

A step function in pigeons' temporal generalization in the peak shift task

MARCIA L. SPETCH

University of Alberta, Edmonton, Alberta, Canada

and

KEN CHENG

Macquarie University, Sydney, New South Wales, Australia

Pigeons were reinforced for pecking a key following one signal duration (S+) but not following another signal duration (S-). The S+ and S- were 2.52 and 5.67 sec, respectively, counterbalanced across birds. Subsequent generalization tests with a range of signal durations revealed a step function, with high response rates for all durations on the S+ side of the distribution, low response rates for all durations on the S- side, and an intermediate rate for the intermediate duration. A comparison group of pigeons trained with only the S+ duration showed a flat generalization function. For the discrimination-trained birds, the delay between signal termination and opportunity to respond was subsequently varied during generalization testing. A step function again appeared, and no evidence of subjective shortening over the delay was found. The overall pattern of results suggests that the birds categorized the temporal signal into two classes and retained a categorical code over the delay.

Understanding how antecedent stimuli gain control over reinforced responding is an important part of the experimental analysis of behavior. Stimulus control is reflected by changes in responding as a function of changes in some aspect of the training stimulus. Generalization (similar responding despite variation in the stimulus) and discrimination (responding differently when the stimulus is changed) are opposite ends of a continuum of stimulus control.

Guttman and Kalish (1956) conducted a classic study of stimulus control. Pigeons were intermittently reinforced for pecking a key in the presence of a monochromatic light source. The spectral stimulus used in training varied across groups of pigeons. Subsequent generalization tests with a range of spectral wavelengths on either side of and including the training value showed orderly generalization gradients, with highest response rates at the training value and decreasing rates of responding at increasingly different test values.

The nature of stimulus generalization has since received a great deal of experimental attention. One general finding has been that generalization gradients become sharper after discrimination training, in which periods of reinforced responding in the presence of a positive training

stimulus (S+) are intermixed with periods during which responding is not reinforced in the presence of a negative stimulus (S-). Moreover, intradimensional discrimination training, in which the S+ and S- differ only along the dimension on which generalization tests will be conducted, produces steeper generalization gradients than does interdimensional training, in which the S+ and S- differ in many dimensions. For example, Jenkins and Harrison (1960) gave pigeons (1) nondifferential training with a 1000-Hz tone, (2) interdimensional training with a 1000-Hz tone as S+ and silence as S-, or (3) intradimensional training with a 1000-Hz tone as S+ and a 950-Hz tone as S-. During subsequent tonal frequency generalization tests, birds given intradimensional training showed the steepest gradients, and birds given nondifferential training showed the flattest gradients.

Intradimensional discrimination training not only produces steeper generalization gradients but also can produce a *peak shift effect*, in which the location of peak responding is shifted away from the training stimulus in the direction opposite the S-. Hanson (1959) trained four groups of pigeons on a wavelength discrimination with 550 nm as the S+ and with 555, 560, 570, or 590 nm as the S-. A control group was trained nondifferentially with the S+ only. During subsequent generalization tests with stimuli ranging from 480 to 620 nm, control subjects showed a peak of responding at the S+, whereas experimental subjects showed peak responding at wavelengths below the S+. Moreover, the magnitude of the shift in the peak varied inversely with the difference between the S+ and S- values.

Subsequent research has established the reliability and generality of the peak shift effect (see Honig & Urcuioli,

This research was supported by a Natural Sciences and Engineering Research Council of Canada research grant held by the first author. The authors thank M. Johnston, D. Kelly, and S. Reid for assistance with the research. Correspondence should be addressed to M. L. Spetch, Department of Psychology, University of Alberta, Edmonton, Alberta, Canada, T6G 2E9 (e-mail: mspetch@psych.ualberta.ca), or to K. Cheng, School of Behavioural Sciences, Macquarie University, Sydney, NSW 2109, Australia (e-mail: kcheng@bunyip.bhs.mq.edu.au).

—Accepted by previous editor, Robert A. Rescorla

1981, Purtle, 1973, and Rilling, 1977, for reviews). The peak shift effect has been observed with various species, including goldfish (Ames & Yarczower, 1965), rats and guinea pigs (Thomas & Setzer, 1972), horses (Dougherty & Lewis, 1991), and humans (Doll & Thomas, 1967), and with various stimulus dimensions, including visual intensity (Ernst, Engberg, & Thomas, 1971), auditory intensity (Thomas & Setzer, 1972), line tilt (Bloomfield, 1967), floor tilt (Ricchio, Urda, & Thomas, 1966), object size (Dougherty & Lewis, 1991), relative numerosity (Honig & Stewart, 1993), and spatial location (Cheng, Spetch, & Johnston, 1997). Even in cases in which a peak shift has not occurred, generalization gradients following intradimensional discrimination training typically have been asymmetrical, with more responding to values on the S+ side of the distribution than to values on the S- side, an effect referred to as an *area shift* (Rilling, 1977).

Spence's (1937) theory of discrimination learning has provided the dominant theoretical framework for understanding the peak shift effect. According to this theory, subjects learn response tendencies to specific absolute stimulus values, and discrimination training results in an excitatory gradient forming around the S+ and an inhibitory gradient forming around the S-. The peak shift effect is thought to arise from the algebraic summation of these two opposite processes.

The peak shift effect also has been interpreted within views that assume subjects learn about relationships between the stimulus values. For example, Thomas and colleagues (e.g., Thomas, 1974; Thomas, Mood, Morrison, & Wiertelak, 1991) have interpreted the peak shift effect in terms of adaptation-level theory (Helson, 1964). They suggested that, during discrimination training, an adaptation level or frame of reference develops on the basis of the average of the stimuli experienced. Subjects learn that the S+ is X units above or below this adaptation level. If the range of stimuli presented during generalization testing is different from that in training, and if no feedback is given, then the adaptation level will shift to reflect the new range. Consequently, the code for S+ will correspond to a different stimulus value, in absolute terms, from the S+ value during training. Application of the rule to respond X units away from the adaptation level can lead to a peak shift effect. In support of this interpretation, studies with humans have found that the peak of responding depends critically on the range of stimuli presented in testing (Thomas et al., 1991). Range effects have been less forthcoming in research with animals, however, perhaps because extended training leads to less volatile adaptation levels (Thomas et al., 1991). Although it is unlikely, therefore, that the peak shift effect typically seen in animals is due to changes in adaptation level, the potential for range effects should be considered in all studies of peak shift.

The present research extended the study of stimulus generalization and the peak shift effect to the temporal dimension. A great deal of research has shown that animals are sensitive to the passage of time and can discriminate between stimuli that differ only in duration. More-

over, both discrimination and generalization processes have been investigated with temporal stimuli. However, the procedures used typically have not been designed to permit an evaluation of peak shift. For example, Church and Gibbon (1982) studied temporal generalization in rats following training in which responses were reinforced following one signal duration but not following either shorter or longer signal durations. The rats showed orderly generalization gradients, with a peak of responding at the reinforced duration. Wearden (1992; Wearden, Denovan, Fakhri, & Haworth, 1997) conducted analogous temporal generalization experiments on humans. Humans made judgments, with feedback, on whether or not a temporal duration equaled an S+ duration. Generalization curves (% "yes" as a function of duration) peaked at the S+ duration and dropped on either side of S+. In these and several other studies, the procedure used differed from those used in studies of peak shift because training included a range of S- durations, rather than just a single S- on one side of the S+.

Although not designed to investigate peak shift, two studies used a procedure appropriate for assessing peak shift. Elmsore (1971) reinforced pigeons on a variable-interval schedule for pecking a key following one duration of timeout (9 or 21 sec, counterbalanced across birds), but not following the other duration of timeout. The pigeons then received generalization tests with timeout durations ranging from 3 to 27 sec. Interestingly, the birds responded as much (or more) to the values on the far side of S+ as they did to the S+, and they showed similar rates of responding to the values on the far side of the S- as they did to the S-. Mellgren, Mays, and Haddad (1983) studied discrimination and generalization by rats of intertrial interval (ITI) duration. In one of their experiments, the rats received discrimination training with 6- and 12-min ITIs. They were rewarded for running down an alley following one of these ITI durations but not following the other (counterbalanced across subjects) and then were tested with durations ranging from 2 to 18 min. Rats trained with 12-min ITIs as the S+ ran just as fast following ITIs longer than 12 min, and rats trained with 6-min ITIs ran just as fast following ITIs shorter than 6 min. Thus, in both of these studies, values on the far sides of the S+ were treated as being at least as good as the S+ duration, and a clear peak in the response distribution was not apparent. However, in both studies, only two values on either side of the S+ and S- were tested.

EXPERIMENT 1

In Experiment 1, pigeons were trained to discriminate between 2.52- and 5.67-sec durations of a signal (house-light illumination). Pecks at a key were reinforced on a fixed-interval (FI) schedule following one duration but not following the other duration (counterbalanced across birds). During subsequent generalization tests, signals of varied duration were presented on occasional unreinforced probe trials. In the first test series, the range of test

stimuli presented was designed to be subjectively symmetrical about the S+ and S- values. Because pigeons judge pairs of different absolute durations that differ by the same ratio as being equally discriminable (e.g., Dreyfus, Fetterman, Smith, & Stubbs, 1988), the values chosen differed from each other by a constant ratio. The next two series of tests used partial ranges on the S+ and S- sides. If range effects influenced responding, peak responding might shift toward the side on which the test stimuli are found. However, we did not expect such effects because test trials were presented only occasionally, and the more frequent training trials should serve to anchor the adaptation level. Thus, our experiments were purposely designed to control for range effects, and a lack of such effects does not constitute evidence against the adaptation-level hypothesis.

Method

Subjects

The subjects were 4 adult Silver King pigeons with varied experimental histories but with no previous experience in generalization tasks. The birds were maintained at approximately 85% of their free-feeding weights by mixed grain obtained during experimental sessions and supplemental feedings of pigeon chow. They were housed in large individual cages under a 12-h light:dark cycle (lights on at 6:00 a.m.). Water and grit were freely available in the home cages.

Apparatus

Procedures were conducted in standard operant pigeon chambers, each containing three horizontally aligned pecking keys (2.5 cm in diameter). Each key was set to detect a force greater than or equal to approximately 0.25 N. Stimulus projectors mounted behind each key were used to transilluminate the keys with uniform fields of red or green. A grain feeder was centered below the pecking keys, and a lamp within the feeder was used to illuminate food presentations. A houselight was centered horizontally above the pecking keys, and a lamp shield was affixed to direct the light toward the chamber ceiling. A fan provided ventilation and background masking noise. Presentation of stimulus events and the recording of responses were accomplished with a microcomputer located in an adjoining room.

Procedure

Preliminary training. All birds received one or two preliminary training sessions with an autoshaping procedure to establish reliable pecking at the right pecking key when it was illuminated with green light and at the left pecking key when it was illuminated with red light.

Discrimination training. Trials began with presentation of green light on the right pecking key as a "warning stimulus." A peck at the warning stimulus terminated it and initiated the temporal signal, which consisted of illumination of the overhead houselight. The temporal signal remained on for 2.52 sec, on a randomly determined half of the trials and for 5.67 sec on the remaining trials. For 2 birds, 2.52 sec was designated as the positive stimulus (S+) and 5.67 sec was designated as the negative stimulus (S-). The opposite assignment was used for the remaining 2 birds. Immediately after termination of the temporal signal, the left key was illuminated with red light, and a 4-sec interval was initiated during which pecks at the red key were recorded but had no consequences. If the S+ had been presented, the first peck after this interval had elapsed terminated the keylight and was reinforced with 3 sec of access to food. If the S- had been presented, the keylight terminated without reinforcement as soon as the interval elapsed. In other words, an FI rein-

forcement schedule was in effect on S+ trials and a fixed-time (FT) extinction schedule was in effect on S- trials. Trials were separated by a 30-sec ITI during which the chamber was dark. Each session consisted of 48 scheduled trials.

All birds received 55 sessions of discrimination training prior to the first test series. For 2 birds that were still showing poor discrimination by Session 30, the duration of the FI and of the FT was reduced to 2 sec for a few sessions until discrimination accuracy improved. The FI/FT duration was then restored to 4 sec for the remainder of training. Following discrimination training, each bird received three series of tests. The birds were returned to baseline training for 20 sessions between test series.

Test Series 1: Generalization tests. The first series provided 24 sessions of generalization tests with a range of values below, between, and above the S- and S+ values. Eleven values were used: 0.5, 0.75, 1.12, 1.68, 2.52, 3.78, 5.67, 8.5, 12.74, 19.12, and 28.67 sec. These values were designed to be symmetrical around the S+ and S- (i.e., there were 4 values on either side of the 2 training values) and equally spaced on a ratio scale (i.e., as one moved from shorter to longer durations, each value was 1.5 times the duration of the preceding value). Each generalization test session included 18 reinforced S+ trials, 18 unreinforced S- trials, and 11 unreinforced test trials, one with each of the 11 test values. Procedurally, the test trials were the same as S- trials (i.e., a 4-sec FT extinction schedule followed termination of the temporal signal).

Test Series 2 and 3: Generalization tests with partial ranges. Test Series 2 and 3 each consisted of 24 sessions of generalization testing that were identical to those presented in Test Series 1, except that only 7 of the 11 test values were presented. In one series (counterbalanced across birds to occur as Series 2 or Series 3), the S+ side of the distribution was emphasized: The values presented ranged from the S- through to the 4 values on the S+ side. In the other series, the values ranged from the S+ through to the 4 values on the S- side.

Results

In all statistical tests, for Experiment 1 and for the subsequent experiments, the alpha level was set at .05.

During training, discrimination ratios were computed by dividing the number of pecks made on S+ trials (excluding the one that produced reinforcement) by the total number of pecks made on both types of trials. Averaged over the last 5 training sessions, these discrimination ratios were .96, .81, .82, and .94 for the 4 birds.

The generalization data on the three series of tests are presented in Figure 1. The data for 2 of the birds have been reflected, so that S+ is at the same location on the graph for all birds. Response rates, which differed greatly between birds, have been reexpressed as a proportion of the average response rate at S+ for each bird in each series and then averaged across birds. The pattern of results is similar in the three series, and it can be described as a step function. For durations on the S+ side, proportional response rates varied nonsystematically around 1. For durations on the S- side, response rates were uniformly low. For the duration between S+ and S-, the response rate was intermediate.

Some formal statistical tests confirm these impressions. The dependent variable analyzed was the proportional response rate shown in Figure 1: rate at each test duration divided by average rate at the S+ duration for each bird. Each series was separately analyzed by a split-plot

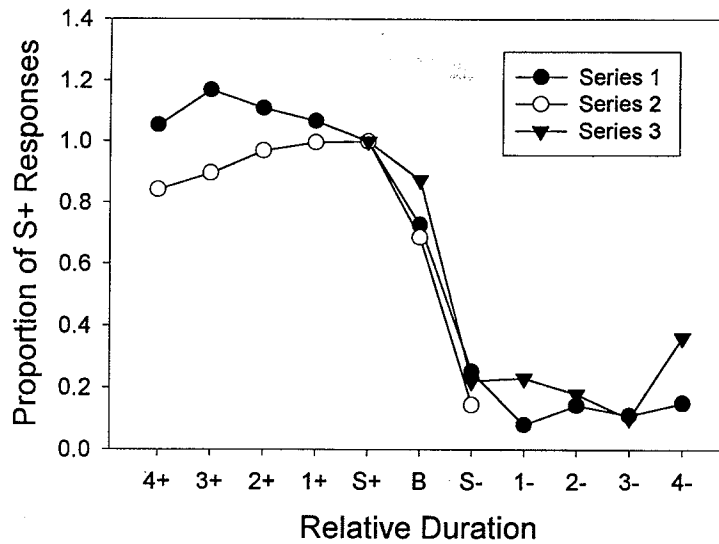


Figure 1. Generalization gradients from the three test series of Experiment 1, averaged across the 4 birds. Response rates are shown as a proportion of rates at S+ for each bird. Labels 1+ to 4+ refer to duration values on the S+ side, and 1- to 4- are duration values on the S- side. B = duration between the S+ and S-. The absolute duration of the S+ and S- differed across birds.

analysis of variance (ANOVA). The data were divided into four blocks of six sessions, which constituted one within-subjects factor. Durations were coded as in Figure 1, starting from 4 units away from S+ and ending at 4 units away from S-, making a second within-subjects factor. In addition, a between-subjects factor was the duration of S+.

The ANOVA on Series 1 revealed main effects of block [$F(3,9) = 8.14$] and duration [$F(10,30) = 27.36$], with no other significant effects. Inspecting the block effect revealed that the middle two blocks had higher average proportional rates of responding. For the partial distribution series on the S+ side, the ANOVA revealed only a duration main effect [$F(6,18) = 7.06$]. Likewise, for the partial distribution series on the S- side, the ANOVA found only a duration main effect [$F(6,18) = 14.43$]. To see whether a step function was found, ANOVAs were conducted on only the 5 durations on the S+ side and on only the 5 durations on the S- side. Two separate tests were done on Series 1 (S+ side and S- side), and one test each was done on the partial distribution series. We ignored the between-subjects factor of S+ duration, because it had no effect in the main ANOVAs. For the partial series, we dispensed with the block factor as well, since it had no effect. These ANOVAs revealed no effects. Thus, proportional response rates did not differ significantly across values on the S+ side or across values on the S- side, although they differed between the S+ and S- sides.

Figure 1 also shows that, in Series 2, proportional response rates appeared lower than corresponding rates in

Series 1. However, the differences were not statistically significant.

A step function suggests two states: one of a high rate of responding and one of a low rate of responding. One good test of the two-state hypothesis is to examine the distribution of rates of responding at Duration B, which is on average intermediate in rate between the rates found on S+ and S- tests. To do this, we examined individual tests with S+, B, and S- durations across all three series. For each bird, the average number of pecks on S+ tests was calculated, and then individual tests for all three durations were binned in terms of proportions of this S+ average, in 20% bins. Thus, the first bin contained tests with 0%–20% of the average number of pecks on S+ tests, and so on. These relative frequency distributions were then averaged across birds. The relative frequency distribution of the number of pecks on S+, B, and S- tests are shown in Figure 2. Each individual bird showed the pattern of the average distribution. In particular, for B trials, each bird showed a double-peaked pattern, with one peak in the 0%–20% bin, and a much flatter peak around 100%.

The distribution for B looks like a weighted average of the distributions for S+ and S-. To verify this, we calculated a weighted average of the S+ and S- distributions. The weighting factor was calculated from the data (and, hence, not a free parameter), and it represents the proportional "distance" at which B lies between S+ and S-. Let a , b , and c be the average number of pecks on S-, B, and S+ tests, respectively. For each bird, the weighting factor for the S+ distribution was $(b - a)/(c - a)$,

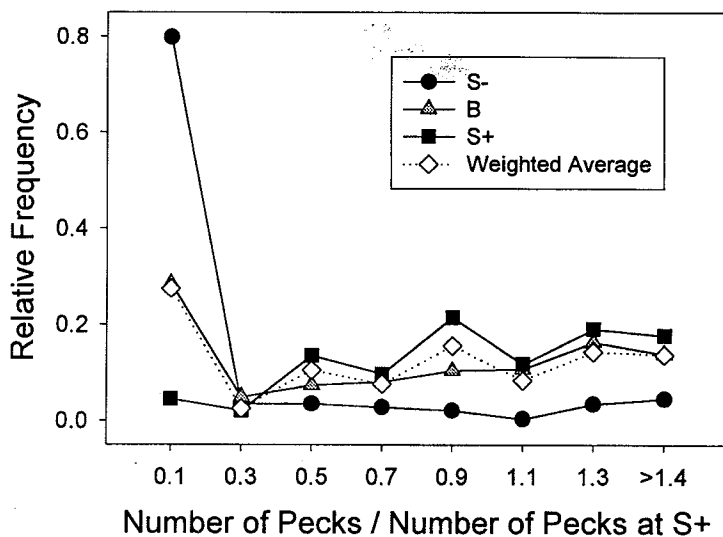


Figure 2. Frequency distribution of pecks on individual tests with the S+ duration, the S- duration, and the intermediate duration (B) in Experiment 1, as well as the weighted average of the S+ and S- distributions. The x-axis plots the number of pecks as a proportion of the average number of pecks on S+ tests. These scores are grouped into bins for the frequency determination (the first bin includes tests with proportions from 0 to .2, the second bin includes tests with proportions from .2 to .4, etc.).

whereas the weighting factor for the S- distribution was $(c - b)/(c - a)$. We took the average of the weighting factors across the 4 birds to compute the weighted average distribution of S+ and S- shown in Figure 2. As is clearly evident, this distribution resembles the distribution for B. This analysis suggests that, on any given test trial, the birds classified a B duration either as S+ or as S-.

Discussion

The results from all the series showed that the peak shift training regime produced step functions in generalization. The birds treated all durations on the S+ side as equivalently positive and all durations on the S- side as equivalently negative. As expected, no stimulus range effects were found, suggesting that the manipulations controlled for effects of adaptation level.

To our knowledge, a step function has never been reported in any study using the peak shift manipulation with other stimulus dimensions. More typically, a distinct peak away from S+ is found in the distribution (e.g., Hanson, 1959). Sometimes, as with spatial generalization, an area shift is found, in which the distribution is asymmetric but the peak remains at S+ (Cheng et al., 1997). The step function is different from either of these two patterns and suggests that durations presented in the task were recoded into two categories by the birds. For one category, the animals respond often; for the other category, they respond rarely. The duration in between was ambiguous. But analysis of the distribution of numbers of pecks on individual tests suggests that it was sometimes classified

as S+ and sometimes as S-, and not as a separate intermediate category.

EXPERIMENT 2

Experiment 1 showed that discrimination training in which one duration signals reinforcement and another does not leads to a step function in which durations on the S+ side produce uniformly high rates of responding and durations on the S- side produce uniformly low rates of responding. However, we do not yet know how this function differs from the generalization function that would emerge if pigeons were trained only with the S+. Experiment 2 was designed to provide such a comparison.

Method

Subjects

The subjects were 4 adult Silver King pigeons with varied experimental histories but with no previous experience in generalization tasks. The birds were maintained and housed as described for Experiment 1.

Apparatus

The apparatus was the same as in Experiment 1.

Procedure

Training. All birds received one or two preliminary training sessions that were identical to those given to the birds in Experiment 1. This was followed by 25 sessions of duration training that was identical to the discrimination training of Experiment 1 except that all 48 trials in each session were S+ trials. For 2 birds, the S+ duration was 2.52 sec; for the remaining 2 birds, the S+ duration was 5.67 sec.

Generalization tests. Thirty-six sessions of generalization testing were conducted, using the same 11 test durations used in Test Series 1 of Experiment 1. Each generalization test session included 37 reinforced S+ trials and 11 unreinforced test trials, one with each of the 11 test values. The procedure on test trials was the same as in Experiment 1 (i.e., a 4-sec FT extinction schedule followed termination of the temporal signal).

Results

The birds failed to show significant discrimination of the duration signal, even by the end of testing. For analysis, the data were collapsed into six blocks of 6 test sessions, and, for each bird, pecks at each duration were calculated as a proportion of pecks at the S+ duration. The data for the 2 birds with the longer duration as the S+ were reflected so that test durations were relative to the S+. An ANOVA showed no significant effect of test duration [$F(10,20) = 0.42$] or of blocks of testing [$F(5,10) = 0.70$] and no significant interaction between test duration and blocks [$F(50,100) = 0.53$]. Average functions of the 4 birds, collapsed across the six blocks of testing, are shown in Figure 3. One of the plotted functions shows mean pecks as a proportion of S+ pecks for each bird. The other function plotted is the mean absolute number of pecks per trial. Standard errors of the mean are also plotted to show that between-subject variability was very high. Neither function showed any tendency to respond more to the S+ duration than to the other durations.

Inspection of absolute rates of pecking suggested that the birds in Experiment 2 differed from those in Experiment 1 primarily in their rate of pecking on the S- (or

hypothetical S-) side of the distribution. To quantify this, we calculated the average rate of pecking across the 5 test durations on the S+ side of the distribution and across the 5 durations on the S- side of the distribution for each bird during the first block of testing. The birds in Experiment 1 pecked at a mean rate of 103.9 pecks/min for durations on the S+ side and 9.3 pecks/min for durations on the S- side. The birds in Experiment 2 pecked at a mean rate of 73.6 pecks/min for durations on the S+ side and 80.8 pecks/min for durations on the S- side. An ANOVA revealed a significant main effect of side [$F(1,6) = 9.39$], but not of group [$F(1,6) = 0.31$], and a significant interaction between group and side [$F(1,6) = 12.70$]. Post hoc comparisons (Newman-Keuls) showed a significant difference between groups in rate of pecking on the S- side but not in rate of pecking on the S+ side. In addition, rate of pecking differed across the two sides for the birds in Experiment 1 but not for the birds in Experiment 2. Thus, the main effect of training with an S- appeared to be a suppression of pecking to values on the S- side of the distribution.

Discussion

The difference between the generalization function obtained in Experiment 2 and the ones obtained in Experiment 1 indicates that the step function obtained in Experiment 1 was a result of discrimination training between an S+ and an S-. The effect of discrimination training with the S- appears to have been a suppression of responding to all values on the S- side of the distribution. It is

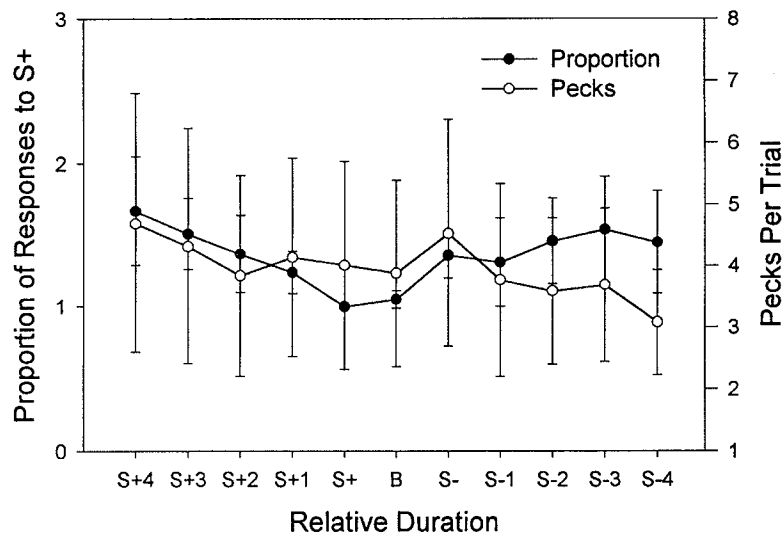


Figure 3. Generalization gradients averaged across the 36 test sessions and the 4 birds in Experiment 2. Functions are shown for both absolute response rates and relative response rates (rates as a proportion of rates at S+ for each bird). For comparison with the functions obtained in Experiment 1, the duration values are labeled as though the birds had been trained with an S-, even though training included only the S+ duration. Labels 1+ to 4+ refer to duration values on the S+ side, and 1- to 4- are duration values on the hypothetical S- side. B = duration between the S+ and hypothetical S-. For 2 birds the S+ was 2.52 sec; for 2 birds, it was 5.67 sec.

interesting that the generalization function remained flat, even by the end of testing, in the birds trained with only the S+. Although a lack of stimulus control by tonal frequency was reported by Jenkins and Harrison (1960) in birds that were previously trained with the S+ frequency only, Cheng et al. (1997) found that pigeons trained with only the S+ in a spatial task showed strong stimulus control by spatial location even on the first few generalization tests. It may be that location is a particularly salient stimulus dimension for pigeons.

EXPERIMENT 3

Experiment 3 examined whether lengthening the delay between termination of the duration signal and opportunity to respond would systematically alter the generalization function obtained following discrimination training. When pigeons are trained in a choice matching-to-sample task to choose one stimulus (e.g., red key) following a short sample and a different stimulus (e.g., green key) following a long sample, asymmetrical retention functions typically occur in response to increases in the delay between the signal and choice opportunity (e.g., Fetterman, 1995; Grant & Spetch, 1991; Spetch & Wilkie, 1982, 1983). Specifically, accuracy following short samples typically stays reasonably high, whereas accuracy following long samples decreases rapidly, often to below chance levels. In other words, after a long retention interval, pigeons appear to classify a long sample as having been short. Spetch and Wilkie (1983) interpreted this finding in terms of a subjective shortening process in which the remembered duration of the signal systematically decreases over the delay interval. Evidence consistent with a subjective shortening process was also found when gaps were inserted in a duration signal in the peak procedure (Cabeza de Vaca, Brown, & Hemmes, 1994).

Evidence of subjective shortening appears to depend, however, on the way in which duration is coded in memory. For example, no evidence for subjective shortening has been seen with successive delayed-matching-to-sample tasks (Grant & Spetch, 1991; Spetch & Grant, 1993; Spetch, Grant, & Kelly, 1996; Wasserman, DeLong, & Larew, 1984), in which only a single test stimulus follows the sample and pecks are either reinforced or not reinforced depending on the combination of sample and test stimulus presented. Grant and Spetch (1991) suggested that, whereas the standard choice task seems to promote retention of duration in terms of an analogical and retrospective representation, the successive task might promote retention in the form of a categorical or prospective code. Coding the duration categorically or prospectively would preclude the operation of a subjective shortening process. Other evidence consistent with this notion is provided by studies showing that the choose-short effect that typically occurs in the choice matching-to-sample task can be eliminated by using a many-to-one mapping arrangement in which more than one sample stimulus is associated with each choice stimulus (Grant & Spetch, 1993; Santi, Bridson, & Ducharme, 1993). Many-to-one

mapping arrangements appear to induce a common-coding process in which samples associated with the same choice stimulus activate the same code, either in the form of a prospective code or in the form of an intermediate code (e.g., Grant & Spetch, 1994; Urcuioli, Zentall, Jackson-Smith, & Steirn, 1989). Again, this recoding of stimulus duration into a nonanalogical form would preclude the operation of a subjective shortening process (see Grant, Spetch, & Kelly, 1997).

The task used in the present research entailed a simple, rather than a conditional discrimination, and, hence, the subjects could determine prior to test stimulus onset whether reinforcement would or would not be available. This task therefore seems particularly likely to promote retention in terms of a prospective code or an outcome expectancy, rather than in terms of an analogical retrospective representation of the signal duration. The form of the generalization function in Experiment 1, a step function, also suggests the coding of durations into two categories. We therefore expected that generalization functions following a retention interval would show no evidence of a subjective shortening process.

Method

Subjects and Apparatus

Experiment 3 was conducted immediately after completion of Experiment 1, and the same subjects and apparatus were used.

Procedure

Prior to the first retention test series, all birds received 20 training sessions with the baseline procedure described in Experiment 1.

Delay tests. This test series included 30 sessions of generalization testing in which delays were sometimes interposed between termination of the temporal signal and onset of the red pecking key. During delays, all stimuli were off, and the chamber was dark, thereby duplicating the stimulus conditions of the ITI. Generalization tests were conducted at each of three delay intervals: 0, 5, and 10 sec. All baseline S+ and S- trials (18 of each per test session) were conducted with the 0-sec delay. Each session provided 11 test trials, one at each of the 11 signal durations used in Test Series 1 of Experiment 1. The delay used for these generalization test trials varied across sessions but was the same for all tests conducted within each session. Each delay was in effect for a total of 10 sessions. Each delay was presented once in each 3-session block, but the order was random within blocks. On all test trials, the red pecking key stayed on for 4 sec, and then terminated without reinforcement.

Retraining with 5-sec delays. This phase consisted of 110 sessions of retraining on the baseline discrimination procedure, but with a 5-sec delay interval on every trial. By the end of this extensive retraining, 2 birds showed excellent discrimination (mean ratios of .89 and .96 over the last 5 sessions), whereas 2 birds showed only moderate discrimination (mean ratios of .74 and .63 over the last 5 sessions).

Delay tests with 5-sec baseline. This test phase was identical to the first delay test series except that the delay on baseline S+ and S- trials was 5 sec rather than 0 sec. The delay on test trials was again 0, 5, or 10 sec.

Results

To determine whether stimulus control by duration declined across test delays, we assessed whether the pigeons responded more after durations on the S+ side of the distribution than after durations on the S- side of the dis-

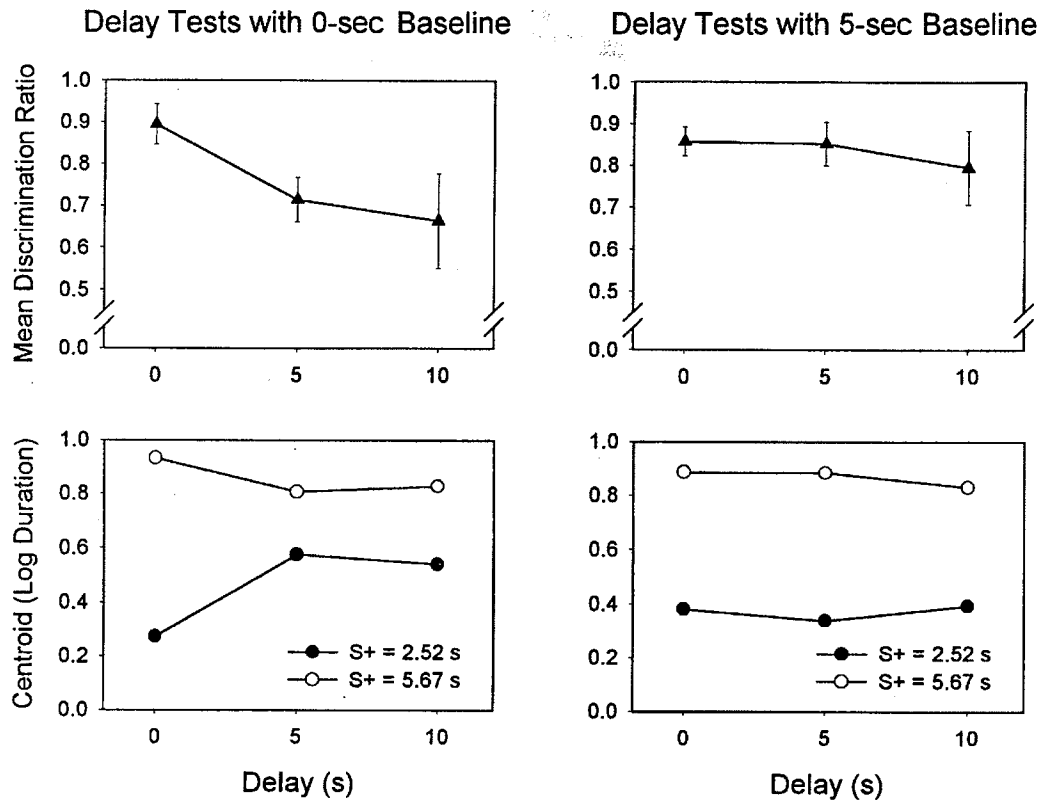


Figure 4. The top panels show mean discrimination ratios at each test delay during the first test series (0-sec baseline) and the second test series (5-sec baseline) in Experiment 3. The bottom panels show calculated centroids of the response distribution as a function of test delay interval for the two delay test series, plotted separately for the birds trained with different S+ durations.

tribution on test trials at each delay for the two test series. Discrimination ratios were calculated by dividing the total number of pecks after the 5 durations on the S+ side (including the S+ duration) by the sum of pecks after the 5 durations on the S+ side and the 5 durations on the S- side (including the S- duration). Lack of control by duration would yield a ratio of .5, whereas maximum control by duration would yield a ratio of 1.0. These discrimination ratios are presented in the top panels of Figure 4. During the first delay test series, control by duration declined across delays, and discrimination was not significantly above .5 at the longest test delay [$t(3) = 1.45$]. During the second delay test series, however, control by duration remained strong and was significantly above .5 even at the longest test delay [$t(3) = 3.33$].

To measure subjective shortening, we need a measure of central tendency for the response rates across durations. To this end, we treated a response distribution for a subject in a condition as a frequency plot, with the rate of response (or, equivalently, number of pecks) as a count for the frequency at that duration. We conducted a log transformation on the durations in order to make the bins equally spaced. Two measures of central tendency were then calculated on this pseudodistribution. A mean or

“centroid” of the distribution was calculated by multiplying each log-transformed duration by number of pecks at that duration and then dividing by the total number of pecks. A median was calculated by determining the bin that contained the median peck, treating the number of pecks as a flat function within each bin.

The centroids across delay conditions of both test series are presented in the bottom panels of Figure 4. The medians produced similar results, and they are not presented. Split-plot ANOVAs were conducted on both measures. Each analysis had the between-subjects factor of duration of S+. Within-subjects factors were baseline delay interval and test delay interval. The ANOVA on medians yielded only a main effect of duration of S+ [$F(1,2) = 54.30$], whereas the ANOVA on centroids revealed a main effect of duration of S+ [$F(1,2) = 40.81$] and a three-way interaction [$F(2,6) = 7.66$]. The three-way interaction is apparent in Figure 4.

The hypothesis of subjective shortening predicts a main effect of test delay interval, with centroids or medians lengthening with increasing delay. That is, a subjective shortening processes would mean that the test durations seemed shorter than they really were following longer delays. To give a concrete example, suppose that, after a

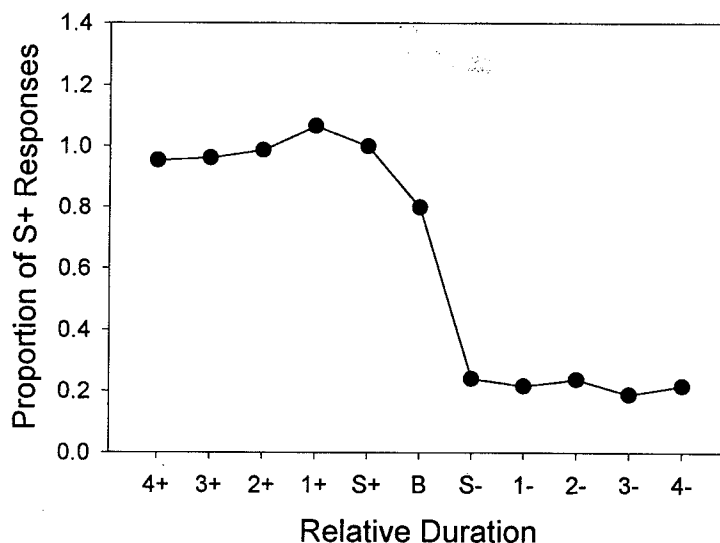


Figure 5. Generalization gradient, averaged across all test delays of the second delay test series (5-sec baseline) of Experiment 3. Response rates are shown as a proportion of rates at S+ for each bird. Labels 1+ to 4+ refer to duration values on the S+ side, and 1- to 4- are duration values on the S- side. B = duration between the S+ and S-.

lengthy delay, a 5.67-sec signal is subjectively shortened so that it seems more like the 2.52-sec signal. Birds trained with 5.67 sec as the S+ should not respond at high rates after this subjectively shortened signal but instead should require a much longer test value before responding at high rates. On the other hand, birds trained with 2.52 sec as the S+ should respond at high rates after this subjectively shortened signal and should require a much longer signal than normal before responding at low rates. In either case, then, a rightward shift in the step function would be found, and the median and centroid of the distribution would therefore occur at a longer value. This was not found. In tests with 0-sec as the baseline delay, test delay produced opposite effects for the birds with a short and a long S+: The centroid lengthened for the birds with short S+, but it shortened for the birds with long S+. In the baseline 5-sec delay phase, the centroids were similar across delays.

Since the distributions in the baseline 5-sec delay phase looked similar across delays, we plotted the response distribution, averaged across test delays, to check the form of the distribution (Figure 5). As in Figure 1, the dependent measure was the rate of responding at each duration relative to the average rate of responding at S+, and the data for 2 of the birds have been inverted so that S+ is at the same location on the graph for all. Again, a step function was clearly evident. A split-plot ANOVA with duration of S+, test delay, and duration of stimulus as factors found only a main effect of duration [$F(10,30) = 46.80$]. We also conducted an ANOVA on only the 5 durations on the S+ side and an ANOVA on only the 5 durations on the

S- side, ignoring S+ duration and test delay. As in Experiment 1, these found no significant duration effects.

The same individual-test analysis conducted in Experiment 1 was carried out on the distributions of number of pecks on S+, B, and S- tests for the 5-sec baseline delay phase. The results (Figure 6) again show that the distribution for B tests looks like a weighted average of the distribution for S+ and S- tests, thus supporting the hypothesis of coding in two categories.

Discussion

In Experiment 3, two results of note were found: (1) varying the retention interval at testing did not systematically lengthen the central tendency of the response distribution, and (2) the distribution of response rates across durations continued to be a step function. The first result rules against the hypothesis of subjective shortening with delays. It suggests that durations in the task were not coded in an analog fashion but were transformed into some other categorical code. This code might be some categorical equivalent of "long" and "short," it might be couched in terms of the response to make, equivalent to "peck often" and "peck rarely," or it might represent an outcome expectancy, equivalent to "food" and "no food." The conclusion that durations were categorically coded is likewise consistent with the second result: The step function suggests that stimulus durations were coded into two categories. This coding scheme appears to have been stable over 30 test sessions or 30 tests at each of the stimulus durations.

We think that the pattern of results in the baseline 0-sec delay phase is also interpretable. What happened was that,

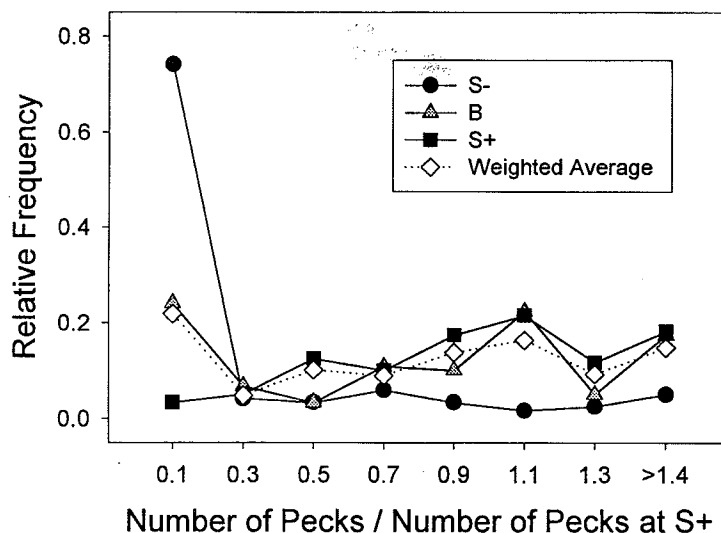


Figure 6. Frequency distribution of pecks on individual tests with the S+ duration, the S- duration, and the intermediate duration (B), as well as the weighted average of the S+ and S- distributions in Experiment 3. Data are from all test delays of the second delay test series (5-sec baseline). The x-axis plots the number of pecks as a proportion of the average number of pecks on S+ tests. These scores are grouped into bins for the frequency determination (the first bin includes tests with proportions from 0 to .2, the second bin includes tests with proportions from .2 to .4, etc.).

with test delays, the central tendency for the pseudo-distributions drifted toward the middle of the stimulus range in all animals (with long or short S+). This is what would be expected given that the test delays brought on a loss of stimulus control by duration. This loss of control may reflect the fact that the birds had never been reinforced after a delay interval. In contrast, with training under a baseline delay of 5 sec, the birds learned that food might be available with or without delays (the latter from earlier training experience). Stimulus control by stimulus duration is thus regained on all test delays.

EXPERIMENT 4

The step function we obtained using the peak shift procedure with duration appears to differ from the generalization functions typically obtained with other stimulus dimensions. However, it is possible that the emergence of a step function reflects some unique aspect of our procedure rather than our use of duration as the stimulus dimension. For example, in most demonstrations of peak shift, generalization functions are obtained during sessions conducted completely in extinction, rather than by presenting intermittent probe tests. In addition, the response window is typically much larger than the 4-sec window we used here. It seemed unlikely that these procedural variables were responsible for the emergence of a step function, however, because we used intermittent probe testing and a reasonably short time window (8 sec per trial) in a peak shift study with the spatial dimension, and we did not observe a step function (Cheng et al., 1997). Nevertheless, we sought to demonstrate that the

step function with duration is robust across variations in the test procedure. Accordingly, we lengthened the response window slightly (from 4 to 6 sec), and we presented generalization tests during a block of sessions conducted in extinction.

Method

Subjects and Apparatus

The 4 subjects from Experiments 1 and 3 participated in Experiment 4 several months following their completion of Experiment 3. The apparatus was the same as that used in Experiments 1 and 3.

Procedure

Initially, each bird received several retraining sessions with the same baseline discrimination procedure and with the same contingencies as in Experiment 1 except that the response window (FI on reinforced trials and FT on nonreinforced trials) was 6 sec rather than 4 sec. Each bird remained on discrimination training until discrimination ratios were .80 or higher for 3 consecutive sessions. For the 1st bird we retrained, we subsequently raised the response window to 8 sec, but the bird's accuracy fell to near chance level and did not recover within 10 sessions. Consequently, we decreased the response window back to 6 sec until the accuracy criterion was again reached. The remaining birds were exposed only to the 6-sec window. Each bird then received 5 sessions of generalization testing conducted in extinction. During these sessions, each of the 11 test durations (0.5, 0.75, 1.12, 1.68, 2.52, 3.78, 5.67, 8.5, 12.74, 19.12, and 28.67 sec) was presented four times, and an FT 6-sec extinction schedule was in effect on all trials.

Results and Discussion

Because 2 birds had almost completely extinguished pecking by the 4th session, only the first 3 extinction test sessions are presented. The data from the 3 extinction ses-

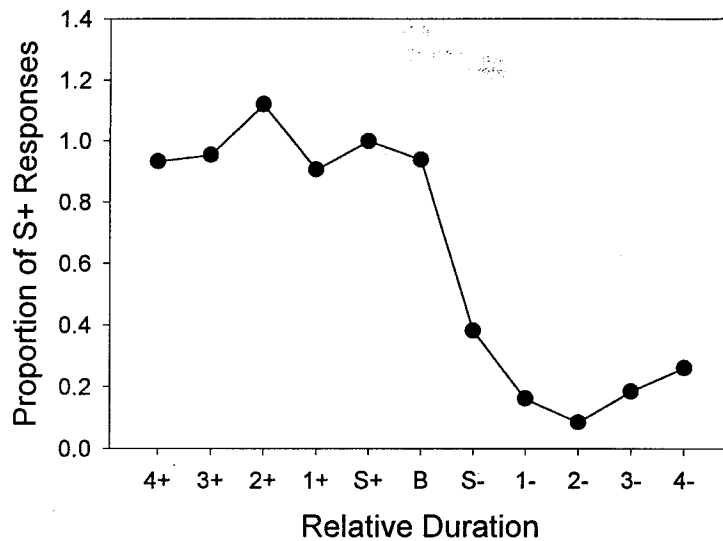


Figure 7. Generalization gradient, averaged across all 4 birds and the three extinction test sessions in Experiment 4. Response rates are shown as a proportion of rates at S+ for each bird. Labels 1+ to 4+ refer to duration values on the S+ side, and 1- to 4- are duration values on the S- side. B = duration between the S+ and S-.

sions were collapsed, and response rates at each test duration were calculated as a proportion of rates at the S+ duration. As in the previous experiments, the data from 2 birds were reflected so that scores are plotted as a function of relative duration. Mean proportional response rates are shown in Figure 7. An ANOVA revealed a significant effect of duration [$F(10,30) = 9.75$]. Subsequent ANOVAs were conducted on the 5 durations on the S+ side and on the 5 durations on the S- side. Neither of these was significant [S+ side, $F(4,12) = 0.34$; S- side, $F(4,12) = 1.53$], consistent with a step function. Thus, the results were similar to those found in Experiments 1 and 3, despite the use of a different testing procedure and a slightly longer response window.

GENERAL DISCUSSION

Two main results emerged from these experiments using a peak shift manipulation on temporal durations. First, the generalization function was a step function. The S+ side had high rates of responding, the S- side had low rates of responding, and the duration intermediate had an intermediate rate. Second, changing the duration of the delay period between the signal offset and the opportunity to respond produced no systematic changes in the patterns of generalization. We interpret the first result to mean that the birds treated durations as one of two categories, equivalent to "go" and "no go," or "food coming" and "food not coming." The subsidiary analysis of the number of pecks on individual tests supports this claim. We interpret the second result to mean that the code held during the delay period was not an analog code of temporal duration but instead was some type of categorical

code. In this section, we compare this peak shift task to other tasks used in research on interval timing and then discuss the nature of temporal coding and mechanisms of generalization that can produce the generalization gradients obtained.

Peak Shift Task

The peak shift manipulation used in our experiments is a variant of the generalization paradigm. The difference is that, with peak shift, a preponderant amount of S- training with a single value on one side of S+ is given. In standard generalization tasks, S- s on both sides of S+ are presented, equally often on each side. Temporal generalization in rats and humans has been modeled with variants of scalar expectancy theory (SET; Church & Gibbon, 1982; Wearden, 1992). In these models, the subject compares the clock reading of the presented duration with a remembered duration of S+ retrieved from reference memory. If the two values are within some threshold of one another, the stimulus is classified as S+, otherwise as S-. The models thus explicitly incorporate a two-state classification of durations, consistent with our results. A SET model with a single reference memory distribution for our peak shift task might work as follows. From the training regime, the animal stores a distribution of S+ durations in reference memory. It also stores a distribution of thresholds for comparing the duration experienced on a trial against reference memory. On a trial, the animal retrieves a duration and a threshold from reference memory and compares the presented duration, timed by its internal clock, against the reference memory. If the timed duration is on the S+ side of the reference memory, it is assumed to be within thresh-

old and is treated as S+. If the timed duration is on the S- side of the reference memory, the animal treats it as S+ if it is within the threshold and as S- if it is beyond the threshold. We see no point, however, in actually carrying out such a quantitative exercise. Accepting the data as a step function leaves only three different data values to account for: one for S+ side, one for S- side, and one for the intermediate duration. The models in question typically have at least three free parameters.

The peak shift task also resembles the bisection or choice task. In the choice task, the subject is presented with either a long or a short stimulus duration and is asked to classify the duration as long or short by its response (e.g., Allan & Gibbon, 1991; Meck & Church, 1983; Roberts, Cheng, & Cohen, 1989). For instance, red might correspond to long, and green might correspond to short. The peak shift task is a go/no-go task rather than a choice task. The bird is required to classify by responding often or rarely. Thus, models in the SET framework accounting for choice task performance can be applied to the peak shift task as well. In these models, the clock reading of the stimulus duration is compared with two different durations retrieved from reference memory: one from a distribution representing the long standard and the other from a distribution representing the short standard. Ratio comparison is made, and the stimulus is classified as long or short according to whether it is nearer to the longer or the shorter standard retrieved from memory. For the same reason as before, it is pointless to actually carry out the quantitative exercise.

Temporal Coding

In any model for the peak shift task based on SET, then, the subject starts with an analog clock reading of the stimulus duration. This representation changes in a continuous and cumulative fashion as a function of duration. One proposal for the clock reading is in terms of the number of counts generated by a pacemaker (e.g., Gibbon & Church, 1984; Meck & Church, 1983). The clock reading is then compared with other durations from reference memory, and, on the basis of the comparisons, one of two states is determined. Duration information is recoded categorically, perhaps in terms of a response decision (e.g., "peck" or "do not peck") or an outcome expectancy ("food" or "no food"). If a delay is imposed between stimulus offset and the chance to respond, we can ask what kind of code the birds hold onto during the delay: the analog duration or the categorical code.

Extant theory suggests that an analog code decays with the passage of time (see Grant et al., 1997; Spetch & Rusak, 1992). It is as if the clock reading consists of pacemaker outputs stored in a leaky container. This means that with the passage of time, the analog code comes to represent a shorter duration, a phenomenon called *subjective shortening* (Spetch & Wilkie, 1983). Categorical codes are subject to forgetting but not to subjective shortening. In fact, it is hard to conceive of how a code such as "respond often" or "peck red key" can in any sense shorten.

We looked for subjective shortening by manipulating the delay on tests. That we did not find any subjective shortening is consistent with the notion that a categorical code was held during the delay period. Apparently, the birds make the comparison between clock reading and reference memory as soon as the stimulus duration ends.

Research on memory for temporal duration suggests that both kinds of codes may be held during the delay period, depending on the task. An analog code appears to be used on standard choice delayed-matching-to-sample tasks and on the peak procedure. In the matching-to-sample task, birds are reinforced for choosing one comparison stimulus following a short sample and another comparison stimulus following a long sample. Subjective shortening is typically found: Pigeons make more choices appropriate to the short sample as the delay increases (e.g., Fetterman, 1995; Spetch & Wilkie, 1983). Pigeons trained on a peak procedure also appear to retain an analog representation of duration. In the peak procedure, subjects are trained on a signaled FI schedule and then are tested on occasional nonreinforced trials with a very long signal presentation. If the signal is interrupted with a gap, the represented duration of a signal appears to shorten over a gap (Cabeza de Vaca et al., 1994).

A categorical code for duration, on the other hand, appears to be used on successive matching-to-sample tasks (Grant & Spetch, 1991; Spetch & Grant, 1993; Spetch et al., 1996), in which responding during the test is reinforced following some combinations of sample and test stimuli and is not reinforced following other combinations of sample and test stimuli. Pigeons also appear to retain a categorical code for duration if more than one type of sample is associated with each comparison stimulus (Grant & Spetch, 1993; Santi et al., 1993). In these tasks, pigeons do not show an increasing tendency to make responses appropriate to the shorter sample as the delay increases. Thus, the type of code for duration retained over a delay varies depending on task demands.

Research on memory for other kinds of events also suggests flexibility in coding. For example, if a pigeon has to remember whether the color of a sample stimulus was red or green, it might remember retrospectively, in terms of the color of the to-be-remembered stimulus, or prospectively, in terms of what response to make. Depending on task demands, evidence for both kinds of coding has been found (e.g., Jackson-Smith, Zentall, & Steirn, 1993; Urcuioli & Zentall, 1986; Zentall, Urcuioli, Jagiolo, & Jackson-Smith, 1989).

Why a Step Function?

One might wonder why we obtained a step function when other studies of temporal generalization typically produce peaked or graded functions. Quite possibly, the occurrence of a step function depends on specific features of the peak shift training procedure we used. In our procedure, one training duration signals reinforcement, and the other signals nonreinforcement. A comparison with reference memory can be made as soon as the stim-

ulus ends, and an outcome expectancy can be formed and retained during the test. We suspect that this may be critical for the emergence of a step function. For example, a graded and peaked function for average response rates is seen during generalization testing in the peak procedure (e.g., Cheng & Westwood, 1993; Church, Meck, & Gibbon, 1994). The peak procedure requires continuous comparison between clock time and criterion duration, so it makes sense to maintain an analog code while the stimulus is on. Temporal bisection tasks, in which subjects are trained to make one choice after a short signal and a different choice after a longer signal, also do not yield a step function. Siegel (1986) tested rats with durations outside of, as well as between, the two training durations. For values between the training durations, the probability of a long response increased systematically as a function of duration. Choice regressed toward indifference for values outside of the range. That is, subjects were less likely to choose "long" after a 16-sec signal than after the 8-sec signal used as "long" in training. This contrasts with the results of Mellgren et al. (1983), in which the rats responded at least as much to durations on the far side of the S+ as they did to the S+ duration. A difference between these studies is that reinforcement was available only following one signal duration in the training procedure used by Mellgren et al., whereas reinforcement was available for correct responses following both signal durations in the training procedure used by Siegel. Being able to form an outcome expectancy may be a critical factor in the form of coding of temporal signals (e.g., Ducharme & Santi, 1993) and, hence, in the form of the generalization function obtained.

Comparing our results with those obtained by Reynolds and Catania (1962) suggests that the number or range of durations presented in training may also be important. Reynolds and Catania trained pigeons with a single S+ duration and a range of S- durations on one side of the S+. They did not test values outside of the training range, but response rates to the training values showed a graded, rather than a step, function. It may be that exposure to a range of stimulus durations from the outset of training enhances sensitivity to the continuous nature of the temporal dimension. Thus, a step function might only appear when subjects are trained with just two stimulus durations, one that signals reinforcement and one that does not. Further research is needed to confirm these speculations.

Another question raised by this work is why temporal generalization in the peak shift task shows a step function, whereas generalization with a wide range of other stimulus dimensions in the peak shift task shows a peaked function. One possibility is that temporal stimuli are processed differently than are other types of stimuli. Certainly, the dynamic nature of temporal stimuli, with values changing constantly while the signal is on, makes the temporal dimension different from other stimulus dimensions tested in the peak shift task. Why this would lead to a step function, however, is not clear. Another possi-

bility is that the inclusion of a memory component in the temporal generalization task plays a role in emergence of a step function. In other studies of peak shift, subjects responded while the discriminative stimulus was present. This was not possible with the temporal dimension because the value changes constantly during the signal presentation. Instead, subjects responded after the signal terminated, which required working memory processes. One way to test the possibility that this memory requirement played a role in the emergence of a step function would be to add a memory requirement to the peak shift task with a more standard stimulus dimension, such as wavelength, and see whether it alters the form of the generalization function obtained.

Mechanisms of Generalization

The most thorough theorizing on generalization comes from the work of Shepard (1958a, 1958b, 1965, 1987, 1988). In Shepard's (1987) theory, generalization is seen as a task of classifying perceptually discriminable stimuli. Classification is made in terms of how to respond to them. In the typical laboratory generalization task, the classification is in terms of whether to respond or not. A mechanism for generalization is the diffusion of excitation across units representing stimuli (Shepard, 1958a, 1958b; Staddon & Reid, 1990). Imagine a linear row of units representing duration in topographic fashion, with neighbors representing similar durations. The topographic row of units constitutes a psychological scale representing duration, and it need not correspond in a linear fashion with physical duration. Suppose that, before any training, all units have zero-level excitation. An S+ trial boosts the value of the appropriate psychological unit to some level. Over the passage of time, this value diffuses across to neighboring units, so that some kind of Gaussian excitation distribution is spread over the units. The next S+ trial boosts the value of the S+ unit again, and this also diffuses over time. The generalization function can thus be conceived of as a sum of Gaussians of varying widths. Shepard (1958a) has argued that such a distribution should be exponential in shape over the psychological scale, and generalization functions across clearly discriminable stimuli obey this form (Shepard, 1958a, 1958b, 1987).

The peak shift task has S- trials as well. It is reasonable to suppose that an S- trial sets the corresponding S- unit to 0, and that this also diffuses with the passage of time. Cheng et al. (1997) showed that one such diffusion model accounts reasonably well qualitatively for area shifts in spatial generalization. The model represented the surface of a monitor topographically as a two-dimensional array. We thought that essentially the same model, transformed to a one-dimensional format to represent durations, was worth trying to see if step functions might be generated.

The model is an adaptation of Staddon and Reid's (1990) diffusion model. The psychological scale for dura-

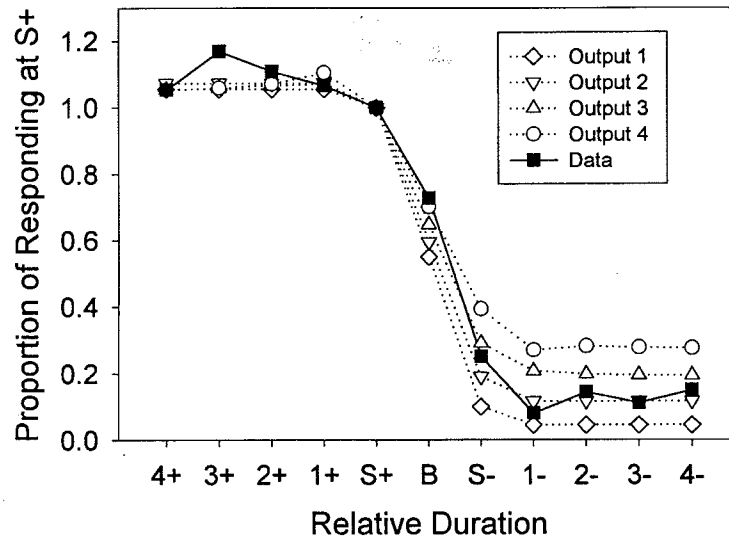


Figure 8. Outputs of a diffusion model compared to the data from Series 1 of Experiment 1. Parameters in the model were n , the number of neighbors on each side over which averaging takes place, and s , the number of steps of diffusion taken. Average absolute difference between model and data constitutes the measure of error. All the models displayed show the s that minimized error for a particular n . For Output 1, $n = 6$, $s = 177$, error = 0.089; for Output 2, $n = 5$, $s = 275$, error = 0.053; for Output 3, $n = 4$, $s = 3693$, error = 0.041; for Output 4, $n = 2$, $s = 7023$, error = 0.074. For Outputs 3 and 4, the model showed a steady state, changing minimally over many hundreds of steps.

tion is assumed to be one-dimensional and logarithmic. This means that our experimental stimulus durations are equally spaced on the psychological scale. Arbitrarily, we imposed four units between each of the units corresponding to stimulus durations. The excitation scale, arbitrarily, ranges from 0 to 1. We fed the model steps consisting of one S+ trial and one S- trial. Two things happen in a step: (1) the S+ trial sets the S+ unit to 1, and the S- trial sets the S- unit to zero, and (2) the resulting distribution is then smoothed, much as scientists might smooth data. Smoothing is one of many ways to instantiate the diffusion process. In our smoothing process, each value is averaged with its nearest n neighbors on both sides ($n = 1, 2, \dots$). Different outputs are produced depending on the smoothing parameter (n) and the number of steps (s) taken.

For a number of small n s, we found the step that produced the best fit to the data from Series 1 in Experiment 1. This series is the best for comparison, since the modeling assumes no effects of test trials. Presumably, the first series of tests best reflects this assumption. Figure 8 shows that, over a range of smoothing widths, we can find a step-like function after a suitable number of steps. Uncannily, a number of subsidiary features in the data turn up in the model. Values beyond S+ are slightly higher than S+. The value for B is closer to that of S+ than to that of S-. S- has a higher value than units beyond S-. Even though these trends in the data may not amount to statis-

tical significance, the correspondence between models and data is still striking.

The smoothing technique used was chosen not for theoretical reasons but because it was easy to do on a spreadsheet (Excel 5.1). The model's performance might be improved by changing the smoothing procedure. It might be more realistic to suppose that the averaging is weighted by the nearness of the neighbor, with nearer neighbors given higher weight than more distant ones. Variations along these lines are endless, and it is not worthwhile refining the model any further to account for such limited data. The performance of our model suggests that some diffusion model can produce step-like gradients in the peak shift paradigm.

In spatial peak shift, we modeled the effects of tests as setting values of the tested unit to zero followed by smoothing. The main effect of this is that the S+ side should drop down over tests, and the distribution should become symmetric about S+. In spatial generalization, the distribution became symmetric rapidly with testing at all locations, in a matter of 4 sessions or 12 tests per location. In temporal generalization, this process, if it happens at all, takes much longer.

Conclusions

The peak shift task produced very orderly temporal generalization gradients, but these gradients do not resemble those seen previously with a wide variety of other

stimulus dimensions. Typically, the generalization gradients seen in the peak shift task are characterized by a peak in responding at a value shifted away from the S+ in the direction opposite the S-. The present temporal generalization gradients showed no clear peak in responding but instead showed a step function, in which all values on the S+ side were treated as being equally positive and all values on the S- side were treated as being equally negative. The birds appeared to classify the durations into two categories and to treat all durations in each category as being equivalent. This is a unique result within the peak shift literature, but it is not inconsistent with models of timing, such as SET. Furthermore, we found that such a function can be generated by a variant of the diffusion model.

REFERENCES

- ALLAN, L. G., & GIBBON, J. (1991). Human bisection at the geometric mean. *Learning & Motivation*, **22**, 39-58.
- AMES, L. L., & YARCZOWER, M. (1965). Some effects of wavelength discrimination on stimulus generalization in the goldfish. *Psychonomic Science*, **3**, 311-312.
- BLOOMFIELD, T. M. (1967). A peak shift on a line-tilt continuum. *Journal of the Experimental Analysis of Behavior*, **10**, 361-366.
- CABEZA DE VACA, S., BROWN, B. L., & HEMMES, N. S. (1994). Internal clock and memory processes in animal timing. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 184-198.
- CHENG, K., SPETCH, M. L., & JOHNSTON, M. (1997). Spatial peak shift and generalization in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **23**, 469-481.
- CHENG, K., & WESTWOOD, R. (1993). Analysis of single trials in pigeons' timing performance. *Journal of Experimental Psychology: Animal Behavior Processes*, **19**, 56-67.
- CHURCH, R. M., & GIBBON, J. (1982). Temporal generalization. *Journal of Experimental Psychology: Animal Behavior Processes*, **8**, 165-186.
- CHURCH, R. M., MECK, W. H., & GIBBON, J. (1994). Applications of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 135-155.
- DOLL, T. J., & THOMAS, D. R. (1967). Effects of discrimination training on stimulus generalization for human subjects. *Journal of Experimental Psychology*, **75**, 508-512.
- DOUGHERTY, D. M., & LEWIS, P. (1991). Stimulus generalization, discrimination learning, and peak shift in horses. *Journal of the Experimental Analysis of Behavior*, **56**, 97-104.
- DREYFUS, L. R., FETTERMAN, J. G., SMITH, L. D., & STUBBS, D. A. (1988). Discrimination of temporal relations by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **14**, 349-367.
- DUCHARME, M. J., & SANTI, A. (1993). Alterations in the memory code for temporal events induced by differential outcome expectancies in pigeons. *Animal Learning & Behavior*, **21**, 73-81.
- ELSMORE, T. F. (1971). Control of responding by stimulus duration. *Journal of the Experimental Analysis of Behavior*, **16**, 81-87.
- ERNST, A. J., ENGBERG, L., & THOMAS, D. R. (1971). On the form of stimulus generalization curves for visual intensity. *Journal of the Experimental Analysis of Behavior*, **16**, 177-180.
- FETTERMAN, J. G. (1995). The psychophysics of remembered duration. *Animal Learning & Behavior*, **23**, 49-62.
- GIBBON, J., & CHURCH, R. M. (1984). Sources of variance in an information processing theory of timing. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 465-488). Hillsdale, NJ: Erlbaum.
- GRANT, D. S., & SPETCH, M. L. (1991). Pigeons' memory for event duration: Differences between choice and successive matching tasks. *Learning & Motivation*, **22**, 180-199.
- GRANT, D. S., & SPETCH, M. L. (1993). Analogical and nonanalogical coding of samples differing in duration in a choice-matching task in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **19**, 15-25.
- GRANT, D. S., & SPETCH, M. L. (1994). Mediated transfer testing provides evidence for common coding of duration and line samples in many-to-one matching in pigeons. *Animal Learning & Behavior*, **22**, 84-89.
- GRANT, D. S., SPETCH, M. L., & KELLY, R. (1997). Pigeons' coding of event duration in delayed matching-to-sample. In C. M. Bradshaw & E. Szabadi (Eds.), *Time and behavior: Psychological and neuro-behavioral analyses* (pp. 217-264). Amsterdam: Elsevier.
- GUTTMAN, N., & KALISH, H. I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology*, **51**, 79-88.
- HANSON, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, **58**, 321-334.
- HELSON, H. (1964). *Adaptation level theory*. New York: Harper & Row.
- HONIG, W. K., & STEWART, K. E. (1993). Relative numerosity as a dimension of stimulus control: The peak shift. *Animal Learning & Behavior*, **21**, 346-354.
- HONIG, W. K., & URCUIOLI, P. J. (1981). The legacy of Guttman and Kalish (1956): Twenty-five years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior*, **36**, 405-445.
- JACKSON-SMITH, P., ZENTALL, T. R., & STEIRN, J. N. (1993). Prospective and retrospective memory processes in pigeons' performance on a successive delayed matching-to-sample task. *Learning & Motivation*, **24**, 1-22.
- JENKINS, H. M., & HARRISON, R. H. (1960). Effect of discrimination training on auditory generalization. *Journal of Experimental Psychology*, **59**, 246-253.
- MECK, W. H., & CHURCH, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, **9**, 320-334.
- MELLGREN, R. L., MAYS, M. Z., & HADDAD, N. F. (1983). Discrimination and generalization by rats of temporal stimuli lasting for minutes. *Learning & Motivation*, **14**, 75-91.
- PURTLE, R. B. (1973). Peak shift: A review. *Psychological Bulletin*, **80**, 408-421.
- REYNOLDS, G. S., & CATANIA, A. C. (1962). Temporal discrimination in pigeons. *Science*, **135**, 314-315.
- RICCIO, D. C., URDA, M., & THOMAS, D. R. (1966). Stimulus control in pigeons based on proprioceptive stimuli from floor inclination. *Science*, **153**, 434-436.
- RILLING, M. (1977). Stimulus control and inhibitory processes. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 432-480). Englewood Cliffs, NJ: Prentice-Hall.
- ROBERTS, W. A., CHENG, K., & COHEN, J. S. (1989). Timing visual and auditory signals in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 23-35.
- SANTI, A., BRIDSON, S., & DUCHARME, M. J. (1993). Memory codes for temporal and nontemporal samples in many-to-one matching by pigeons. *Animal Learning & Behavior*, **21**, 120-130.
- SHEPARD, R. N. (1958a). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review*, **65**, 242-256.
- SHEPARD, R. N. (1958b). Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. *Journal of Experimental Psychology*, **55**, 509-523.
- SHEPARD, R. N. (1965). Approximation to uniform gradients of generalization by monotone transformations of scale. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 94-110). Stanford: Stanford University Press.
- SHEPARD, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, **237**, 1317-1323.
- SHEPARD, R. N. (1988). Time and distance in generalization and discrimination: Reply to Ennis (1988). *Journal of Experimental Psychology: General*, **117**, 415-416.
- SIEGEL, S. F. (1986). A test of the similarity rule model of temporal bisection. *Learning & Motivation*, **17**, 59-75.
- SPENCE, K. W. (1937). The differential response of animals to stimuli differing within a single dimension. *Psychological Review*, **44**, 430-444.
- SPETCH, M. L., & GRANT, D. S. (1993). Pigeons' memory for event duration in choice and successive matching-to-sample tasks. *Learning & Motivation*, **24**, 156-174.

- SPETCH, M. L., GRANT, D. S., & KELLY, R. (1996). Procedural determinants of coding processes in pigeons' memory for duration. *Learning & Motivation*, *27*, 179-199.
- SPETCH, M. L., & RUSAK, B. (1992). Time present and time past. In W. K. Honig & J. G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 47-87). Hillsdale, NJ: Erlbaum.
- SPETCH, M. L., & WILKIE, D. M. (1982). A systematic bias in pigeons' memory for food and light durations. *Behaviour Analysis Letters*, *2*, 267-274.
- SPETCH, M. L., & WILKIE, D. M. (1983). Subjective shortening: A model of pigeons' memory for event duration. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 14-30.
- STADDON, J. E. R., & REID, A. K. (1990). On the dynamics of generalization. *Psychological Review*, *97*, 576-578.
- THOMAS, D. R. (1974). The role of adaptation level in stimulus generalization. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 91-145). New York: Academic Press.
- THOMAS, D. R., MOOD, K., MORRISON, S., & WIERTELAK, E. (1991). Peak shift revisited: A test of alternative interpretations. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 130-140.
- THOMAS, D. R., & SETZER, J. (1972). Stimulus generalization gradients for auditory intensity in rats and guinea pigs. *Psychonomic Science*, *28*, 22-24.
- URCUIOLI, P. J., & ZENTALL, T. R. (1986). Retrospective coding in pigeons' delayed matching-to-sample. *Journal of Experimental Psychology: Animal Behavior Processes*, *12*, 69-77.
- URCUIOLI, P. J., ZENTALL, T. R., JACKSON-SMITH, P., & STEIRN, J. N. (1989). Evidence for common coding in many-to-one matching: Retention, intertrial interference, and transfer. *Journal of Experimental Psychology: Animal Behavior Processes*, *15*, 264-273.
- WASSERMAN, E. A., DELONG, R. E., & LAREW, M. B. (1984). Temporal order and duration: Their discrimination and retention by pigeons. In J. Gibbon & L. Allen (Eds.), *Timing and time perception* (Annals of the New York Academy of Sciences, Vol. 423, pp. 103-115). New York: New York Academy of Sciences.
- WEARDEN, J. H. (1992). Temporal generalization in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*, 134-144.
- WEARDEN, J. H., DENOVAN, L., FAKHRI, M., & HAWORTH, R. (1997). Scalar timing in temporal generalization in humans with longer stimulus durations. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*, 502-511.
- ZENTALL, T. R., URCUIOLI, P. J., JAGIELO, J. A., & JACKSON-SMITH, P. (1989). Interaction of sample dimension and sample-comparison mapping on pigeons' performance of delayed conditional discriminations. *Animal Learning & Behavior*, *17*, 172-178.

(Manuscript received December 10, 1996;
revision accepted for publication July 8, 1997.)