was trained to discriminate

between the presence and absence of a given stimulus and then was tested with variations in some aspect of that stimulus. For example, pecking may be reinforced in the presence of a vertical line projected on the pigeon's key but not in its absence. Or pecking may be reinforced in the absence of the line but not in its presence. The rate of pecking is then tested with the line tilted at various angles with regard to the vertical and horizontal axes. This procedure is sometimes known as interdimensional training. In more conventional studies, the stimuli to be discriminated are two points lying along the same dimension as that on which the test is to be conducted. For example, the positive and negative stimuli may be two frequencies, two intensities, or two sizes, and the rate of responding is then tested at a succession of points along the dimension of frequency, intensity, or size, in order to assess the subsequent gradient of generalization. In these cases, the rate is high in the presence of one of the training stimuli, low in the presence of the other, and intermediate at points in between, and it is easy to slip into the comfortable assumption that nothing much has happened beyond the strengthening of the response at one point and its weakening at another. But when the dimension running from one training stimulus to the other is perpendicular to the dimension along which the test is conducted, so that only one of the training stimuli ever appears in the series used for that test, this illusion is shattered. Although the training schedule may have affected the level of responding in the presence of the one stimulus, and consequently the absolute level of the rest of the gradient, there is no way in which it could have had a direct strengthening and weakening effect on the subject's response to any of the other stimuli in the test series. Differential reinforcement has not been applied to any two points along the test dimension and cannot account in any immediate sense for the difference in rates. Yet this form of training does increase the influence of the test dimension, as indicated by a steepening of the gradient.

Several examples are available. Newman and Baron (1965) trained their pigeons in Group I to discriminate between the presence and absence of a vertical white line on an otherwise green key. Group III received the same stimulus, the white line on the green background, throughout their training, during both periods of extinction and periods when reinforcement was scheduled. For the two groups, everything was the same except the relationship between the presence or absence of the line and the schedule of reinforcement, yet during a generalization test the pigeons in Group I responded at substantially different rates in the presence of lines of different tilt, while the birds in Group III showed no systematic trend.

A similar finding was obtained by Vetter and Hearst (1968), using the rotary position or orientation — equivalent to a tilt — of a parallelogram as their test dimension. Again, the birds that received differential reinforcement with respect to the presence or absence of the parallelogram were sensitive to its orientation during testing; those that did not receive differential training were not.

Switalski, Lyons, and Thomas (1966) used a slightly more complex procedure. All birds were first given 14 sessions of training with the key illuminated by a wavelength of 555 nm (green). Birds in one group were then given differential training. Responding in the presence of 555 nm (S+) was still reinforced on a variable-interval schedule, but responding in the presence of a white line on a dark surround (S−) was not reinforced. The gradient of generalization along the wavelength dimension was steeper, indicating greater control by this aspect of the stimulus, following the differential training. Birds in the other group received nondifferential reinforcement. Although the two stimuli alternated on the key, as for the first group, there was no correlation between stimulus and reinforcement. With these birds, the gradient of generalization was less steep, indicating a reduction in stimulus control resulting from the additional training.

Finally, Lyons and Thomas (1967) carried out a series of reversals, using the subjects that had been given the differential training in the preceding study. First, the birds were given a

number of sessions of nondifferential training, followed by a generalization test, then a number of sessions of differential training followed by a test, and so on, until they had completed six tests in all, three following each type of training. The data from each test were summarized by calculating the percentage of total responding that occurred in the presence of the 555-nm test stimulus, the one that served as the positive stimulus during the differential training. Each of the 9 subjects exhibited the same pattern: a consistent increase following each block of differential training and a consistent decrease following each block of nondifferential training. All 54 determinations supported the conclusion that a series of wavelengths projected on the key exerts more control over the pigeon's behavior when the wavelength used in training has been positively correlated with the receipt of food than when it has been left uncorrelated.

DIFFERENTIAL REINFORCEMENT AND OBSERVING

As I have indicated earlier, this increase in the control exerted by some dimension of a stimulus, following differential reinforcement in its presence and its absence, cannot be explained in terms of the strengthening of the response in the one case and its weakening in the other. The training dimension and the test dimension are orthogonal to each other, and for this reason the old analysis does not apply. However, a reasonable suggestion, it seems to me, is that the subject is more likely to be influenced by variations in some aspect of the stimulus if it is observing that stimulus. In fact, the more accurate the observing and the greater the proportion of time that the subject engages in it, the greater should be the degree of stimulus control. The increase in control produced by orthogonal or interdimensional discrimination training, then, can be explained if it can be shown that the level of observing increases during discrimination training.

But that, of course, is exactly what was demonstrated in the very first study in which an artifical observing response was employed. The observing procedure was originally de-

vised by L. B. Wyckoff, Jr., as the methodological basis for a doctoral dissertation completed in 1951. (The empirical part of the dissertation, however, was not published until 1969.) Noting that a natural observing response, such as looking at the key, would be very difficult to monitor while the pigeon was pecking, Wyckoff (1969) substituted a response that was entirely different in its topography but that was relatively easy to record. He provided his artifical observing response — standing on a pedal situated on the floor of the experimental chamber — with the same functional properties as the natural one. Only when the bird stood on this pedal could it see the discriminative stimuli, red and green illumination of the key, because only when it stood on the pedal did the red and green actually appear on the key. The rest of the time the key was illuminated with white light, regardless of whether or not reinforcement was scheduled. The pedal operated a switch that turned on the appropriate stimulus. And, in keeping with its status as an analogue to looking at the key, that is all that it did. It did not affect the schedule of reinforcement.

For Group I in Wyckoff's study, there was no correlation between red and green and the receipt of food, and the proportion of time the birds spent on the observing pedal declined from its initial level. For Group II-a, red was the positive stimulus for a discrimination and green was the negative stimulus. In this case, the proportion of time spent on the pedal went up. These birds also learned the discrimination. For Group II-b, the discrimination was reversed after five sessions of training. The proportion of time spent on the pedal declined at this point but then recovered as the new correlation took hold. In short, we see that the proportion of time spent observing the stimulus increases under the same conditions as those producing an increase in control — a correlation between the stimuli and food — and decreases under the same conditions as those producing a decrease in stimulus control — absence of such a correlation. (See also, e.g., Bower, McLean, & Meacham, 1966; Brown, 1968; Driscoll, Lanzetta, & McMichael, 1967; Fantino & Case, 1983; Fantino, Case, & Altus,

1983; Kendall, 1973; Lanzetta & Driscoll, 1966; Prokasy, 1956; Wilton & Clements, 1971). This correspondence and others to follow suggest that the increase in observing may account for the increase in stimulus control. The data are correlational, to be sure, but so are the data implicating the smoking of cigarettes in developing lung cancer. As we shall see, after consideration of some other matters, there are several additional lines of evidence that converge on the same conclusion.

AVERSIVE CONTROL

For aversive control the data are meager, but essentially the same analysis may apply as in the case of control by positive reinforcement. Particularly pertinent is a study by Honig (1966) that appears to parallel the previously cited studies of orthogonal discrimination training. First, pigeons were trained to peck at approximately equal rates when a set of parallel lines was projected on the key at various tilts with respect to the vertical-horizontal axes. Equal responding served as the baseline for the determination of a gradient of generalization following discrimination training. Then, in the discrimination group, pecking was punished with brief electric shocks when vertical lines were displayed on the key but was not punished when the key was blank. Therefore, for this group absence of the lines was what is conventionally know as a safety signal. For the no-discrimination group, the same vertical lines were present during the unpunished (safe) periods as during the periods of punishment, so that there was no correlation between the lines and the shocks. For the no-discrimination group and for another, more complex control group, the tilt of the lines had no systematic effect during the subsequent generalization test. On the other hand, the discrimination group yielded a U-shaped gradient, pecking fewer than 40 times per minute in the presence of the vertical line but more than 70 times per minute in the presence of some of the lines of lesser tilt.

There seems to be a widespread impression that experimenal organisms will not, or at least should not, observe stimuli that are positively correlated with the receipt of shock and that for this reason increasing control of behavior by such stimuli cannot be explained in terms of increments in observing. But this was the very problem that originally lured me into research on the observing response. What we found was very simple: Although the stimulus indicating the punishment of instrumental responding did not, by itself, maintain an observing response, the stimulus indicating no punishment did maintain such responding (Dinsmoor, Flint, Smith, & Viemeister, 1969; for reviews of the general area, see Badia, Harsh, & Abbott, 1979; Imada & Nageishi, 1982.) Under a discriminative punishment procedure, then, the observing response is intermittently reinforced by the production of safety signals. It may also be punished by the production of warning signals, but unless the subject is furnished with yet another set of stimuli, correlated with the first, it cannot on some occasions produce the stimuli negatively correlated with shock without also on other occasions producing the stimuli positively correlated with the shock. .

OBSERVING AND ATTENTION

I have discussed the correspondence between the circumstances under which the behavior of observing arises and those under which stimulus control develops without making reference to the inferred process of attention. Much as I would like to continue the analysis without mention of any process that is not directly observed, I find myself persuaded by the arguments put forth by hypothetico-deductive theorists like Mackintosh (1974, 1975) and Sutherland and Mackintosh (1971) to the effect that purely peripheral adjustments cannot account for all of the data. As one example, consider the well known experiment by Jenkins and Harrison (1960). The design was similar to those we have already considered, except that for the control group only one stimulus was used during the training (single stimulus training). When a tone of 1000 Hz was sounded more or less continuously throughout the period when pecking was being acquired, the subject showed no difference in

reaction during the subsequent test periods to this tone and to tones of much higher or much lower frequency. None of the individual gradients of generalization showed systematic deviation from a straight horizontal line. On the other hand, when another group of birds was trained with the presence of the tone as the S+ and its absence as the S−, each of the subjects yielded a steep gradient, indicating a substantial sensitivity to the frequency of the test stimulus. There are two conclusions to be drawn from these data. First, although the subject might be able to increase or decrease its exposure to standing waves by an appropriate adjustment of its position within the chamber, I find it difficult to believe that the magnitude of such an effect, when the tone presumably was not aversive, could account for the magnitude of the difference in behavior between the two groups. This forces me to accept the existence of some process of a more central nature, which might appropriately be called attention. The second conclusion is that this form of attention, at least, seems to arise under the same circumstances as do observing responses. The issue is necessarily somewhat speculative, but I suspect that in general the two processes go hand in hand: If the subject observes a stimulus for the sort of reason we customarily set up in the animal laboratory, I assume that it also attends to that stimulus.

Theorists who derive their explanatory principles exclusively from the data on stimulus control itself (e.g., Honig, 1970; Mackintosh, 1974, 1975; Sutherland & Mackintosh, 1971) tend to downplay or even to disregard the role of observing responses. But I think that the failure to consider this form of behavior is unwise on two counts. First, a systematic science cannot afford to neglect one of its major categories. There is no question but what such responses do occur, that they comprise a distinct functional category, that they have important effects on other behavior, and that they have sufficient practical significance to be worthy of investigation in applied settings (e.g., Rayner, 1978; Thomas, 1968). Second, if the principles governing what the organism observes turn out to be different from those governing attention, then we will need to understand both if we are to make the best possible prediction of what the organism will do in a given situation. On the other hand, if the principles are the same, then the strength of the observing behavior maintained by various stimuli can serve as a concrete, measurable index to the somewhat evanescent construct of attention (e.g., Dinsmoor, Sears, & Dout, 1976; Eimas, 1969a, 1969b; Hamlin, 1975; Singh & Beale, 1978). Such an index would offer a number of advantages, not the least of which is that it would provide a valuable safeguard against purely ad hoc interpretation.

PROGRESSIVE TRAINING

As noted earlier, the original Skinner-Spence analysis, based on the strengthening of the response in the presence of one stimulus and its weakening in the presence of the other, did not deal with the changes in the magnitude of the transfer or induction between the two stimuli and cannot account for the development of stimulus control under orthogonal (interdimensional) discrimination training. Also worthy of note is the fact that the traditional formulation makes no provision for the more rapid acquisition of a discrimination under progressive training, in which the difference between the positive and the negative stimulus is initially made relatively large and is later reduced to its final value (e.g., Lawrence, 1952; Pavlov, 1927/1960, pp. 121-123). In earlier publications from my laboratory (Dinsmoor et al., 1976, 1983; Dinsmoor, Mueller, Martin, & Bowe, 1982), we have reported that more observing occurs when the difference between the two stimuli (their disparity)—or, more importantly, the difference between the positive stimulus and its background stimulation (its salience)—is relatively large than when these differences are small. Both factors appear to be at work in most of the experiments demonstrating the superiority of progressive discrimination training. I have already suggested that the amount of induction between the two stimuli depends upon the level of observing; it follows, therefore, that the discrimination should be

acquired more rapidly when a substantial level of observing is established early in training by exposure to a more effective pair of stimuli before the final pair is encountered.

The effectiveness of progressive training was demonstrated in particularly dramatic form by the work of Terrace (1963a, 1963b) on what he has called errorless discrimination learning. Basically, what Terrace did, with minor variations from one experiment to the next, was to train his "errorless" birds to peck the key when it was illuminated with red light. Then, very early in training, he introduced brief interruptions of the key illumination as the S− and gradually increased their duration and changed their content from a darkening of the key to its illumination with green light. In other words, green was faded in as the S−. Birds trained in this fashion made few, if any errors − that is, they rarely pecked the key at any point during their training except when it was red. This finding was extremely bothersome to theoretical accounts that had emphasized the necessity of extinguishing responding in S− in order to form a discrimination, because in this case there was virtually no such responding to be extinguished. Somehow, the need for extinction seemed to have been bypassed. The discrimination appeared to have been fully formed before the traditional process ever began. However, errorless training does not present the same difficulty to an analysis that suggests that the formation of a discrimination depends, first of all, on learning to observe and/or attend to the stimuli.

Terrace's procedure also included special features that I think facilitated the rapid formation of a discrimination. For example, throughout all of the relevant work, red and green illumination of the pigeon's key were used as the discriminative stimuli. Although wavelength of key illumination is widely used in studies of stimulus control, it is nevertheless a very special dimension (see Richardson & Evans, 1975). Control is readily and rapidly established. In my laboratory, while preparing our subjects for a discrimination based on punishment (Dinsmoor et al., 1969), Richard Smith and I tried to get birds to peck the key both in the presence of red and in the presence of green. We often got errorless learning in spite of ourselves. After pecking had been shaped in the presence of red, it failed to transfer to the presence of green. Each time the color changed on the key, the bird stopped pecking, and this difficulty often persisted for hours of session time. We could not reinforce pecking in the presence of green because no such pecks occurred. Sometimes we had to "trick" our subject by timing the switch in colors so that it came in the midst of a rapid burst of responding and then reinforcing a peck that landed before the bird reacted to the change.

Note two things about the color on the key. First, it is displayed on the surface of the device the pigeon pecks. I would assume that the subject frequently glances at the key before striking it. Second, the color is not restricted to some localized portion of the key but covers its entire surface. No accommodation is required, and no matter where the pigeon looks, if it looks at any part of the key it will see the color. This is not true of a visual pattern, for example, and Terrace noted (1963b, p. 229) that the procedure described above did not work when two lines of different tilt were used as the discriminative stimuli. Prior training with the colors was needed. What I think happens here is that the pigeon frequently looks at some spot on the key before pecking it. It sees the color. Since the food is delivered as a reinforcer only following pecks on the key, it is selectively associated with looking at the key and seeing the color as well. (Here my argument follows those advanced by Heinemann & Rudolph, 1963, and Richardson & Evans, 1975.) Therefore, even single stimulus training with color on the key may produce an increase in observing. Subsequent training in which reinforcement accompanies the color red and nonreinforcement a dark key would be expected to add further strength to the observing response, so that by the time the color green is introduced as an S− the performances in S+ and S− could be expected to be quite independent. Pecking is never strengthened in the presence of green illumination of the key, either by direct reinforcement or by induction from reinforcement in the presence of red.

Other factors that may have played signifi-

cant roles are superstitious reinforcement of interfering behavior by the early appearance of S+, as noted by Terrace (1963a, p. 26), and selective nonreinforcement, by delaying the appearance of S+, following any pecks that did occur in S− (Dinsmoor, 1950, 1951).

Let me add one caveat, however. Nothing that I have said in attempting to explain the efficacy of Terrace's procedure should be construed as an endorsement of the suggestion, first advanced by Lashley and Wade (1946) and later discussed at some length in Terrace's review (1966), that no stimulus control can be found without prior differential reinforcement. What I have said need not imply such a position. It seems reasonable that the subject might occasionally observe the relevant stimulus before selective reinforcement has occurred. In fact, it is difficult to see how a correlation between stimulus and reinforcer could be effective if the subject did not sometimes make contact with the stimulus during the training.

CONTROL BY THE NEGATIVE STIMULUS

In all of the studies of orthogonal (interdimensional) discrimination training that I have so far cited, it has been the presence of the line or other stimulus event that has served as the S+ and the absence that has served as the S−. It is entirely feasible to reverse this relationship, however, and to use the presence of the stimulus event as the S− and its absence as the S+. Then the S− can be varied in some way during subsequent testing and a gradient can be obtained that shows the rate of responding as a function of similarity to that stimulus. If absence of a line has been used as the S+, for example, variations in the rate of responding to lines of different tilt cannot readily be attributed to variations in their similarity to S+ but must be a function of their similarity to S−. The resulting gradient shows how much control is exerted by this characteristic of the negative stimulus. Attempts have also been made to determine the relative contributions of S+ and S− to the control of behavior in simultaneous (choice) discriminations, but I find that situation too complex, the deductive

chains by which the conclusions are reached too long, and the results too variable to place much reliance on their findings. I greatly prefer the comparison of gradients of generalization following the two forms of orthogonal discrimination training.

In their review of research on stimulus generalization, Honig and Urcuioli (1981) concluded that "the balance of the literature suggests that negative gradients are . . . flatter than positive gradients" (p. 422). I agree with this assessment, but I think that the most convincing evidence comes from comparisons in which S+ and S− gradients were obtained under closely matched conditions. I know of five such studies. First, Jenkins and Harrison (1962) obtained negative gradients following training in which the positive stimulus was either absence of tone (Experiment I) or white noise (Experiment III). In both cases, the S− was a tone of 1000 Hz. The stimuli used in Experiment I were the same, although reversed in assignment, as those used in an earlier study that I have already discussed (Jenkins & Harrison, 1960). The rate of pecking by each of 6 birds was then tested in the presence of each of seven auditory frequencies. In the study in which the positive gradients were obtained, very little pecking (typically about 5%, equivalent to that in S−) occurred in the presence of each of three stimuli toward the extremes of the continuum, but something like 25 to 45% (varying for individual birds) occurred in the presence of S+. In the study in which the negative gradients were obtained, the highest proportion of responding by any bird in the presence of any one of the tones was about 20%. In general, the negative gradients appeared much shallower than the positive gradients, indicating substantially less control of responding by auditory frequency when this was a dimension of the negative stimulus than when it was a dimension of the positive stimulus.

A similar result was obtained by Marsh (1972), using as his test dimension the wavelength displayed on the key. When a stimulus of 550 nm was used as the S+ in training, the subsequent gradient of generalization showed a peak at that point but dropped to very low

values toward either end. The slope was steep, indicating a substantial degree of control by the wavelength dimension. On the other hand, when a wavelength of 560 nm was used as the S−, the gradient was relatively shallow. It dipped in the middle but never reached as low a point as did the gradient for S+, and it rose at the sides, but never reached the same height as did the gradient for S+.

Two more sets of gradients, this time based on the tilt of a black line on a white surround, were obtained in successive replications with two different sets of birds by Honig et al. (1963). Again, both of the gradients for S+ were steeper than and crossed over the corresponding gradients for S−. Similar gradients were also obtained by Hearst (1968), with different tilts used for S− in each of three different groups. In this work, the gradients did not cross, but all three gradients for S− were substantially flatter than the corresponding gradients for S+. Boneau and Honig (1964) obtained both positive and negative line-tilt gradients from the same individuals, following training with a conditional discrimination procedure. When the background was 550 nm, a vertical line was S+ and no line was S−, but when the background was 570 nm, the vertical line was S− and no line was S+. By the third day of testing, at least, the proportion of total test responding in the presence of various tilts yielded a reasonably peaked gradient with the 550-nm background but still yielded a completely flat gradient with the 570-nm background.

On the face of it, the data from all five studies support the conclusion that the negative stimulus exerts less control over the subject's responding following successive discrimination training than does the positive stimulus. However, several writers either have offered alternative interpretations of this type of datum or have questioned the legitimacy of the measures that have been employed. For example, in their study, Honig et al. (1963) maintained that the true or ideal positive and negative gradients were really the same in shape but that the observed gradient produced variations in the negative stimulus suffered from a "floor effect." (This conclusion was subsequently rejected by Honig and Urcuioli, 1981, but in the meantime has been accepted by a variety of other investigators.) As I understand their argument, it is that if the rate had been higher at points distant from the negative stimulus along the test dimension, it would still have been close to zero in the presence of the negative stimulus, and the gradient would have been steeper. But in response to that argument, one has to ask why the rate was not, in fact, much higher in the presence of the other stimuli than it was in the presence of S−. Evidently the subjects were sensitive to the fact that the test stimulus was not S+, but were not sensitive to its difference from S−.

Hearst, Besley, and Farthing (1970, pp. 392 ff.) have suggested that steeper gradients might be obtained with S− if responding were intermittently reinforced during the test, to maintain higher absolute levels. Indeed, under certain conditions fairly steep gradients can be obtained using this technique (e.g., Karpicke & Hearst, 1975); but under other conditions they remain relatively flat (e.g., Rilling, Caplan, Howard, &. Brown, 1975). Much may depend on the stimulus dimension and the length of the testing period, but to the best of my knowledge no controlled comparison of S+ and S− gradients is available in which the resistance-to-reinforcement technique has been employed.

Other writers have raised questions concerning the proper measure to use when comparing stimulus control gradients that differ in their absolute levels. For example, Lea and Morgan (1972) have suggested that the slopes of such gradients should be measured in terms of differences in absolute rates at different points along the continuum, rather than, as is more customary, in terms of relative rates (ratios). Blough (1965) has questioned whether any measure derived from rate of responding can be adequate. He has also suggested that "the right way to measure responses is that way which shows invariances most clearly" (1965, p. 32). Elsewhere (Dinsmoor, 1973b), I have noted that one measure, relative rates, often does remain invariant under a variety of parametric settings, suggesting that this measure may have a high degree of generality

and may represent the best we have for standard use. When changes in relative rate do occur, they are likely to be meaningful, rather than artifacts of measurement (see also Anderson, 1978, p. 369). In any case, when the gradient of generalization for S+ extends from an absolute value lower than any obtained with S− to an absolute value higher than any obtained with S−, encompassing the entire range for the latter stimulus, it is difficult to believe that the difference in slopes can be attributed to some distortion resulting from the technique used for their measurement. It seems to me, then, that gradients of generalization following orthogonal discrimination training do indicate that the negative stimulus exerts less control over the subject's behavior than does the positive stimulus.

SELECTIVE OBSERVING

A simple explanation for the relatively poor control exhibited by S− following presence versus absence training is that during the training and presumably during the test sessions, the subject spends less of the available time observing S− than it does observing S+. We discovered this fact for the artifical observing response, at least, some years ago. Dinsmoor, Browne, Lawrence, & Wasserman (1971) conducted an experiment that replicated Wyckoff's original procedure with one important difference: We kept separate records of the times when S+ was present and the times when S− was present. We found that when stepping on the pedal produced green illumination of the key, S+, the bird typically stayed on the pedal most of the time until the programming circuit switched to nonreinforcement and illumination of the key switched from green to red. The mean durations for 3 birds were 21.32 s, 25.57 s, and 50.89 s. The typical performance with red, however, was to step on the pedal for a few seconds, step off it for a similar period of time, step on it, step off it, and so on, until green reappeared. The mean durations for each exposure to S− were 5.58 s, 6.12 s, and 5.58 s.

It is difficult to be certain that any specific form of response that might be used to represent observing is entirely free of possible facilitation by or interference from other responses, such as the response that produces the deliveries of food (for discussion, see Browne & Dinsmoor, 1974; Hirota, 1972; Kelleher, Riddle, & Cook, 1962; Kendall, 1965a, 1965b), interim and terminal behavior (see Green & Rachlin, 1977), and approaching and eating the grain (see Karpicke, 1978). Furthermore, these forms of behavior occur with different frequencies during S− and S+. Accordingly, to make the case for selective observing as secure as possible, we have subsequently employed a variety of response topographies as indicators to this function. Browne and Dinsmoor (1974) used standing in the right half (or subsequently, standing in the left half) of the experimental chamber, with food delivered on a response-independent schedule. Dinsmoor et al. (1982) used holding down a crossbar mounted on the wall of the chamber. Another crossbar mounted on the same wall, to the other side of the food-delivery mechanism, provided concurrent control data. And in some work yet to be published, we have used a peck on a key to the left of the food hopper to turn on a discriminative stimulus and a peck on a key to the right to turn one off. To prevent approaches to or withdrawals from S+ or S− from influencing our data, we used diffuse illumination of the entire chamber by overhead lamps differing either in their intensity or in their spectral distribution (color). With each of these varied arrangements, we again obtained selective observing. The temporal pattern may be illustrated by a sample, covering approximately the last quarter of a session, taken from an operations record late in training (see Figure 1). Time flows from left to right, and successive segments are arranged from top to bottom. Pecking the left key turned off the mixed-schedule stimulus and turned on either the positive stimulus, as in the first instance on the record, or the negative stimulus. Thus, pecking this key was sometimes reinforced and sometimes punished. Note that S−, which presumably was punishing, was produced many more times than S+, which presumably provided the reinforcement. Yet the behavior was maintained indefinitely.
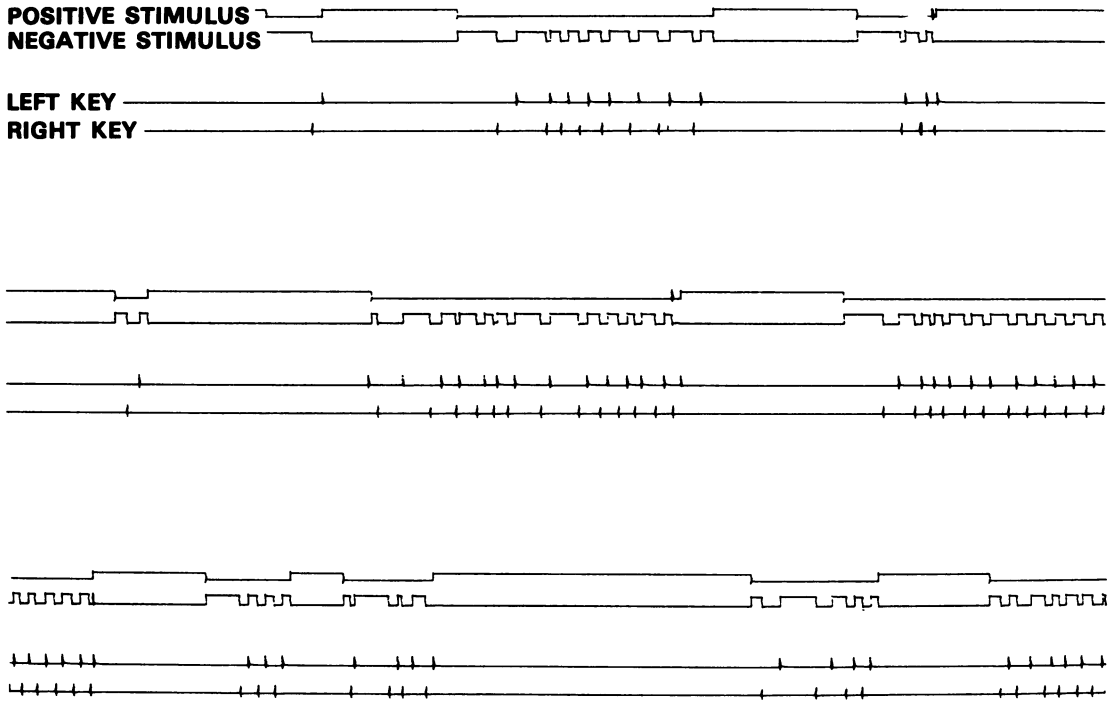
Fig. 1.   Last 15 min of an operations record late in training. Time flows from left to right. Upward displacement of the top line shows when the stimulus accompanied by a variable-interval schedule of reinforcement is present. Upward displacement of the second line shows when the stimulus accompanied by no reinforcement is present. When neither line is displaced, the mixed-schedule stimulus is present. Blips on the third line show pecks on the left (on) key, and blips on the fourth line show pecks on the right (off) key.

(For an extended discussion of how observing is maintained, see Dinsmoor, 1983.) Pecks on the right key turned off S+ or S−, whichever was present, and turned on the mixed stimulus. Although the experimental arrangements do not provide discriminative stimuli for pecking the left key, S+ and S− themselves serve that function for pecking the right key. Note that S− was usually turned off within a few seconds after it appeared but that there are only two instances on the record in which the S+ was turned off. In general, termination of S+ was an extremely rare event, and in both of the instances shown, it came about within a fraction of a second following a switch by the programming apparatus from the negative stimulus to the positive one. Perhaps the bird failed to react in time to the change in circumstance. In accord, then, with simple, straightforward behavioral principles, the subject selects relatively large amounts of exposure to the positive stimulus and relatively small amounts of exposure to the negative

stimulus. This difference in the relative proportion of the time during which a stimulus is available that the subject spends observing it may account for the greater degree of control exerted by the positive stimulus when we compare gradients of generalization for S+ and S− following presence versus absence training.

As my analysis up to this point has been based wholly on data we have obtained with artificial observing responses, perhaps it would be wise to look for some confirmation of the assumption that the same pattern of events holds during more conventional laboratory procedures, in which only natural observing responses are required to make contact with the stimuli. Rand (1977) monitored her birds visually during the formation of a discrimination between a horizontal and a vertical line displayed on the key. She recorded, among others, "a class of behaviors that served to remove the visual stimulus from view" (p. 106). With each of her subjects, she noted a substantial increase in the proportion of time

spent in that behavior during periods of S− presentation as the experiment progressed. In other words, her subjects learned to turn away from the key when S− was displayed. The same kind of relationship held during generalization testing. The more a given stimulus resembled S−, the greater was the proportion of time the bird spent on this stimulus-avoiding class of behavior. It may be objected that the very response of turning away itself indicated control by the stimulus displayed on the key. But such control need be assumed only as a momentary condition, required to account for the initiation of the response of turning away. Turning away is comparable to stepping off the floor pedal or crossbar or pecking the stimulus-terminating key in a study using artifical observing responses. Later, the bird may return for another look, as it does in the artificial case. Note that I have not argued that there is no control by the negative stimulus in this instance, only that it is reduced. I have suggested that the subject makes less contact with S−, when it is present, than it does with S+, and that this may be the reason why gradients of generalization around S− are flatter than those around S+.

FEATURE-POSITIVE EFFECT

Poor observing of S− may also explain a curious phenomenon that has captured the interest of a number of investigators in animal discrimination learning. Although similar effects have been obtained with human subjects (Newman, Wolff, & Hearst, 1980), the bulk of the work has been conducted with pigeons. The feature-positive effect is obtained under special circumstances, to be characterized below, but the basic phenomenon goes as follows: When the stimulus element or feature that distinguishes the positive trials from the negative trials is programmed to occur on the positive trials, pigeons may learn a discrete-trial or an autoshaping discrimination quite readily. But when it is on the negative trials that the feature appears, they may fail to learn the same discrimination. (Although the amount of information presented is logically equivalent in the two cases, this metric ap-

pears to be of no value in predicting the subject's behavior.) The phenomenon is most striking when the distinguishing feature is a small visual detail (Morris, 1977), low in contrast (W. T. Wolff, personal communication, May 1977), which shifts substantially in its locus from trial to trial (Sainsbury, 1971). In other words, the phenomenon is most likely to appear when the critical stimulus element is one that the bird may not see unless it looks directly at it. Athough the visual angle subtended by the typical 2.5-cm key opening may be quite small for the human observer, looking from a distance of, say, 30 cm or more, it is quite large when the same opening is inspected from a distance of perhaps 1 or 2 cm. Consequently, when the pigeon is pecking the key, the individual elements displayed on that key may be some distance apart in the bird's visual field and may more appropriately be treated as independent stimuli, separately seen, than as components of a single pattern. The results reported by Hearst (1975) are consistent with this interpretation, in that he obtained the same effect even when the stimuli were quite salient but were displayed still farther apart, on three separate keys.

In addition to the feature, there typically are other stimuli, similar in kind, that appear on every trial. These stimuli serve to distinguish the period when the trial is in effect from the interval between trials. As half of the trials are positive and half are negative, these common stimuli correctly predict reinforcement 50% of the time. If we assume on the basis of the data from experiments on observing that the pigeon learns to look at the stimulus predicting the highest density of reinforcement (Auge, 1974; Branch, 1970; Jwaideh & Mulvaney, 1976; Kendall, 1973, Experiment II; Kendall & Gibson, 1965; Ward, 1971), we can predict that this will be the positive stimulus, where there is such a feature to look at, but that it will be a common, or trial stimulus, when the distinguishing feature appears only on the negative trials. Such common stimuli are not requisite for a small feature-positive effect to emerge, but they do substantially increase its magnitude (Wolff, 1983). Presumably, they serve as distractors, controlling the pigeon's

attention during the trial periods, when the discrimination is being measured.

Empirical evidence supporting this analysis is provided by records of where the pigeon pecks. Jenkins and Sainsbury (1969) divided their key into four independent sectors or quadrants, each of which could be pecked separately (see also Farthing, 1971; Hearst, 1975; Jenkins & Sainsbury, 1970; Sainsbury, 1971; Wasserman & Anderson, 1974). On any given trial, one sector was left blank, but either a circular spot, 0.3 cm in diameter, or a star of the same size was projected on each of the other three. One of these figures served as the distinguishing feature and was present on one sector or another only during positive trials or only during negative trials, depending on the procedure to which the particular bird was assigned. The other figure was displayed on the two remaining quadrants, as the common or trial stimulus. Jenkins and Sainsbury found that birds that received the feature on positive trials pecked mainly on the sector bearing that feature, but birds that received the feature on negative trials pecked mainly on one of the sectors bearing the common, or trial stimuli. If we can assume, as is indicated by the auto-shaping literature (Hearst & Jenkins, 1974; Schwartz & Gamzu, 1977), that the conditions under which the bird will peck a key are very similar to the conditions under which it will approach and look at that key and the conditions under which it will produce ("observe") stimuli displayed on that key, then the Jenkins and Sainsbury data support the observing analysis. (For relatively direct evidence that looking and pecking are intimately related, see Jenkins, 1973; Maki, Gillund, Hauge, & Siders, 1977; Skinner, 1965.)

Recognition of the power of an analysis based on observing may have been delayed by the objections that Jenkins and Sainsbury (1969, 1970) raised to related lines of thinking in their initial presentations of the feature-positive effect. In their 1969 study, the authors attempted to formulate a version of the prevailing hypothetico-deductive theories of attention that might account for their results. But they concluded that the models then current, which postulated attention to an entire dimension rather than to the concrete stimuli, could not cope with the asymmetry between the effects produced by the presence and the effects produced by the absence of a particular stimulus. Attention to an entire dimension is not, of course, assumed in the present approach, which is based on observation of or attention to a specific stimulus and does not require or even suggest a higher level of abstraction.

In their 1970 study, Jenkins and Sainsbury considered and rejected what they called a "search theory," essentially a chaining analysis in which the behavior of searching for the distinguishing element was strengthened under the feature-positive procedure but not under the feature-negative procedure. Again, the present analysis is different; it does not attribute the success or failure of the training procedure to the presence or absence of searching but to the outcome of the search—whether the subject ends up observing the feature or some other aspect of the stimulus display. But the present interpretation, too, might seem to be precluded by the objection that Jenkins and Sainsbury raised to their own form of search theory. After all, these authors argued, in their initial selection of which sector of the key to peck, the birds must have seen the feature in order to avoid it so consistently. The paradox is resolved, however, if we recognize that the seemingly contradictory performances occurred during different segments of time. At the beginning of a trial the birds may often have looked momentarily at more than one sector of the key, perhaps including the one bearing the negative feature, before settling on one of those bearing the common element. But after that point, during the remainder of the trial, there is no reason to believe that they again looked at the distinguishing feature. To judge from matching-to-sample studies (Grant, 1976; Maki & Leith, 1973; Roberts & Grant, 1974), a brief exposure to the stimulus is not nearly as effective as more prolonged contact in controlling the pigeon's subsequent behavior. During the time that the discriminative performance was being measured, the birds were pecking one of the sectors bearing the common element, and it seems reasonable to conclude that this was the element they

were looking at and that the element they were looking at was the one that controlled their behavior.

REFERENCES

Anderson, N. H. (1978). Measurement of motivation and incentive. *Behavior Research Methods & Instrumentation*, **10**, 360-375.

Auge, R. J. (1974). Context, observing behavior, and conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, **22**, 525-533.

Badia, P., Harsh, J., & Abbott, B. (1979). Choosing between predictable and unpredictable shock conditions: Data and theory. *Psychological Bulletin*, **86**, 1107-1131.

Blough, D. S. (1965). Definition and measurement in generalization research. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 30-37). Stanford: Stanford University Press.

Boneau, C. A., & Honig, W. K. (1964). Opposed generalization gradients based upon conditional discrimination training. *Journal of Experimental Psychology*, **68**, 89-93.

Bower, G., McLean, J., & Meacham, J. (1966). Value of knowing when reinforcement is due. *Journal of Comparative and Physiological Psychology*, **62**, 184-192.

Branch, M. N. (1970). The distribution of observing responses during two VI schedules. *Psychonomic Science*, **20**, 5-6.

Brown, P. L. (1968). Auto-shaping and observing responses (R$_o$) in the pigeon [Summary]. *Proceedings of the 76th Annual Convention of the American Psychological Association*, **3**, 139-140.

Browne, M. P., & Dinsmoor, J. A. (1974). Wyckoff's observing response: Pigeons learn to observe stimuli for free food but not stimuli for extinction. *Learning and Motivation*, **5**, 165-173.

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

DePaulo, P., DeWald, L., & Yarczower, M. (1977). General attentiveness effects of discriminative training. *Journal of the Experimental Analysis of Behavior*, **27**, 245-253.

Dinsmoor, J. A. (1950). A quantitative comparison of the discriminative and reinforcing functions of a stimulus. *Journal of Experimental Psychology*, **40**, 458-472.

Dinsmoor, J. A. (1951). The effect of periodic reinforcement of bar-pressing in the presence of a discriminative stimulus. *Journal of Comparative and Physiological Psychology*, **44**, 354-361.

Dinsmoor, J. A. (1973a). A theory of discrimination learning [Summary]. *Proceedings of the 81st Annual Convention of the American Psychological Association*, **8**, 861-862.

Dinsmoor, J. A. (1973b). Operant conditioning. In B. B. Wolman (Ed.), *Handbook of general psychology* (pp. 501-514). Englewood Cliffs, NJ: Prentice-Hall.

Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences*, **6**, 693-728. (Includes commentary)

Dinsmoor, J. A., Bowe, C. A., Dout, D. L., Martin, L. T., Mueller, K. L., & Workman, J. D. (1983). Separating the effects of salience and disparity on the rate of observing. *Journal of the Experimental Analsyis of Behavior*, **40**, 253-264.

Dinsmoor, J. A., Browne, M. P., Lawrence, C. E., & Wasserman, E. A. (1971). A new analysis of Wyckoff's observing response [Summary]. *Proceedings of the 79th Annual Convention of the American Psychological Association*, **6**, 679-680.

Dinsmoor, J. A., Flint, G. A., Smith, R. F., & Viemeister, N. F. (1969). Differential reinforcing effects of stimuli associated with the presence or absence of a schedule of punishment. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 357-384). Homewood, IL: Dorsey Press.

Dinsmoor, J. A., Mueller, K. L., Martin, L. T., & Bowe, C. A. (1982). The acquisition of observing. *Journal of the Experimental Analysis of Behavior*, **38**, 249-263.

Dinsmoor, J. A., Mulvaney, D. E., & Jwaideh, A. R. (1981). Conditioned reinforcement as a function of duration of stimulus. *Journal of the Experimental Analysis of Behavior*, **36**, 41-49.

Dinsmoor, J. A., Sears, G. W., & Dout, D. L. (1976). Observing as a function of stimulus difference. *Journal of Experimental Psychology: Animal Behavior Processes*, **2**, 154-162.

Driscoll, J. M., Lanzetta, J. T., & McMichael, J. S. (1967). Preference for information under varying conditions of outcome uncertainty, intensity, and delay. *Psychological Reports*, **21**, 473-479.

Eimas, P. D. (1969a). Multiple-cue discrimination learning in children. *Psychological Record*, **19**, 417-424.

Eimas, P. D. (1969b). Observing responses, attention, and the overtraining reversal effect. *Journal of Experimental Psychology*, **82**, 499-502.

Fantino, E., & Case, D. A. (1983). Human observing: Maintained by stimuli correlated with reinforcement but not extinction. *Journal of the Experimental Analysis of Behavior*, **40**, 193-210.

Fantino, E., Case, D. A., & Altus, D. (1983). Observing reward-informative and -uninformative stimuli by normal children of different ages. *Journal of Experimental Child Psychology*, **36**, 437-452.

Farthing, G. W. (1971). Discrimination of compound stimuli involving the presence or absence of a distinctive visual feature. *Journal of the Experimental Analysis of Behavior*, **16**, 327-336.

Ferster, C. B. (1951). The effect on extinction responding of stimuli continuously present during conditioning. *Journal of Experimental Psychology*, **42**, 443-449.

Grant, D. S. (1976). Effect of sample presentation time on long-delay matching in the pigeon. *Learning and Motivation*, **7**, 580-590.

Green, L., & Rachlin, H. (1977). Pigeons' preferences for stimulus information: Effects of amount of information. *Journal of the Experimental Analysis of Behavior*, **27**, 255-263.

Hamlin, P. H. (1975). Observing responses as an index of attention in chickens. *Journal of Experimental Psychology: Animal Behavior Processes*, **1**, 221-234.

Hearst, E. (1968). Discrimination learning as the summation of excitation and inhibition. *Science*, **162**, 1303-1306.

Hearst, E. (1975). Pavlovian conditioning and directed movements. In G. H. Bower (Ed.), *The*

psychology of learning and motivation (Vol. 9, pp. 215-262). New York: Academic Press.

Hearst, E., Besley, S., & Farthing, G. W. (1970). Inhibition and the stimulus control of operant behavior. *Journal of the Experimental Analysis of Behavior, 14*, 373-409.

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

Heinemann, E. G., & Rudolph, R. L. (1963). The effect of discriminative training on the gradient of stimulus-generalization. *American Journal of Psychology, 76*, 653-658.

Hirota, T. T. (1972). The Wyckoff observing response—a reappraisal. *Journal of the Experimental Analysis of Behavior, 18*, 263-276.

Honig, W. K. (1966). The role of discrimination training in the generalization of punishment. *Journal of the Experimental Analysis of Behavior, 9*, 377-384.

Honig, W. K. (1970). Attention and the modulation of stimulus control. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis* (pp. 193-238). New York: Appleton-Century-Crofts.

Honig, W. K., Boneau, C. A., Burstein, K. R., & Pennypacker, H. S. (1963). Positive and negative generalization gradients obtained after equivalent training conditions. *Journal of Comparative and Physiological Psychology, 56*, 111-116.

Honig, W. K., & Urcuioli, P. J. (1981). The legacy of Guttman and Kalish (1956): 25 years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior, 36*, 405-445.

Hull, C. L. (1950). Simple qualitative discrimination learning. *Psychological Review, 57*, 303-313.

Imada, H., & Nageishi, Y. (1982). The concept of uncertainty in animal experiments using aversive stimulation. *Psychological Bulletin, 91*, 573-588.

Jenkins, H. M. (1973). Noticing and responding in a discrimination based on a distinguishing element. *Learning and Motivation, 4*, 115-137.

Jenkins, H. M., & Harrison, R. H. (1960). Effect of discrimination training on auditory generalization. *Journal of Experimental Psychology, 59*, 246-253.

Jenkins, H. M., & Harrison, R. H. (1962). Generalization gradients of inhibition following auditory discrimination learning. *Journal of the Experimental Analysis of Behavior, 5*, 435-441.

Jenkins, H. M., & Sainsbury, R. S. (1969). The development of stimulus control through differential reinforcement. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 123-161). Halifax: Dalhousie University Press.

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

Jwaideh, A. R., & Mulvaney, D. E. (1976). Punishment of observing by a stimulus associated with the lower of two reinforcement densities. *Learning and Motivation, 7*, 211-222.

Karpicke, J. (1978). Directed approach responses and positive conditioned suppression in the rat. *Animal Learning & Behavior, 6*, 216-224.

Karpicke, J., & Hearst, E. (1975). Inhibitory control

and errorless discrimination learning. *Journal of the Experimental Analysis of Behavior, 23*, 159-166.

Kelleher, R. T., Riddle, W. C., & Cook, L. (1962). Observing responses in pigeons. *Journal of the Experimental Analysis of Behavior, 5*, 3-13.

Kendall, S. B. (1965a). Competing behavior and observing responses. *Psychonomic Science, 3*, 279-280.

Kendall, S. B. (1965b). The distribution of observing responses in a mixed FI-FR schedule. *Journal of the Experimental Analysis of Behavior, 8*, 305-312.

Kendall, S. B. (1973). Effects of two procedures for varying information transmission on observing responses. *Journal of the Experimental Analysis of Behavior, 20*, 73-83.

Kendall, S. B., & Gibson, D. A. (1965). Effects of discriminative stimulus removal on observing behavior. *Psychological Record, 15*, 545-551.

Lanzetta, J. T., & Driscoll, J. M. (1966). Preference for information about an uncertain but unavoidable outcome. *Journal of Personality and Social Psychology, 3*, 96-102.

Lashley, K. S., & Wade, M. (1946). The Pavlovian theory of generalization. *Psychological Review, 53*, 72-87.

Lawrence, D. H. (1952). The transfer of a discrimination along a continuum. *Journal of Comparative and Physiological Psychology, 45*, 511-516.

Lea, S. E. G., & Morgan, M. J. (1972). The measurement of rate-dependent changes in responding. In R. M. Gilbert & J. R. Millenson (Eds.), *Reinforcement: Behavioral analyses* (pp. 129-145). New York: Academic Press..

Lyons, J., & Thomas, D. R. (1967). Effects of interdimensional training on stimulus generalization: II. Within-subjects design. *Journal of Experimental Psychology, 75*, 572-574.

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

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

Maki, W. S., Jr., Gillund, G., Hauge, G., & Siders, W. A. (1977). Matching to sample after extinction of observing responses. *Journal of Experimental Psychology: Animal Behavior Processes, 3*, 285-296.

Maki, W. S., Jr., & Leith, C. R. (1973). Shared attention in pigeons. *Journal of the Experimental Analysis of Behavior, 19*, 345-349.

Marsh, G. (1972). Prediction of the peak shift in pigeons from gradients of excitation and inhibition. *Journal of Comparative and Physiological Psychology, 81*, 262-266.

Morris, R. C. (1977). Spatial variables and the feature-positive effect. *Learning and Motivation, 8*, 194-212.

Newman, F. L., & Baron, M. R. (1965). Stimulus generalization along the dimension of angularity. *Journal of Comparative and Physiological Psychology, 60*, 59-63.

Newman, J., Wolff, W. T., & Hearst, E. (1980). The feature-positive effect in adult human subjects. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 630-650.

Pavlov, I. P. (1960). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*

(G. V. Anrep, Trans.). New York: Dover. (Original work published 1927)

Prokasy, W. F., Jr. (1956). The acquisition of observing responses in the absence of differential external reinforcement. *Journal of Comparative and Physiological Psychology*, **49**, 131-134.

Rand, J. F. (1977). Behaviors observed during S− in a simple discrimination learning task. *Journal of the Experimental Analysis of Behavior*, **27**, 103-117.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin*, **85**, 618-660.

Restle, F. (1955). A theory of discrimination learning. *Psychological Review*, **62**, 11-19.

Richardson, W. K., & Evans, M. S. (1975). The effect of localizing the stimulus at the response key on generalization along the hue dimension in the pigeon. *Animal Learning & Behavior*, **3**, 119-122.

Rilling, M., Caplan, H. J., Howard, R. C., & Brown, C. H. (1975). Inhibitory stimulus control following errorless discrimination learning. *Journal of the Experimental Analysis of Behavior*, **24**, 121-133.

Roberts, W. A., & Grant, D. S. (1974). Short-term memory in the pigeon with presentation times precisely controlled. *Learning and Motivation*, **5**, 393-408.

Sainsbury, R. S. (1971). Effect of proximity of elements on the feature-positive effect. *Journal of the Experimental Analysis of Behavior*, **16**, 315-325.

Schwartz, B., & Gamzu, E. (1977). Pavlovian control of operant behavior. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 53-97). Englewood Cliffs, NJ: Prentice-Hall.

Sidman, M. (1969). Generalization gradients and stimulus control in delayed matching-to-sample. *Journal of the Experimental Analysis of Behavior*, **12**, 745-757.

Singh, N. N., & Beale, I. L. (1978). Attentional changes during discrimination learning by retarded children. *Journal of the Experimental Analysis of Behavior*, **29**, 527-533.

Skinner, B. F. (1933). The rate of establishment of a discrimination. *Journal of General Psychology*, **9**, 302-350.

Skinner, B. F. (1934). A discrimination without previous conditioning. *Proceedings of the National Academy of Sciences*, **20**, 532-536.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1965). Stimulus generalization in an operant: A historical note. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 193-209). Stanford, CA: Stanford University Press.

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

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

Spence, K. W. (1940). Continuous versus noncontinuous interpretations of discrimination learning. *Psychological Review*, **47**, 271-288.

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

Switalski, R. W., Lyons, J., & Thomas, D. R. (1966). Effects of interdimensional training on stimulus generalization. *Journal of Experimental Psychology*, **72**, 661-666.

Terrace, H. S. (1963a). Discrimination learning with and without "errors." *Journal of the Experimental Analysis of Behavior*, **6**, 1-27.

Terrace, H. S. (1963b). Errorless transfer of a discrimination across two continua. *Journal of the Experimental Analysis of Behavior*, **6**, 223-232.

Terrace, H. S. (1966). Stimulus control. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 271-344). New York: Appleton-Century-Crofts.

Thomas, E. L. (1968, August). Movements of the eye. *Scientific American*, **219**(2), 88-95.

Vetter, G. H., & Hearst, E. (1968). Generalization and discrimination of shape orientation in the pigeon. *Journal of the Experimental Analysis of Behavior*, **11**, 753-765.

Ward, E. F. (1971). Acquisition and extinction of the observing response as a function of stimulus predictive validity. *Psychonomic Science*, **24**, 139-141.

Wasserman, E. A., & Anderson, P. A. (1974). Differential autoshaping to common and distinctive elements of positive and negative discriminative stimuli. *Journal of the Experimental Analysis of Behavior*, **22**, 491-496.

Wilton, R. N., & Clements, R. O. (1971). Observing responses and informative stimuli. *Journal of the Experimental Analysis of Behavior*, **15**, 199-204.

Wolff, W. T. (1983). *Discrimination learning based on the presence and absence of predictive stimuli: An analysis of the feature positive effect.* Unpublished doctoral dissertation, Indiana University.

Wyckoff, L. B., Jr. (1952). The role of observing responses in discrimination learning: Part I. *Psychological Review*, **59**, 431-442.

Wyckoff, L. B., Jr. (1969). The role of observing responses in discrimination learning. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 237-260). Homewood, IL: Dorsey Press.