ON THE DEVELOPMENT OF STIMULUS CONTROL¹

ERIC G. HEINEMANN AND EDWARD AVIN

BROOKLYN COLLEGE OF THE CITY UNIVERSITY OF NEW YORK

Pigeons were trained to peck one key when presented with white noise at any of five intensities lower than a reference intensity, and to peck another key when presented with white noise at any of five intensities greater than the reference intensity. The shape of the stimulus control curves (proportion of responses to one key *versus* stimulus intensity) changed from a horizontal line at the beginning of training to the sigmoid form of typical psychometric functions at the end of training. The development of stimulus control is described in terms of a model based on the theory of signal recognition and a concept of attention.

Several recent experiments have shown that choice "generalization" or "stimulus control" curves for intensity of white noise, obtained after training to discriminate between two intensities, have the form of normal ogives or typical psychometric functions (Heinemann, Avin, Sullivan, and Chase, 1969; Heinemann and Chase, 1970a). Choice curves of this form have also been obtained for the continuum of luminance (Heinemann and Chase, 1970b, Chase and Heinemann, 1972). These curves are strikingly different from most other published generalization gradients in that they do not have peaks; rather, the choice probability is a *monotonic* function of stimulus intensity.

The experiments cited were done by a discrete trial, two-choice, successive discrimination method. During discrimination training, a single peck on one of the keys led to reinforcement in the presence of one of the stimuli, S_1 , and a single peck on the alternate key led to reinforcement in the presence of the second stimulus, S_2 . Generalization tests were done by presenting intensities other than S_1 or S_2 on trials for which no reinforcement was scheduled. The generalization curves under discussion relate the proportion of responses to one of the choice keys to stimulus intensity.

Generalization curves obtained after training on easy and difficult discriminations had

the same general form, though the ogives were somewhat steeper after training on difficult discriminations. After training on difficult discriminations, the generalization curves show large transposition effects; *i.e.*, the smallest and largest values of response probability do not occur at the values of the training stimuli, but beyond them. This phenomenon is similar to the positive and negative peak shifts, but, since the curves have no peaks, the peak-shift terminology is not appropriate here.

In these experiments, there were neither peaks nor other features (such as differences in choice latency) that distinguish the two stimuli used in training from other stimuli on the dimension. These facts do not accord with the general view of conditioning, which holds that the procedures used in discrimination training impart to the specific stimuli used in training the power to evoke the conditioned responses (or to prevent them from being evoked) and that this power spreads, with some decrement, to other stimuli that are similar to those used in training.

Several authors have suggested that the theory of signal recognition may provide the basis for an alternate approach to problems of stimulus control (Boneau and Cole, 1967; Blough, 1969; Heinemann *et al.*, 1969). The stimulus control curves found by Heinemann *et al.* (1969) can be described quite well by a model based on the theory of signal recognition and on the assumption that the experimentally manipulated stimuli do not govern responding on all trials; in other words, that the subjects do not "attend" to the relevant

¹This research was supported by United States Public Health Service Grant MH 13955. Reprints may be obtained from Eric G. Heinemann, Department of Psychology, Brooklyn College, Brooklyn, New York 11210.

stimuli on all trials. In the experiments cited, the stimulus control curves were obtained after the discriminations had been well learned. The present experiments examined the changes that such curves undergo during the course of training, and evaluated the applicability of the signal recognition and attention model to the description of acquisition processes.

GENERAL METHOD

Subjects

Eight White Carneaux pigeons, approximately 2 yr old at the beginning of the experiment, were kept at 80% of their free-feeding weight throughout the experiment. Four subjects served in each of two experiments.

Apparatus

The apparatus was a three-key Grason Stadler pigeon chamber with automatic control and recording equipment. A Grason-Stadler noise generator, Model 901B, in conjunction with a series of attenuators, was used to produce white noise stimuli covering an intensity range of 65 to 100 dB re 0.0002 dyne/cm². Sound intensities were measured with a General Electric sound level meter, Type 1565-A (C network), placed in front of the response panel in the approximate position of the pigeon's head.

EXPERIMENT I

Preliminary Training

After the pigeons were trained to eat from the food magazine, the center key was illuminated and the pigeons were trained to peck this key by the method of successive approximation. A single peck on the key caused the keylight to go off and the food magazine to be raised for 2.3 sec. After 15 reinforcements, an intertrial interval varying in duration from 5 to 35 sec was introduced. During the intertrial interval, the center key was not illuminated and key pecking was not reinforced. At the start of a trial, the center keylight came on and remained on until the bird pecked the key. After 10 such trials on the center key, there were 45 trials on which one of the three keys, selected at random, was lighted and a response to the lighted key was the only one reinforced.

Discrimination Training

The start of each trial was signalled by the illumination of the center key and the simultaneous presentation of white noise at one of 10 levels of intensity. These levels were 67, 72, 75, 80, 83, 86, 90, 93, 96, and 99 dB re 0.0002 dyne/cm². A peck on the center key caused the keylight to go off and the illumination on both of the side keys to come on. A subsequent response to either of the side keys ended the trial by causing the side keylights and the auditory stimulus to go off. In addition, if the response was to the key (left or right) defined as correct for the auditory stimulus presented during the trial, it was followed by reinforcement. Responding to one of the keys was defined as correct if it occurred in the presence of an intensity level of 83 dB or less, and responding to the alternate key was defined as correct if it occurred in the presence of an intensity level of 86 dB or greater. For Birds 21 and 23, responding to the left key was reinforced in the presence of any of the five high-intensity stimuli and responding to the right key was reinforced in the presence of any of the five low-intensity stimuli. These relations between the intensity levels and correct responses were reversed for Birds 22 and 24.

The auditory stimuli were presented in successive blocks consisting of 10 intensities arranged in different random orders. A trial that ended with a correct response was always followed by the next stimulus in the predetermined sequence. However, if an incorrect response was made the stimulus presented during that trial was repeated on the next trial. A training session ended when 10 reinforced responses had been made in the presence of each stimulus.

Daily training sessions were continued until every bird had attained a level of performance that was clearly asymptotic. Discrimination training was stopped after 54 days.

RESULTS

Throughout the experiment, stimuli that were not responded to correctly were repeated on the next trial. All the results reported refer to responses made on the first presentation of each stimulus. Responses to the key that was correct for the high-intensity stimuli are designated R₂; responses to the alternate key

are designated R_1 . Figure 1 shows how the proportion of R_2 responses varies with stimulus intensity at different stages of training. Each of the first nine curves represents the results obtained on three successive days of training (30 responses per point). The last curve, located at the front of each graph, is based on the last 27 days of training and represents the asymptotic form of the choice curve.

At the very beginning of training, the response proportions would be expected to scatter about a horizontal line. For some subjects there appears to be some stimulus control at the end of the first three days of training. As training proceeds, the proportions of R_2 responses occurring in the presence of the high-intensity stimuli increase and those occurring in the presence of the low-intensity stimuli decrease. The choice curves gradually change their shape until they attain the sig-

moid form of typical psychometric functions. The final choice curves are virtually indistinguishable from choice curves obtained from pigeons that are trained to discriminate only between two intensity levels approximating the extremes used in the present experiment. This can be seen in Figure 2, where the curve representing the mean performance of the subjects of the present experiment is compared with the mean choice curves of Groups 1 and 1a of Heinemann *et al.* (1969). The latter curves were obtained after the subjects were trained to discriminate between intensity levels of 68 dB and 97 dB re 0.0002 dyne/cm². The method used to obtain the choice curves for Group 1a was identical with the method used in the present experiment, whereas the method used for Group 1 differed from the present one in that responding in the presence of intensities other than the two levels used in training was never reinforced. Heinemann

Fig. 1. Proportion P of R_2 responses as a function of stimulus intensity at various stages of training.

et al. obtained choice curves that are steeper than those shown in Figure 2 from pigeons that were trained on more difficult discriminations (when the choice curves were determined by the same method used in the present experiment). Since easy and difficult discriminations were scheduled concurrently in the present experiment it might have been expected that the choice curves would be steeper than those found after training on an easy discrimination alone. The fact that there is no apparent difference in the steepness of the curves shown in Figure 2 suggests that the steepness of the choice curves obtained in the present experiment is determined by the easiest discrimination presented.

Heinemann *et al.* (1969) found that the

choice curves obtained after training has been completed can be described quite accurately by a model based on the theory of signal recognition and a correction for inattention. There follows a brief outline of this model. A more detailed discussion of the theory of signal recognition may be found in Green and Swets (1966).

Signal Recognition

It is assumed that the sensory effect produced by each of the stimuli presented during the experiment varies from trial to trial and is normally distributed. The distributions produced by repeated presentations of the different stimuli are assumed to have different means but the same variance. The sub-

Fig. 2. Comparison of the mean choice curve of the subjects in the present experiment with mean choice curves obtained by Heinemann *et al.* (1969) for two groups of subjects trained to discriminate between intensity levels of 68 dB and 97 dB.

ject is assumed to make response R_1 when the sensory effect falls below some criterion value, and to make response R_2 when this effect exceeds the criterion value.

A set of assumed distributions for different stimuli (1 to 7) is shown in Figure 3 (A). The point on the x-axis marked c represents the criterion. The proportion of R_2 responses occurring over a large series of presentations of a given stimulus is given by the portion of area under the curve for that stimulus that falls to the right of the criterion. If the proportions of R_2 responses for each of the distributions shown in Figure 3 (A) are plotted against the means of the distributions, the resulting curve is a normal ogive (Figure 3 (B)) that is identical with the ogive obtained by cumulating the area under the particular distribution with mean equal to the criterion, *i.e.*, distribution 4 in Figure 3 (A) (*cf.* Boneau and Cole, 1967). If the y-axis of Figure 3 (B) is changed from proportions to normal deviates (*z*-scores) the ogive becomes a straight line with slope equal to $1/\sigma$, where σ is the common standard deviation of the distributions shown in Fig-

Fig. 3. (a) Assumed distributions of sensory effects produced by repeated presentations of each of seven different stimuli. (b) Proportion of R_2 responses when the criterion is at c . (c) A replot of (b) with the y-axis scaled in normal deviate units rather than proportions.

ure 3 (A). Note that for the distribution with mean equal to the criterion, the proportion of R_2 responses has the value 0.5. Conversely, the stimulus value that yields $p(R_2) = 0.5$ has a distribution with mean equal to the criterion. This stimulus value then represents the criterion on the stimulus axis and is also referred to as "the criterion" in this paper.

The parallel curves shown in Figure 4 were constructed in the same way as the curve shown in Figure 3 (B), but each curve is based on a different assumed value of the criterion. Figure 4 illustrates that the effect of a change in criterion is to move the entire ogive parallel to the stimulus axis, leaving the slope unchanged.

Within the framework of signal recognition theory, the slope of the ogives is a pure measure of sensitivity to stimulus differences. Referring to Figure 3 (C), it is not difficult to show that the difference in the ordinate values for any two stimuli is equal to the quantity $d' = M_2 - M_1/\sigma$, where M_2 and M_1 are the means of the distributions corresponding to the two stimuli under consideration, and σ is the common standard deviation. In other words, $z_2 - z_1 = d'$, where z_2 and z_1 are the ordinate values for any two stimuli, read off from Figure 3 (C).

Because analyses using signal recognition theory generally involve ROC curves rather than psychometric functions, it may be useful to show the relation between these two kinds of representations. Consider the construction in Figure 5 (A). I_1 and I_2 are two stimulus values for which points on an ROC curve are to be found. The proportions A_1 , A_2 , and A_3 represent the conditional probabilities of response R_2 given stimulus I_1 , $p(R_2|I_1)$, for three different values of the criterion. Similarly, B_1 ,

Fig. 4. Normal ogives that differ from each other in the value of the criterion (thin lines). The thick line represents the arithmetic mean of the four thin-line curves.

B_2 , and B_3 represent the corresponding values of $p(R_2|I_2)$. Plotting these two sets of conditional probabilities against each other yields three points on an ROC curve as shown in Figure 5 (B). The same method may be used, of course, to construct an ROC curve based on any other pair of stimulus values. It can be seen that ROC curves and psychometric functions are simply alternate ways to represent the same data, though there is one important difference, *viz* ROC curves show the relationship between two dependent variables whereas a family of psychometric functions shows the relationship of these dependent variables to the stimulus values as well.

A final matter that is important in the analysis that follows is the effect of averaging a number of psychometric functions that differ from each other only in the value of the criterion. A set of such curves is shown in Figure 4, together with a curve representing the mean proportions (heavy line). Note that the mean curve is less steep than the individual curves. This flattening effect produced by averaging must be taken into account when interpreting results obtained over periods during which the criterion may have been changing systematically. The variations in criterion value do not have to be systematic, of course, for the flattening to occur. Thus, random trial-to-trial fluctuations in the value of the

criterion would also result in a stimulus control curve less steep than it would have been in the absence of such criterion fluctuations.

The Correction for Inattention

A normal ogive has its lower asymptote at $p = 0$ and its upper asymptote at $p = 1.0$. Even at the end of training, the asymptotes of empirical choice curves often depart from these values. Heinemann *et al.* (1969) assumed that the failure of the asymptotes to lie at the values required by the signal recognition model indicates that on some trials, the subjects' choice behavior is not governed by the stimuli manipulated by the experimenter but by other unspecified stimuli; in other words, that the subjects do not attend to the relevant stimuli on every trial. On those trials on which the subject does attend to the auditory dimension, it is assumed to behave in accordance with the signal recognition model; on those trials on which it does not attend, the auditory stimuli are assumed to have no influence on the choice of response. If it is assumed further that the probability of attending to the auditory dimension is independent of the intensity of the stimulus presented, it is possible to derive a simple correction that removes the effects of inattention.

The observed probability of response to one of the keys, $p(R)$, can be expressed as the sum of two joint probabilities as follows:

$$p(R) = p(A) \cdot p(R|A) + p(\bar{A}) \cdot p(R|\bar{A}) \quad (1)$$

where $p(R|A)$ is the conditional probability of response given attention to the auditory stimulus dimension, $p(R|\bar{A})$ is the conditional probability of response given inattention, $p(A)$ is the probability of attention to the auditory dimension and $p(\bar{A}) = 1 - p(A)$.

Using the specification by the signal recognition model that $p(R|A)$ is zero for stimuli at the lower asymptote and unity for stimuli at the upper asymptote, it is possible to get an estimate of the probability of attention, $p(A)$. Writing Equation 1 once for $p(R|A) = 0$, and again for $p(R|A) = 1.0$, and then subtracting the first equation from the second shows that

$$p(A) = p(R_u) - p(R_l) \quad (2)$$

where $p(R_u)$ is the obtained value of the upper asymptote and $p(R_l)$ is the obtained value of the lower asymptote.

Fig. 5. The relation between psychometric functions and ROC curves (See text.)

From

$$p(R_1) = [1 - p(A)] p(R|\bar{A})$$

one gets

$$p(R|\bar{A}) = \frac{p(R_1)}{1 - p(A)} \quad (3)$$

Substituting the right-hand parts of Equations 2 and 3 for the appropriate quantities in Equation 1 and solving for $p(R|A)$ leads to

$$p(R|A) = \frac{p(R) - p(R_1)}{p(R_0) - p(R_1)} \quad (4)$$

Equation 4 can be used to compute a corrected proportion $p(R|A)$ for each observed proportion $p(R)$, and will be referred to as the "correction for inattention".

Formal Description of the Data

The results shown in Figure 1 strongly suggest that during the course of training, the probability of attention changes from an initial value of zero to a final value near 1.0. Many of the curves show clear lower and upper asymptotes, with the difference between these asymptotes increasing as training progresses (*e.g.*, all but the first and third curves for Bird 22), or show fairly clear lower asymptotes whose value decreases as training progresses. Even if the curves do not exhibit clear upper and lower asymptotes, *i.e.*, if these asymptotes lie beyond the ends of the stimulus range, the values of these asymptotes can be calculated, given the assumptions of the model that has been presented.

In principle, then, it is possible to apply the correction for inattention to each of the curves obtained during training. If the only effect of training were to change the probability of attention to the relevant stimulus dimension, all of the *corrected* curves would be identical to each other. A detailed analysis of this sort was attempted but the empirical curves, particularly for the first few blocks of training days, are too irregular to permit any firm conclusions.

EXPERIMENT 2

In an attempt to obtain more orderly choice curves at all stages of training, the experiment was repeated with several modifications, the principal one being that correct responses were not reinforced on every trial. This was

done in the expectation that the probability of attention would change more slowly, thus permitting the construction of curves based upon larger blocks of training trials.

METHOD

The preliminary training procedure was the same as that used in Experiment 1. The procedure used in discrimination training differed from that of Experiment 1 in the following ways. The 10 levels of white noise used covered the range of 60 to 96 dB re 0.0002 dyne/cm² in steps of 4 dB. Responding to one of the side keys was reinforced in the presence of all intensities less than or equal to 76 dB, and responding to the alternate key was reinforced in the presence of all intensities greater than or equal to 80 dB. For Birds 90 and 92, responding to the left key was reinforced in the presence of the high-intensity stimuli and responding to the right key was reinforced in the presence of the low-intensity stimuli. These relations between intensity levels and correct responses were reversed for Birds 91 and 93. Whereas in Experiment 1 all correct responses were reinforced, in the present experiment, reinforcement for correct responses was available only on 70% of all trials. The 10 stimuli were presented in random order with the restriction that each appear on 10 trials during each session, with reinforcement for correct responding available on seven of these 10 trials. If an incorrect response occurred on a trial for which reinforcement was scheduled, the stimulus incorrectly responded to was presented again on the next trial. A daily session consisted of 100 trials that ended with correct responses. Training was continued for 100 days.

RESULTS

The results of Experiment 2 are shown in Figure 6. Each curve represents the results obtained on 10 successive days of training (100 responses per point). Starting from the left, the first curve represents the first block of 10 days, the second curve represents the second block of 10 days, and so on. To avoid overlap, all curves except the first one have been displaced to the right by arbitrary amounts, as specified in the legend under Figure 6.

When the data shown in Figure 6 are corrected for inattention, the corrected points can be fitted quite well by normal ogives. The

Fig. 6. Proportion of R_2 responses as a function of stimulus intensity, at various stages of training. Moving from left to right, the curves represent the results obtained on successive blocks of 10 days. The first curve is correctly placed in the coordinate system. The second curve has been moved to the right by 30 dB, and each of the remaining curves has been moved to the right 15 dB more than the immediately preceding one.

smooth curves shown were obtained by applying the correction in reverse to the ogives fitted to the corrected points, *i.e.*, by using equation 1 to obtain values of $p(R)$ that correspond to the theoretical values of $p(R|A)$. The actual fitting was done with the aid of a digital computer. The program used an iterative procedure to determine the values of the upper

and lower asymptotes of the choice curves and the parameters of the normal ogive fitted to the corrected proportions that yielded the least sum of weighted (by the reciprocal of the variance) squared deviations of the uncorrected proportions about the theoretical curve. In two instances, the curves for the first block of 10 days were so close to hori-

zontal lines that a meaningful fit was not possible.

To describe the curve-fitting procedure in somewhat more detail: it was decided, first of all, to fit a normal ogive to each empirical curve. Normal ogives might be fitted to the points without reference to any correction for inattention (this is equivalent to assuming that the probability of attention has the value 1.0 for all curves), or they might be fitted after a correction for inattention has been applied to the data. Whether or not to correct for inattention, and by how much, was decided on the basis of a goodness of fit criterion for each individual set of empirical points. The computer program² first fitted an ogive assuming the lower and upper asymptotes of the curve to be 0 and 1.0, respectively, and then repeated the process for each of a large series of other assumed values of the upper and lower asymptotes, finally selecting those values for the asymptotes that yielded the least sum of weighted (by the reciprocal of the variance) squared deviations of the empirical points about the theoretical curve.

Figure 7 shows how the difference between the upper and lower asymptotes of the best-fitting curves (the probability of attention) varies during training for each subject. For all subjects, the data representing the first few 10-day blocks of training were fitted better with a correction for inattention than without it, and it can be seen that the probability of attention values required to obtain the best fit increase systematically as training proceeds, attaining a value near 1.0 after 40 to 60 days. For many of the choice curves obtained early in training, the asymptotes derived from the computer analysis lie beyond the ends of the stimulus range and are thus not evident to inspection. In a few cases, however, the raised lower asymptotes can be seen quite clearly, *e.g.*, the third curve for Bird 91, and the second curves for Birds 92 and 93.

The stimulus value that corresponds to a response proportion of 0.5 on the corrected (for inattention) choice curves will be called the criterion. The way the value of the criterion varies during the course of training is shown in Figure 8. The results for all birds show an initial bias in favor of the key that

is correct for the low-intensity stimuli. As training proceeds, the criterion is lowered until, for all subjects except Bird 93, it comes to lie between the highest intensity for which R_1 was reinforced and the lowest intensity for which R_2 was reinforced. It is obvious that this placement of the criterion is approximately optimal in the sense that it maximizes the number of reinforcements the bird receives per session.

Figure 9 shows the variations in the slope of the corrected choice curves, expressed in units of z per dB, during the course of training. The corrected choice curves for all subjects become steeper as training proceeds. This increase in slope undoubtedly reflects in part the systematic changes in criterion value that are shown in Figure 8. Note that each of the choice curves can be considered as the average of 10 choice curves, one for each day. If the criterion changes systematically during the 10-day period, the 10 one-day curves will differ from each other in their placement along the x-axis, and as was illustrated in Figure 4, the mean curve will be flatter than the individual curves. Moreover, the more widely separated the individual curves are from each other, *i.e.*, the greater the change in criteria over the 10-day period, the flatter the average curve will be. Thus, average curves obtained early in training, when the criterion is changing rap-

Fig. 7. Changes in the probability of attention during training.

²A copy of the computer program is available from the first author on request.

Fig. 8. Changes in the mean value of the criterion during training.

Fig. 9. Changes in the slope of the corrected choice curves during training.

idly, can be expected to have lesser slopes than those obtained later.

However, systematic changes in the value of the criterion are probably not the sole cause of the observed changes in slope. It can be shown that if the criterion varies randomly from trial to trial, the slope of the choice curve will vary inversely with the standard

deviation of the distribution of criteria (*cf.* Eijkman and Vendrik, 1963; Wickelgren, 1968). It may well be that one of the effects of discrimination training is to reduce criterion variance. This is an assumption introduced previously by Heinemann *et al.* (1969) to account for the finding that training on difficult discriminations results in steeper choice curves

than training on easy ones. That such a reduction in criterion variance contributes to the slope changes is suggested by the results of Birds 91 and 93. The curves for these birds (Figures 8 and 9) appear to indicate that the slope of the choice curves continues to increase late in training after the mean value of the criterion has already attained an asymptotic level. Unfortunately, the results for single days are too irregular to permit a more detailed evaluation of the relative importance of changes in the mean and the variance of the criterion distribution.

DISCUSSION

To summarize briefly, the choice curves obtained after varying amounts of discrimination training appear to be well described by the signal recognition and attention model. In the terms of this model, the principal effects of the training procedure are to increase the probability of attention to the relevant stimulus dimension, to cause the mean of the criterion distribution to be adjusted to a value that is approximately optimal, and possibly to reduce the variance of the distribution of criteria.

In unpublished experiments, the authors have trained pigeons to discriminate between only two values of intensity of white noise, and stopped separate groups of subjects for generalization testing after varying amounts of training. The curves obtained also have the monotonic sigmoidal form, indicating that the proposed model probably describes the acquisition of a discrimination with the traditional two-stimulus method also.

It is not easy to relate the present findings or the proposed model to the large body of data obtained with the single-key free-operant method. The principal difficulty is that it is not known what accounts for the difference in the basic form of the post-discrimination gradients obtained (monotonic-sigmoidal with the choice method *versus* peaked with the free-operant method). This difference in the basic form of the curves probably springs at least in part from differences in the stimulus continua studied. On the basis of signal recognition theory, differences in the structure of the stimulus continua would be expected to produce differences in the form of the choice curves obtained after training to discriminate

between two values from the continuum. For example, compare a unidirectional continuum such as intensity with a circular continuum, such as the angular orientation of a line that is rotated about its midpoint. For the intensity continuum, a consistent rule for selecting one of two responses can be based on a single criterion or cutoff point that divides the continuum into two segments. However, in order to divide the circular continuum into two segments it is necessary to use two criteria, and it can be shown that in this case the choice curve derived from the signal recognition model will not be sigmoidal but peaked. A two-criterion situation of this sort is considered in detail by Blough (1969).

While the differences in the form of the stimulus control curves obtained with the choice and free-operant methods may well be attributable largely to differences in the stimulus continua studied, there is no doubt that differences between the methods also play a role. Ernst, Engberg, and Thomas (1971), working with an intensive continuum, luminance, and the single-key free-operant method, found post-discrimination gradients that were peaked in form, though very asymmetric. The possible reasons for the differences obtained with two methods are too numerous to justify a speculative discussion at this point.

What we have put forth is intended to be a formal description of the data, not a theory of learning. Since concepts of "attention" play a role in a number of current theories (*e.g.*, Zeaman and House, 1963, Lovejoy, 1965), it may be worth noting that our usage of the term "attention" is much closer to Skinner's (Skinner, 1953) than to that of most two-stage theorists, in that attention is not treated as a response, but as a controlling relation.

Further examination of changes in the parameters of choice curves under various conditions of acquisition may well provide better starting points for the construction of theory than examination of such overall measures of performance as the proportion of correct responses.

REFERENCES

- Blough, D. S. Generalization gradient shape and summation in steady-state tests. *Journal of the Experimental Analysis of Behavior*, 1969, 12, 91-103.
- Boneau, C. A. and Cole, J. L. Decision theory, the

- pigeon and the psychophysical function. *Psychological Review*, 1967, 74, 123-135.
- Chase, S. and Heinemann, E. G. Choices based on redundant information: an analysis of two-dimensional stimulus control. *Journal of Experimental Psychology*, 1972, 92, 161-175.
- Eijkman, E. and Vendrik, A. Detection theory applied to absolute sensitivity of sensory systems. In J. A. Swets (Ed.), *Signal detection and recognition by human observers*. New York: John Wiley & Sons, 1964. Pp. 392-409.
- Ernst, J., Engberg, L., and Thomas, D. R. On the form of stimulus generalization curves for visual intensity. *Journal of the Experimental Analysis of Behavior*, 1971, 16, 177-180.
- Green, D. M. and Swets, J. A. *Signal detection theory and psychophysics*. New York: John Wiley & Sons, 1966.
- Heinemann, E. G., Avin, E., Sullivan, M. A., and Chase, S. Analysis of stimulus generalization with a psychophysical method. *Journal of Experimental Psychology*, 1969, 80, 215-224.
- Heinemann, E. G. and Chase, S. Conditional stimulus control. *Journal of Experimental Psychology*, 1970, 84, 187-197. (a)
- Heinemann, E. G. and Chase, S. On the form of stimulus generalization curves for auditory intensity. *Journal of Experimental Psychology*, 1970, 84, 483-486. (b)
- Lovejoy, E. P. An attention theory of discrimination learning. *Journal of Mathematical Psychology*, 1965, 2, 342-362.
- Skinner, B. F. *Science and human behavior*. New York: MacMillan, 1953.
- Wickelgren, W. A. Unidimensional strength theory and component analysis of noise in absolute and comparative judgments. *Journal of Mathematical Psychology*, 1968, 5, 102-122.
- Zeaman, D. and House, B. The role of attention in retardates discrimination learning. In N. P. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill, 1963. Pp. 159-223.

Received 13 July 1970.
(Final Acceptance 30 April 1973.)