Proactive interference in the retention of "attentional sets" in pigeons

DAVID R. THOMAS

University of Colorado, Boulder, Colorado 80309

D. E. SCOTT BURR

University of Central Florida, Orlando, Florida 32816

and

MARK B. VOGT

University of Colorado, Boulder, Colorado 80309

In Experiment 1, six groups of pigeons ($n=8$) were tested for wavelength generalization either immediately or 24 h after learning a successive discrimination, with 550 nm reinforced and a black vertical line extinguished. The groups differed in the stimulus present during single stimulus pretraining, which was 550 nm (pretrain S+), the vertical line (pretrain S-), or a neutral dim white light (pretrain Sn), respectively. The three immediate generalization gradients were steep and indistinguishable, reflecting only the immediately preceding discrimination training condition. The three delay gradients were flatter, with the flattening particularly marked in the pretrain S- group. This was interpreted as proactive interference (PI) resulting from the memory that both the 550-nm and the line stimuli had previously been reinforced. In Experiment 2, two (TD) groups of pigeons ($n=16$) were given single stimulus training with a 555-nm keylight followed by eight sessions of discrimination training with two line angles, then one session of non-differential (ND) training with the same two lines, and then a wavelength generalization test either immediately or after a 24-h delay. Two other (hold) groups ($n=16$) received similar training, except for the TD line angle training sessions. In these hold groups, the wavelength gradient was flatter in a delayed test; in the TD groups it was steeper, indicating PI from the prior TD training. These two experiments suggest that the "attentional sets," which purportedly result from TD and ND training, may fruitfully be viewed as target memories subject to the principles of interference theory.

A stimulus generalization test may be viewed as constituting a test of recognition memory in which responding to a given stimulus value is analogous to a human adult subject's making the verbal response, "That stimulus is the one to which I was trained to respond." Since responding during generalization testing tends to be continuous rather than all or none, different rates of responding to different stimuli may be taken to reflect the subject's degree of certainty that a given stimulus is the former training stimulus. As forgetting occurs, the degree of certainty is pre-

sumably reduced, resulting in a flattened generalization gradient. The location of the peak (or mode) of the gradient at the value of the training stimulus gives an indication of what is being remembered, whereas the slope of the gradient indicates how well it is remembered.

Thomas and Lopez (1962) trained pigeons to peck a key illuminated by a monochromatic light of 550 nm and then, after 10 sessions of variable interval 1-min (VI 60-sec) training, three different groups were tested for wavelength generalization in extinction. One group was tested "immediately" (i.e., after a 1-min delay), another was tested 24 h later, and the third was tested after a 1-week delay. The relative generalization gradient of the immediate test group was significantly steeper than that of the other two groups, which did not differ. The finding of greatest forgetting in the period soon after learning is, of course, typical of forgetting curves since the experiments of Ebbinghaus (1913).

Given the finding that forgetting with the passage of time may be reflected in flattened generalization

Requests for reprints should be addressed to David R. Thomas, Department of Psychology, Campus Box 345, University of Colorado at Boulder, Boulder, Colorado 80309. This research was supported by NSF Research Grant BNS-7801407. Experiment 1 was performed by D. E. Scott Burr as part of a doctoral dissertation submitted to the University of Colorado in 1972. Experiment 1 and Replication 1 of Experiment 2 were reported by David R. Thomas at the November 1980 meeting of the Psychonomic Society in St. Louis. Replication 2 of Experiment 2 was reported by Mark B. Vogt at the April 1981 meeting of the Rocky Mountain Psychological Association in Denver.

gradients, Burr and Thomas (1972) predicted that forgetting resulting from conflicting memories would have a similar effect. Interference theory (cf. Keppel, 1968) postulates that prior learning of a conflicting habit can interfere with the retention of a subsequently learned habit (proactive interference, PI). Presumably, unlearning (extinction) of the first habit occurs during the acquisition of the second habit, later spontaneous recovery of the first habit over time results in interference with the second habit, causing PI (Underwood, 1948; Underwood & Postman, 1960). In verbal learning, Keppel (1964) reported that spacing (i.e., distributing) practice on a second task over several days reduced and perhaps eliminated PI, apparently through repeated extinction of the spontaneously recovering prior habit. It is also possible that massed practice on the second task is conducive to the demonstration of PI because it leads to less effective learning of Task 2, rendering it more subject to interference. The study by Burr and Thomas (1972) included the massing, or spacing, of training on Task 2 as an independent variable in addition to the presence or absence of (spaced) Task 1 learning, and immediate vs. delayed generalization testing. The two habits, Task 1 and Task 2, were directly conflicting, that is, Task 1 was a successive discrimination with responding to a light of 570 nm reinforced and responding to 550 nm extinguished; Task 2 was the reversal of this problem, that is, 550 nm was S+ and 570 nm was S-.

All groups were tested for wavelength generalization in extinction either immediately or 1 day after the completion of Task 2 training, and the results were clearly in accord with prediction. The group showing the most interference, as reflected in the flattest gradient, received (1) training on both tasks, (2) massed practice on Task 2, and (3) a delayed generalization test.

Experiment 1 in this paper is an extension of the Burr and Thomas (1972) study. It was designed to extend the range of habits that might be considered "conflicting" from the perspective of interference theory. Our discussion, up to this point, has treated "what is learned" as a simple response tendency or set of tendencies (e.g., peck at 550 nm, do not peck at 570 nm). It has been proposed (cf. Thomas, 1970) that, in learning a discrimination, the subject also learns about the importance of stimuli in general and the benefits of paying attention to stimulus differences. Such learning of a set (called "general attention") may be reflected in sharpened generalization gradients even along dimensions not involved in original training. It does not take place under nondifferential (i.e., equivalence) training with two stimuli, or under single stimulus training. Indeed, by showing subjects that stimulus differences are unimportant, nondifferential training purportedly produces an inattentive set. This set is then reflected in a relatively

flat generalization gradient, in comparison with that produced by single stimulus training, a more neutral condition (cf. Thomas, 1970).

Suppose, then, that the conflicting habits learned in a PI paradigm were (1) to be highly attentive (or responsive) to stimulus differences vs. (2) to be less attentive. An immediate generalization test would presumably reflect the attentional set produced by the most recent training, just as Burr and Thomas (1972) found that an immediate test reflected only Task 2 learning in their study. With the passage of time before testing, however, the attentional set produced by Task 1 training might increase in relative strength and be reflected in an appropriate change in generalization slope. Experiment 1 examined this hypothesis, and was concerned with another matter as well. While the Burr and Thomas (1972) study compared the gradients of groups exposed to Task 1 and Task 2 with those exposed only to Task 2, in the present study all subjects received a single-stimulus pretraining experience as Task 1, but they differed in the relationship of the pretraining stimulus to the stimuli used in Task 2 interdimensional training. It was expected that this factor would affect the amount of PI observed.

EXPERIMENT 1

Method

Subjects. The subjects were 48 experimentally naive domestic pigeons obtained from a local supplier. All subjects were maintained at 70%-75% of their ad-lib weights for the duration of the experiment. The birds were housed individually in a colony room on a 12-h light/dark cycle and with continuous access to grit and water.

Apparatus. The apparatus was the same as that employed in the Burr and Thomas (1972) study. Two similar pigeon Skinner boxes were used. The response key in each box was transilluminated by a Bausch and Lomb diffraction grating monochromator, Model 33-86-02, equipped with a tungsten light source. Masking noise was provided by a white-noise generator.

Procedure. The subjects were assigned at random to one of three prediscrimination training conditions, which will hereafter be identified according to the pretraining stimulus. These three conditions were: (1) pretrain on the stimulus that was to become the positive stimulus in the later interdimensional discrimination problem (pretrain S+), (2) pretrain on the stimulus that was to become the negative stimulus (pretrain S-), and (3) pretrain on a "neutral" stimulus (i.e., a dim white light) that was orthogonal to the stimuli to be used in the later discrimination problem (pretrain Sn).

Magazine and keypeck training were accomplished on the first day of training. A reinforcement consisted of 3-sec of access to a tray of mixed grain. From Day 2 to Day 10, the pigeons received 20 reinforcements daily, delivered on an increasingly long intermittent schedule such that, by Day 10, a VI 60-sec schedule was in effect. Commencing on the 4th day, 5-sec periods of blackout were alternated with 50-sec stimulus-on periods. Responses made during blackouts were not reinforced. Since the stimulus lights were the only source of illumination in the chambers, little or no responding occurred during blackout periods.

Discrimination training began on the 11th day for each subject in each pretraining condition. A 550-nm light was the positive stimulus (S+), and a black vertical line on a white surround was the negative stimulus (S-). Discrimination training was adminis-

tered in three consecutive 10-min 50-sec blocks. Each block contained five 60-sec S+ and five 60-sec S- periods in a sequence that was randomized with the restriction that no more than three S+ or S- periods occur consecutively. Each 60-sec stimulus-on period was separated from the preceding period by a 5-sec blackout. Interdimensional discrimination training was continued until each subject attained a criterion of at least 90% of all responses in a given block being emitted in the presence of the positive stimulus. Upon reaching the criterion, each subject in a given pretraining group was assigned at random to either an immediate or a 24-h delay test group, resulting in six groups of eight subjects each.

Following attainment of the discrimination criterion, all subjects were tested for wavelength generalization in extinction. The subjects in the immediate test condition each received one additional 60-sec S+ period after reaching criterion, during which the apparatus was prepared for administering the testing procedure. A 5-sec blackout period separated the last training trial from the first test stimulus presentation. For the 24-h delay groups, testing was begun 5 sec after each subject was placed in the blacked-out chamber.

The test consisted of six random series of 11 chromatic stimuli ranging from 490 to 610 nm in 10-nm steps, omitting 500 and 600 nm. Stimulus-on periods of 30 sec duration were alternated with 5-sec blackout periods.

Results

The pretraining conditions had a substantial effect on the time required to achieve the discrimination

Figure 1. Mean absolute generalization gradients in immediate and 24-h delay tests following interdimensional discrimination training in groups with three different pretraining conditions.

Figure 2. Mean relative generalization gradients in immediate and 24-h delay tests following interdimensional discrimination training in groups with three different pretraining conditions.

criterion. The pretrain S+ group required a mean of 15.0 min, compared with 34.4 min in the pretrain Sn group and 70.6 min in the pretrain S- group. The group differences were highly significant by one-way analysis of variance [$F(2,45)=66.1$, $p < .01$], and subsequent Duncan's multiple range tests indicated that each group mean differed significantly ($p < .01$) from the other two.

Figure 1 presents the mean absolute generalization gradients, based upon the number of responses emitted to the various test stimuli. Figure 2 presents the mean relative gradients, based upon the conversion of each subject's absolute gradient into the percentage of total responses emitted to each stimulus. As is traditional in the operant stimulus control literature, statistical analyses were performed only on the relative gradients, since there are extreme within-group differences in response rates. The major experimental result was equally apparent in both the relative and the absolute gradients.

The generalization gradients obtained in immediate testing are shown in the upper panels of the two figures, and the similarity of the gradients of the

three pretraining groups is quite remarkable. This similarity indicates that Task 2 learning (discrimination) was comparable in the three different pretraining groups. Note, however, that the gradients of the three groups were no longer comparable when testing was carried out after a 24-h delay. Although the relative gradients of the pretrain S+ and pretrain Sn groups were only slightly flattened by the delay, the gradient of the pretrain S- delay group was substantially flatter than all others. This conclusion is supported by statistical analysis. Analysis of variance, using percentage of responding to S+ as an index of gradient slope, indicated that the effect of pretraining condition was significant [$F(2,42) = 5.22$, $p < .01$], as was the effect of immediate vs. delayed testing [$F(1,42) = 11.38$, $p < .01$]. The interaction was also significant [$F(2,42) = 3.39$, $p < .05$]. Subsequent Duncan's comparisons indicated that the gradient slope of the pretrain S- delay test group was significantly ($p < .01$) different from the slopes of the five other group's gradients, which did not differ significantly among themselves ($p > .10$). It is important to emphasize that in this experiment generalization was measured along the dimension of the S+. The S- was on a different dimension and was not included in the generalization test, thus the flattening of gradients cannot be directly attributed to the recovery of S- responding.

Discussion

Proactive interference may be defined as a decrement in retention performance as a consequence of a training experience that precedes the acquisition of the target memory. In the present case, the target memory is conceived of as an attentive set, a non-stimulus-specific effect of the interdimensional discrimination training that constituted Task 2 for all subjects. The acquisition of that set is reflected in the relatively steep generalization gradients from all three groups in their immediate generalization tests. That set was retained quite well in two of the three pretraining groups, but forgetting (and, hence, evidence for PI) was massive in the pretrain S- group. Although this group had far less total exposure to 550 nm than did the pretrain S+ group, they had twice as much exposure as did the pretrain Sn group, so that total exposure to S+ per se cannot be critical. It is, of course, well known that the amount of PI obtained depends upon the nature of the relationship between Task 1 and Task 2 (cf. Keppel, 1968), but those relationships are so complex in the present experiment that their critical components are not readily discernible.

All three pretraining groups received single stimulus training as Task 1 and interdimensional discrimination training as Task 2; thus, prior experience with single-stimulus training per se is an insufficient condition for obtaining substantial PI. Since the pre-

train S+ and pretrain Sn groups both showed little evidence of PI, the use of the same training stimulus in both tasks is also insufficient. It was necessary that the pretraining stimulus become the subsequent S- for the PI effect to be obtained. This condition is a reversal of sorts; that is, an S+ becomes an S- when a new S+ is introduced.

Let us assume that, at the time of the generalization test, the subject performs a backward scan of its training experiences. In the immediate test condition, the recency effect is so strong that only the memory of the immediately preceding training experience is revealed. Consider the case of a subject in the pretrain S- delayed test group. On the previous few days, responding to 550 nm was reinforced, whereas responding to the vertical line was extinguished. This experience constitutes discrimination training, which should produce an attentive set and a steep wavelength gradient. However, the weakening of this memory with time (and/or the strengthening of the prior memory) permits the subject to recall that both the 550-nm and the vertical-line stimuli have accompanied reinforcement. In this conception, there are, indeed, two conflicting memories in the pretrain S- group, one based on discrimination training experienced on the last few days of training and one based on the composite discrimination training and (single stimulus) pretraining. The composite memory, which would be relatively stronger after a 1-day delay (when the "recency" effect is weaker) is of reinforcement to both training values. If such a memory is functionally like that of the usual nondifferential training, which employs the random alternation, within a session, of two equally reinforced stimuli, the expected result would indeed be an inattentive set, as reflected in a flattened generalization gradient. A potential problem for this interpretation is the fact that the pretrain Sn group also experienced reinforcement with two different stimuli, yet they showed little evidence of PI. Note, however, that the nondifferentially reinforced stimuli for this group, green and dim white, were not the same as the stimuli used in discrimination training, green and line. Thus, there was less basis for direct conflict between the memories of the two training experiences in the pretrain Sn group than in the pretrain S- group.

Rather than pursuing the question of why PI in the retention of an attentional set was so striking in the pretrain S- group in Experiment 1, we decided that it would be more fruitful to investigate the retention of attentional sets in those training paradigms we have used to study the formation of attentional sets, that is, extradimensional "true" discrimination (TD) and nondifferential (ND) training. Although "general attention" was originally conceived of as a learning construct in the sense that it affects what is learned about subsequently experienced stimuli (cf. Thomas, 1970), experiments reported by Honig (1974) and

Turner and Mackintosh (1972) demanded an extension of this conceptualization by showing that extra-dimensional discrimination and nondifferential training can influence (in the expected direction) the control exercised by a previously experienced stimulus. This finding is not inconsistent with the basic assertion that a TD or ND training experience will produce some effects that are nonspecific with regard to the training stimuli or the stimulus dimension involved in that training. Presumably, the memory of a training experience will contain elements of both stimulus-specific and nonspecific effects, both of which may be altered by the incorporation of subsequent information. In other words, in an extradimensional training paradigm, the subject acquires knowledge of the value of the training stimulus on the test dimension and also acquires an attentive or inattentive set via orthogonal TD or ND training. The sequence of these two kinds of learning is unimportant. Knowledge of the value of the dimensional stimulus provides the vehicle through which the development and retention of the attentional set can be manifested. The opportunity to provide training on the dimensional stimulus first allows us strict control of the time between the establishment of an attentional set and the test for its retention.

A particularly intriguing test of the validity of these notions is possible for the case in which responding is first established via single-stimulus training with a wavelength value, then TD training is given with two line angles, followed by brief ND training with the same line angles. An immediate wavelength generalization test should reveal a recency effect, that is, a relatively flat gradient. In a delayed test, however, the relative strengthening of the memory of earlier TD training should result in a *sharper* gradient for this group. For a control group that never experienced TD line-angle training prior to their ND line angle training, no such sharpening of the gradient with a delay is expected because there would be no earlier conflicting memory to be strengthened. In this (control) group, one might expect a flattening of the gradient to the extent that a 24-h delay between training and testing leads to uncertainty about the appropriate response tendency and possibly some additional forgetting of the value of the original wavelength training stimulus. Such effects would occur in the experimental group as well, but in that group they would be opposed by PI from prior TD training.

Experiment 2 followed the design just described, with one exception. Pilot work had indicated that control subjects exposed to line angle stimuli for the first time after brief wavelength training followed by an 8-day ("hold") period without training responded very sporadically during their single ND session with the two line angles. In order to assure stable responding and thus exposure to the ND contingency, all

subjects were given 2 days of single stimulus training with a 60-deg line interspersed during initial wavelength training and prior to the inception of the TD or hold phase of the experiment.

EXPERIMENT 2

Method

Subjects. The subjects were 64 experimentally naive domestic pigeons obtained from a local supplier. All subjects were maintained at 70%-75% of their ad-lib weight for the duration of the experiment. The birds were housed individually in a colony room on a 16/8-h light/dark cycle with continuous access to grit and water.

Apparatus. The experimental chambers were four identical sound-attenuating ice chests with interior dimensions of 31 × 35 × 36 cm. Each chamber had a wire-mesh floor, and one wall constituted an "intelligence panel" constructed of aluminum with a 2.2-cm-diam translucent plastic response key centered 20 cm above the floor. Approximately 15 cm directly below the key, a 6.3-cm square aperture allowed access to mixed grain from a solenoid-operated hopper. Behind the front wall, directly to the right of the aperture, was a 10-cm-diam speaker which delivered 80-dB (SPL) white noise inside the chamber. A fan mounted behind the front wall provided ventilation as well as additional masking noise.

Stimuli were projected onto the keys by Industrial Electronic Engineers in-line display cells (No. 10-01-3040-1815) mounted .64 cm behind the keys. Stimuli of nominal peak transmitted wavelengths 490, 538, 555, 576, and 606 nm were produced by Kodak Wratten filters 75, 74, 99, 73, and 72 B, respectively. White lines with angular orientations of 30 and 60 deg (from horizontal) were 3 mm wide × 21 mm long and bisected the otherwise dark keys. The filters were transilluminated by individual G.E. miniature lamps (No. 1815) drawing 200 mA current. The only other source of illumination in the chambers was provided by an ESB 12-V lamp situated above the food hopper and illuminated for the entire duration of each hopper presentation. Programming of scheduled events was accomplished by standard relay circuitry located in an adjoining control room.

Procedure. Experiment 2 was run in two procedurally identical replications. In each replication, 32 pigeons were assigned randomly to one of four groups, which will hereafter be identified according to their training experience and the time of their generalization test: TD-immediate, TD-delay, hold-immediate, and hold-delay.

On Day 1, all pigeons were hand-shaped to keypeck for access to mixed grain. Following continuous reinforcement of the first 20 responses, the next 30 reinforcers were earned on a gradually increasing VI schedule of reinforcement. In the following two daily sessions, pigeons were allowed to earn 50 reinforcers by responding to the constantly lit key on the same gradually increasing VI schedule. During Sessions (days) 4-7, the subjects received 32 1-min trials consisting of a 50-sec stimulus-on period VI 60-sec reinforced, followed by a 10-sec blackout, during which the key was not lit and responses were not reinforced. In Sessions 1, 2, 4, 5, and 7, the key was illuminated with a 555-nm (green) light; in Sessions 3 and 6, the key stimulus was a white 60-deg line on a dark surround.

Discrimination training (TD) was initiated on the 8th day. Subjects in the TD-immediate and TD-delay groups were given eight daily sessions of successive discrimination training in which the 60-deg line was the positive stimulus (S+, VI-30 reinforced) and a 30-deg line was the negative stimulus (S-, extinguished). Training was administered in four consecutive 8-min blocks, for a total of 32 min daily. Each block contained four 50-sec S+ and four 50-sec S- periods in a random sequence with the restriction that no more than three S+ or S- periods occur consecutively. Each 50-sec stimulus-on period was separated from the preceding period

by a 10-sec blackout. The subjects in the hold-immediate and hold-delay groups were maintained in their home cages without training during the 8 days in which the TD subjects were given discrimination training.

On the day following the TD subjects' eight session of TD training, all subjects were given one session of nondifferential (ND) training on the 60- and 30-deg line angles. This training was similar to the TD subjects' discrimination training, except that pecks to either key stimulus were reinforced on a VI 60-sec schedule. Following ND training, all subjects were tested for wavelength generalization in extinction. Following their last ND training trial, the subjects in the TD-immediate and the hold-immediate groups received a 15-sec blackout, during which the apparatus was prepared for administering the testing procedure. For the TD-delay and hold-delay groups, testing was begun 24 h after their ND training.

The test consisted of five chromatic stimuli of 490, 538, 555, 576, and 606 nm, respectively. Stimulus-on periods of 30-sec duration were alternated with 10-sec blackouts. Each generalization test comprised 12 different random series of the five test stimuli, resulting in a 39-min 50-sec test sequence.

Results and Discussion

One subject in the hold-immediate group in Replication 1 and one in the hold-delay group in Replication 2 were eliminated from the study for failure to respond on more than half of the trials during the single ND session. One subject in the TD-immediate group, Replication 1, and one subject in the hold-immediate group, Replication 2, were eliminated for failure to respond enough to give reliable generalization gradients; that is, they responded less than 30 times and in fewer than 6 of the 12 blocks of test stimuli. Thus, for purposes of statistical analysis, there were 14 subjects in the hold-immediate group, 15 in the hold-delay and TD-immediate groups, and 16 in the TD-delay group.

The TD subjects learned the TD contingency during eight sessions of TD training. On the last session of TD training, these subjects gave 77.5% of their total responses to the 60-deg (S+) line. Furthermore, the ND contingency was taking hold during the one session of ND training that these subjects eventually received. During the last block of eight ND trials, the TD subjects made 70% of their total responses to the 60-deg line. This figure is significantly lower than the 77.5% emitted during the previous (TD) session [$t(30) = 2.73$, $p < .02$], and it is an underestimate of the effect of the ND contingency. Unfortunately, no data were taken separately for the last block of eight TD trials, but the percentage of responses given to the S+ was almost certainly higher than that for the session taken as a whole; thus, the decrease in the percentage of responses given to 60 deg during the ND session was probably somewhat greater than indicated.

The subjects in the hold groups gave approximately half of their responses to each line angle during their one ND session. There is, of course, no way to tell from this level of performance whether the ND contingency was controlling these subjects' behavior,

since similar performance would have resulted from a failure to note that there were two line angles rather than one.

The data of primary interest are the wavelength generalization gradients. The absolute generalization gradients are shown in Figure 3, and the relative gradients, in Figure 4. The results were similar for both absolute and relative gradients, and, as in Experiment 1, statistical analysis was restricted to the relative gradients.

For each gradient, the percentage of total responses emitted to 555 nm was used as an index of generalization slope, and a three-way analysis of variance was carried out with training condition (TD vs. hold), test condition (immediate vs. delay), and replication as factors. The two replications gave entirely comparable results ($F < 1$ for main effect of replication and interactions involving replication). The major prediction of this study was that the effect of a delay between training and testing would depend upon the prior training experience. Subjects without prior TD experience should show the typical flattening of generalization slope with time, whereas in subjects with prior TD experience this effect should be overcome and a sharpening of generalization slope should be obtained. Figures 3 and 4 indicate that this did indeed occur, and the difference in the effect of delay is revealed statistically in the significant interaction between training and test conditions [$F(1,52) = 9.81$, $p < .01$]. The reliable interaction does not, of course, indicate that the sharpening of generalization slope with delay in the TD group is, taken by itself, a statistically significant effect. Although this sharpening was observed in both replications, it did not quite reach conventional significance (Duncan's multiple range test, $.05 < p < .10$). The flattening with delay in the hold group was significant, however (Duncan's multiple range test, $p < .05$), and this flattening has been obtained in every published study since it was first reported by Thomas and Lopez (1962). This pattern of statistical results indicates that prior TD training had the anticipated effect of overcoming the ex-

Figure 3. Mean absolute generalization gradients of TD-immediate, TD-delayed, hold-immediate, and hold-delayed groups in Experiment 2.

Figure 4. Mean relative generalization gradients of TD-immediate, TD-delayed, hold-immediate, and hold-delayed groups in Experiment 2.

pected flattening of generalization with delayed testing but that the magnitude of the effect was insufficient to produce a reliable effect in the opposite direction.

Additional analysis of the generalization gradients obtained in this study revealed some intriguing effects that will warrant further investigation. Figure 4 suggests that the immediate test gradient was flatter in the TD group than in the hold group, and the difference in percentage of total responses emitted to 555 nm was significant, according to Duncan's multiple range test ($p < .05$). One session of ND training may not have been sufficient for the ND contingency to gain control over responding in the hold groups. In any case, the contingency was clearly more effective in the TD group, presumably because of the contrast with the TD contingency previously in effect. It is remarkable that this was the case in view of the fact that the TD subjects were still discriminating between the two (now) nondifferentially reinforced stimuli at the end of the ND session whereas the hold subjects never did.

GENERAL DISCUSSION

In Experiment 1, the effect of (purported) non-differential training was to proactively interfere with the memory of subsequent discrimination training. In Experiment 2, the effect of (explicit) TD training was to proactively interfere with the memory of subsequent ND training. The two experiments taken together suggest that, from the perspective of interference theory, attentional sets may be treated fruitfully as target memories in their own right. As such, they should be subject to the same principles of retention and forgetting as are target memories of simple response tendencies. Recent studies by Thomas and McKelvie (1982) and Thomas, McKelvie, Ranney, and Moye (1981) have demonstrated the significant role of contextual retrieval cues in the retention of simple response tendencies in pigeons. It remains to

be seen whether retrieval cues will be similarly effective in reinstating attentional sets.

In the case of simple response tendencies, the fact that a generalization gradient obtained after a delay generally peaks at the value of the training stimulus indicates *what* the subject remembers. The slope of the gradient, which is flattened by the delay, gives some indication of *how well* the training stimulus is remembered. Since attentional sets are reflected in (and inferred from) differences in generalization slope, this measure enables us to identify the memory that is being manifested in performance. Thomas (1970) has proposed that nonspecific transfer of training provides an alternative indication of an attentional set; that is, TD subjects will show positive transfer to an orthogonal discrimination problem, whereas ND subjects will yield negative transfer. This suggests the possibility that in future research a new learning (i.e., transfer of training) measure might usefully supplement the recognition memory (i.e., generalization) measure in investigations of factors affecting the retention and forgetting of attentional sets.

REFERENCES

- BURR, D. E. S., & THOMAS, D. R. Effect of proactive inhibition upon the postdiscrimination generalization gradient. *Journal of Comparative and Physiological Psychology*, 1972, **81**, 441-448.
- EBBINGHAUS, H. *Memory: A contribution to experimental psychology* (H. A. Reiger & C. E. Bussanuis, trans.). New York: Bureau of Publications, Teachers College, Columbia University, 1913.
- HONIG, W. K. Effects of extradimensional discrimination training upon previously acquired stimulus control. *Learning and Motivation*, 1974, **5**, 1-15.
- KEPPEL, G. Facilitation in short- and long-term retention of paired associates following distributed practice in learning. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 9-11.
- KEPPEL, G. Retroactive and proactive inhibition. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.
- THOMAS, D. R. Stimulus selection, attention, and related matters. In J. H. Reynierse (Ed.), *Current issues in animal learning*. Lincoln: University of Nebraska Press, 1970.
- THOMAS, D. R., & LOPEZ, L. J. The effect of delayed testing on generalization slope. *Journal of Comparative and Physiological Psychology*, 1962, **44**, 541-544.
- THOMAS, D. R., & MCKELVIE, A. R. Retrieval of memory in the pigeon by context manipulations. *Animal Learning & Behavior*, 1982, **10**, 1-6.
- THOMAS, D. R., MCKELVIE, A. R., RANNEY, M., & MOYE, T. B. Interference in pigeons' long-term memory viewed as a retrieval problem. *Animal Learning & Behavior*, 1981, **9**, 581-586.
- TURNER, C., & MACKINTOSH, N. J. Stimulus selection and irrelevant stimuli in discrimination learning in pigeons. *Journal of Comparative and Physiological Psychology*, 1972, **78**, 1-19.
- UNDERWOOD, B. J. "Spontaneous recovery" of verbal associations. *Journal of Experimental Psychology*, 1948, **38**, 429-439.
- UNDERWOOD, B. J., & POSTMAN, L. Extra-experimental sources of interference in forgetting. *Psychological Review*, 1960, **67**, 73-95.