*STIMULUS GENERALIZATION AS A FUNCTION OF THE
DELAY BETWEEN TRAINING AND TESTING
PROCEDURES: A REEVALUATION¹*

DAVID R. THOMAS AND D. E. SCOTT BURR

UNIVERSITY OF COLORADO

Three groups of 12 pigeons each were trained to discriminate between lights of 550 m μ (S^P), correlated with 1-min variable-interval reinforcement and 570 m μ (S^A), correlated with extinction. Group A was tested for wavelength generalization in extinction 1 min after meeting the discrimination criterion; Group B was tested 24 hr later; Group C was tested 24 hr later after a 3-min (reinforced) warm-up with the S^P. The post-discrimination gradient of Group B was significantly flatter and showed significantly greater area shift than that of Groups A and C. The gradient of Group C was similar to that of Group A, indicating that the warm-up eliminated the effect of the delay period.

The present study was designed to reconcile the discrepancy between the results of several earlier experiments on the effects of a delay period between training and testing for generalization.

In 1958 Perkins and Weyant reported a study in which (response speed) generalization gradients obtained from rats one week after training on a runway were shown to be reliably flatter than those obtained immediately after training. Four groups were used, one tested immediately on the training stimulus, one tested immediately on the novel stimulus (alley and goal box of different color), one group tested after a one-week delay on the original alley, and one group tested after a one-week delay on the novel alley.

Subsequently Thomas, Ost, and Thomas (1960) reported results inconsistent with those of Perkins and Weyant. In this study, pigeons pecked at a monochromatic light of 550 m μ during variable-interval (VI) reinforcement and then were trained to discriminate 550 m μ (VI reinforcement) from 570 m μ (no reinforcement) when the two stimuli were presented successively. One day, one week, or three

weeks after reaching a criterion, subjects were tested for generalization carried out in extinction in a manner introduced by Guttman and Kalish (1956). A wide range of spectral stimuli was successively presented with subjects' rate of responding to each providing the measure of generalization. There was no evidence that the three delay groups produced post-discrimination gradients which differed in any way whatever.

There are many possible reasons for the discrepancy in the findings of these two studies. Different species were used, a different type of task, a different stimulus dimension, *etc.* Unlike Perkins and Weyant (1958), Thomas *et al.* (1960) tested for generalization after administering discrimination training. It is possible that one effect of this training was to increase the "stability" of the generalization gradient over a time period. In addition, Thomas *et al.* (1960) earliest test was given 24 hr after training was completed. Any changes in the gradient which take place in the first few hours after training would have been overlooked by such a procedure. To explore these procedural differences, Thomas and Lopez (1962) performed an experiment in which three groups of pigeons were trained to respond to a monochromatic light during VI reinforcement and then were tested for generalization immediately, one day later, or one week later. The results revealed a significant flattening of the gradient after one day of delay, with no further flattening after one week. Thus, the finding of Perkins and Wey-

¹This study was supported by N.I.H. research grant HD-03486-01 under the direction of the first author. Preparation of the manuscript was also aided by N.I.H. training grant MH-10427. The study was reported by the second author at the May 1968 meetings of the Rocky Mountain Psychological Association in Denver, Colorado. Reprints may be obtained from David R. Thomas, Dept. of Psychology, University of Colorado, Boulder, Colorado 80302.

ant (1958) was replicated, despite substantial differences in procedure, species, response measure, *etc.*

In the Thomas and Lopez (1962) study, the delayed tests were carried out without any prior warm-up. Thomas *et al.* probably used a short reinforced warm-up to the positive stimulus before initiating generalization testing, as was typical practice in the Duke laboratory at that time. It remains to be determined whether the failure to find an effect of the delay period in the Thomas *et al.* (1960) study was due to the initiation of testing one day after meeting criterion, a "stabilizing" effect of discrimination training, the use of a warm-up, or some other factor. The present study was designed for this purpose.

METHOD

Subjects

Thirty-six experimentally naive domestic pigeons were maintained at 70 to 75% of their free-feeding weights.

Apparatus

Two operant conditioning chambers, described in detail by Hiss and Thomas (1963), were used. The response key in each box was illuminated by a Bausch and Lomb diffraction grating monochromator, Model 33-86-02, equipped with a tungsten light source.

Procedure

Upon arrival at the laboratory, all birds were given free access to food and water for 10 days to allow their weights to stabilize. At this time, total food deprivation was initiated and continued until the appropriate weight level was achieved. Subjects were then magazine and key-peck trained, receiving 50 continuous reinforcements on the first training day. On the second through sixth days of training all subjects earned 30 reinforcements per day on a gradually lengthening temporal interval such that by the sixth day, a VI 1-min schedule of reinforcement was in effect. During this preliminary training the key was illuminated by a 550-m μ light. Starting on the third day of training, 10-sec periods of blackout were alternated with 60-sec work periods. The subjects' pecks were never reinforced during blackouts, and little or no responding occurred during blackout periods.

Starting with the seventh day, successive discrimination training was administered. The positive stimulus (S^D) was the original training stimulus, 550 m μ ; the negative stimulus (S^A) was 570 m μ . During discrimination training, positive and negative periods of 60-sec duration were randomly alternated, with the restriction that no more than three S^D or S^A periods appeared in succession. Work periods were alternated with 10-sec blackout periods. The VI 1-min reinforcement schedule was in effect during all positive periods; responses in the presence of the negative stimulus and during blackouts were not reinforced.

Discrimination training was administered for 30 min daily until each subject reached a criterion of no responses during five consecutive negative periods while maintaining responding to the positive stimulus. As the birds achieved this criterion, those in Group A were immediately tested for generalization in extinction; Groups B and C were tested 24 hr later, the latter after a 3-min reinforced warm-up with the positive stimulus. During generalization testing, stimulus-on periods were of 30-sec duration alternated with 10-sec blackouts. Eleven stimuli (490 to 610 m μ in 10-m μ steps, omitting 500 and 600 m μ) were randomized within a series, and six different random series comprised each generalization test. The different wavelengths were employed without correction for intensity differences in the output of the monochromator.

Because the subjects in Groups A and C were to be tested for generalization after (at least part of) a daily training session, whereas those in Group B were not, a control was required for the difference in deprivation between these groups. For this purpose half of the subjects in Group B were pre-fed their daily ration 15 min before generalization testing was begun. As a consequence, their rate of responding during test was slightly (insignificantly) reduced, but their relative generalization gradients were indistinguishable from those of the other six birds in their group.

RESULTS

The number of minutes required to achieve the criterion of discrimination ranged from 54 to 131 with a mean of 96.6 for Group A, 101.8 for Group B, and 103.4 for Group C. ($F < 1$, not significant) Similarly, the three

groups failed to differ reliably in total number of responses emitted during generalization testing. The respective means were 671.3 for Group A, 866.8 for Group B, and 672.3 for Group C ($F(2,33) = 1.36$, n.s.). Thomas *et al.* (1960) and Thomas and Lopez (1962) also failed to note any change in total response output over the delay intervals they sampled.

For each subject, an individual relative post-discrimination gradient was obtained based on the per cent of total responses emitted to each of the generalization test stimuli. Figure 1 presents the mean of these individual gradients for Group A (immediate), Group B (24-hr delay), and Group C (24-hr delay plus 3-min warm-up).

All groups show the typical effect of successive intra-dimensional discrimination training on the post-discrimination gradient, a displacement of the mode of responding from the S^D so as to be farther removed from the S^A . The figure suggests, however, that the peak shift is greater after a 24-hr delay if no warm-up is given. Figures 2, 3, and 4 present the individual gradients of the subjects in the three groups. All of the gradients obtained from Groups A and C showed peaks at 540 $m\mu$; in Group B, five gradients peaked at 540 $m\mu$, six peaked at 530 $m\mu$, and one peaked at 520 $m\mu$. To analyze statistically the difference in area shift, each individual gradient was treated as a grouped frequency distribution and the mean of that distribution

Fig. 1. Mean relative generalization gradients of groups of subjects tested immediately after achieving criterion or 24 hr later.

Fig. 2. Individual relative generalization gradients of subjects tested immediately after achieving criterion.

was obtained. If the gradient were symmetrical about the positive stimulus, 550 $m\mu$, the mean would fall at this value. With the nega-

Fig. 3. Individual relative generalization gradients of subjects tested 24 hr after achieving criterion.

tive stimulus of longer wavelength than the positive, the greater the area shift the lower the mean value of the gradient would be. The mean of the gradient mean values was 540.08 for Group A, 533.03 for Group B, and 541.96 for Group C ($F(2,33) = 17.58, p < 0.01$). The value for Group B was significantly lower ($p < 0.01$) than that of the other two groups, which in turn did not differ.

To compare the slopes of the gradients of the three groups, the mean per cent of total responses given to the stimuli to which the peak had shifted was taken as an index of slope; the higher that per cent, the steeper the gradient. Because not all gradients peaked at the same stimulus value, it was decided to use the per cent of total responses at the peak of each individual gradient, wherever it lay. To make this comparison, a relative post-discrimination gradient was plotted for each subject to include the peak (as the central stimulus) and three values on each side of it. (Averaged over all subjects in the experiment, a mean of 98.27% of all responses fell within this range.) The group means of the individual gradients treated in this manner are shown in Fig. 5, and it is clear that the group difference in

gradient steepness evidenced in Fig. 1 remains. The difference between per cent values at the peaks of the three group gradients is significant ($F(2, 33) = 3.49, p < 0.05$). The value for Group B was significantly lower ($p < 0.05$) than that of the other two groups, which in turn, did not differ.

DISCUSSION

Quite clearly, a 24-hr delay between training and testing for generalization produced both a greater peak shift and a flatter gradient than that obtained on an immediate test. The data also indicate that a 3-min exposure to reinforcement during the positive stimulus immediately before testing eliminated the effect of the intervening delay. It is therefore possible to conclude that the failure of Thomas *et al.* (1960) to obtain differences in generalization gradients following delay periods of one day, one week, and three weeks was due to: (1) the initiation of testing after a 24-hr delay after which most (if not all) measurable forgetting had already taken place, and (2) the use of a warm-up to S^D which counteracts the effects of the delay period. Of course, neither of these effects was known at the time their experiment was performed.

At this point, it may be useful to speculate on why the gradient shifts farther from S^A during the 24-hr delay and why the warm-up

Fig. 4. Individual relative generalization gradients of subjects tested 24 hr after achieving criterion following a 3-min (reinforced) warm-up to the S^D .

Fig. 5. Mean relative generalization gradients of Groups A, B, and C, based on peak value $\pm 10, 20$, and $30 \text{ m}\mu$.

with the positive stimulus counteracts this effect. One possibility is that after discrimination training there is differential forgetting of the "effect" of the positive and the negative training stimuli, such that the "negativity" or "aversiveness" of the S^A is enhanced relative to the "positivity" of the S^D . The reinforced warm-up on the positive stimulus value may then reestablish the balance between the two. This interpretation assumes an explanation of the peak shift itself which is clearly subject to question. Nevertheless, such speculations may have value in pointing the way for subsequent research endeavor.

Gleitman and Steinman (1963) observed that, although conditioning principles play a significant role in much theorizing about forgetting, there has been little systematic research carried out in this area with animal subjects and in simple conditioning situations. In studies by Gleitman and Bernheim (1963) and Gleitman, Steinman, and Bernheim (1965), forgetting was defined by a failure of rats to improve in fixed-interval performance following a 24- or 25-day interruption in fixed-interval training. The use of changes in generalization as a forgetting index may be superior in that testing is uncontaminated by effects of reinforcement. Generalization test procedures have indicated, in agreement with the human literature, that the bulk of forgetting occurs in the first 24 hr after acquisition. The distinction between response forgetting (reflected in a change in total response output) and stimulus-response connection forgetting (reflected in changes in the shape or location of the generalization gradient) also has precedent in the human paired-associate situation (*c.f.*, McGovern, 1964). Whether or

not the present procedures will prove more sensitive to retroactive and proactive interference effects than have past conditioning studies with animals, of course, remains to be determined. In any case, these procedures suggest intriguing and significant theoretical questions which would probably not have arisen in the context of more typical approaches to the study of forgetting.

REFERENCES

- Gleitman, H. and Bernheim, J. W. Retention of fixed interval performance in rats. *Journal of Comparative and Physiological Psychology*, 1963, **56**, 839-841.
- Gleitman, H., Steinman, F., and Bernheim, J. W. Effect of prior interference upon retention of fixed interval performance in rats. *Journal of Comparative and Physiological Psychology*, 1965, **59**, 461-462.
- Guttman, N. and Kalish, H. I. Discriminability and stimulus generalization. *Journal of Experimental Psychology*, 1956, **51**, 79-88.
- Hiss, R. H. and Thomas, D. R. Stimulus generalization as a function of testing procedure and response measure. *Journal of Experimental Psychology*, 1963, **65**, 587-592.
- McGovern, J. B. Extinction of associations in four transfer paradigms. *Psychology Monographs*, 1964, **78**, 1-21 (whole #593).
- Perkins, C. C., Jr. and Weyant, R. G. The interval between training and testing and test trials as a determiner of the slope of generalization gradients. *Journal of Comparative and Physiological Psychology*, 1958, **51**, 596-600.
- Thomas, D. R. and Lopez, L. J. The effects of delayed testing on generalization slope. *Journal of Comparative and Physiological Psychology*, 1962, **55**, 541-544.
- Thomas, D. R., Ost, J., and Thomas, D. H. Stimulus generalization as a function of the time between training and testing procedures. *Journal of the Experimental Analysis of Behavior*, 1960, **3**, 9-14.

Received 5 June 1968.