

Systematic errors in pigeons' memory for event duration: Interaction between training and test delay

MARCIA L. SPETCH

Dalhousie University, Halifax, Nova Scotia, Canada

Pigeons trained to choose different stimuli following short- and long-duration signals make disproportionately more "short" choices (i.e., "choose-short errors") following an increase in the retention interval and more "choose-long errors" following a decrease in this delay. The present experiment provided a systematic investigation of how these selective errors depend on the relationship between the training delay and the test delay. Pigeons were first trained with a 0-sec delay between the signal (2- or 8-sec food presentations) and the choice stimuli (red- and blue-lit keys). On subsequent test trials with 5- and 10-sec delays, choose-short errors predominated. Next, the birds were trained with a constant 10-sec delay and then tested with shorter or longer delays on some trials. The birds now responded accurately and without selective errors at the 10-sec training delay, but made choose-long errors at shorter delays and choose-short errors at longer delays. Finally, the birds were trained with a constant 20-sec delay and then tested with shorter and longer delays. Choose-long errors again appeared at shorter test delays, choose-short errors at longer test delays, and no differential errors at the 20-sec training delay. The selectivity of these errors generally increased with the absolute difference between the training and test delay. Theoretical implications of these results are discussed.

It has long been known that animals can make discriminative responses on the basis of the duration of a past event (e.g., Cowles & Finan, 1941; Reynolds & Catania, 1962; Woodrow, 1928). For example, in a version of the delayed symbolic matching-to-sample (DSMTS) task, pigeons can learn to peck a red key following a short (e.g., 2-sec) sample event and a green key following a longer (e.g., 10-sec) duration of the same event (Spetch & Wilkie, 1981). When the delay between the sample event and the choice period is very brief (e.g., less than 5 sec), pigeons can perform very accurately under this procedure. However, if the delay interval is suddenly lengthened on test trials, pigeons show not only a drop in overall choice accuracy, but also a systematic tendency to make many incorrect "short" choices (i.e., choosing the test stimulus that is correct for short samples). This "choose-short effect" at long delays has been found both with food access as the sample event (Spetch & Wilkie, 1982, 1983) and with light as the sample event (Spetch & Treit, 1984; Spetch & Wilkie, 1982, 1983). Rats also have been reported to make more short choices when tested with a lengthy delay following two-choice duration discrimination training (Church, 1980).

This research was supported by Natural Sciences and Engineering Research Council of Canada Operating Grant U0301. The data were presented in part at the 1985 Canadian Psychological Association conference in Halifax and at the 1985 Psychonomic Society meeting in Boston. I am grateful to D. Davis, D. E. Armstrong, and S. Stewart for their assistance with the research, and to B. Moore, B. Rusak, and D. Treit for their comments and suggestions. Reprint requests should be sent to: M. Spetch, Department of Psychology, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada.

The occurrence of these choose-short errors appears to depend critically on the way in which delays are tested (Spetch & Wilkie, 1983). When a variable delay procedure is used in which the long delays occur on only some trials of each session, the choose-short errors tend to persist and remain relatively stable across sessions. However, with a constant delay procedure, in which a given delay is in effect on all trials for a number of sessions, the choose-short errors rapidly disappear. Moreover, an opposite tendency, toward making choose-long errors, may emerge after a shift from a constant long delay (i.e., 20 sec) to a shorter one (i.e., 5 sec): In Spetch and Wilkie (1983), pigeons that had received lengthy training with the constant 20-sec delay made more choose-long errors on the first session after a decrease in the delay.

These systematic errors are potentially very interesting for the study of animal memory. They offer an additional dependent measure that could provide valuable clues about underlying memory processes (see Kraemer, Mazmanian, & Roberts, 1985; Spetch & Wilkie, 1983). I therefore undertook to explore these bidirectional choice errors more fully to firmly establish that they depend on the relationship between the training and test delay. The occurrence of choose-short errors following long test delays has now been demonstrated several times (Church, 1980; Spetch & Treit, 1984; Spetch & Wilkie, 1982, 1983). However, there has been only one demonstration of a systematic choose-long effect (Spetch & Wilkie, 1983), and this demonstration relied on results obtained during the single sessions that followed a decrease in the delay. Thus, whether the choose-long errors would also

persist over sessions if a variable delay procedure were used was not known. In addition, the two types of effects had never been shown to co-occur within a session: In Spetch and Wilkie (1983), the appearance of choose-long errors was always separated by several sessions from the appearance of choose-short errors.

The present experiment extended previous work by demonstrating that systematic errors of both types (i.e., choose-short and choose-long) would occur within the same test sessions. The magnitude and direction of these errors were shown to vary systematically as a function of the relationship between the test delay and the most recent training delay.

METHOD

Subjects

The subjects were 6 adult White Carneaux pigeons. Birds S3B0, S3B1, S3B2, and S3B3 were experimentally naive, whereas Birds S5B0 and S5B1 had served briefly in an autoshaping experiment. Mixed grain obtained primarily during experimental sessions maintained the pigeons at 85% to 90% of their free-feeding weights. The birds were housed in individual wire-mesh cages; water and grit were freely available.

Apparatus

The experimental chambers consisted of rectangular Grason-Stadler animal chests, which measured 25 cm high, 32 cm wide, and 35 cm deep on the inside. Each chamber contained three horizontally aligned response keys requiring a force of approximately .25 N to operate. The keys could be transilluminated with white, red, or blue light by stimulus projectors mounted behind each key. A solenoid-operated grain feeder was located below the center key, and grain presentations were accompanied by illumination of a lamp located in the feeder. The houselight, which was located behind the response panel, was used only during preliminary magazine-training sessions. Experimental contingencies and data recording were controlled by a PDP-8e computer located in an adjacent room.

Procedure

Preliminary training. All birds received one or two sessions of magazine training, followed by a few sessions of autoshaping in which illumination of the center key with white light was paired with grain presentation. Once pecking at the center key had been established, the birds were given one or two additional training sessions in which the side keys were individually illuminated with red or blue light in a randomly determined order; pecks to the illuminated key were reinforced according to a continuous reinforcement schedule, until the bird had made 20 reinforced pecks per stimulus.

Basic DSMTS procedure. The following features of the DSMTS procedure used were common to all phases of the experiment proper. Trials began with the presentation of white light on the center key as a trial-initiating stimulus. A single peck to this stimulus terminated it and resulted in presentation of the illuminated grain feeder as the sample stimulus. On a random half of the trials, the sample was "long" (8 sec). On the remaining trials, the sample was "short" (2 sec). Termination of the sample was followed (in some phases) by a delay interval during which the chamber was dark, and then red and blue lights were presented on the side keys as comparison stimuli. One color was designated correct for short samples, the other correct for long samples. This color designation was counterbalanced across birds, and the right-left position of red and blue was counterbalanced across trials within each session. A peck to

either comparison stimulus resulted in termination of both stimuli. If the correct stimulus was pecked, a 4-sec presentation of the illuminated feeder occurred as reinforcement. If the incorrect stimulus was pecked, the trial ended without reinforcement. (Pecks to the stimulus that was designated as correct for short samples will be referred to as "short choices," and pecks to the one correct for long samples as "long choices.") All trials were separated by a 45-sec intertrial interval (ITI), during which the chamber was dark. Sessions were conducted 5 or 6 days a week and continued until 48 trials had been completed, or for a maximum of 52 min.

Phase 1a: 0-sec delay training. Each bird was trained with no programmed delay (i.e., 0 sec) between sample offset and presentation of the comparison stimuli until its overall accuracy was 85% or better for 5 consecutive sessions. During this training phase, a correction procedure was used: the stimulus events were duplicated on the next trial whenever an incorrect response was made. However, accuracy levels were determined by performance on non-correction trials only. This phase required 14, 19, 14, 18, 11, and 13 sessions for Birds 1 through 6, respectively.

Phase 1b: Delay testing with 0-sec baseline. This phase consisted of 20 test sessions in which the baseline 0-sec delay occurred on a randomly determined half of the trials of each session, and a 5- or a 10-sec delay occurred equally often on the remaining trials. No correction trials occurred during this phase.

Phase 2a: 5- and 10-sec delay training. This phase consisted of 20 sessions in which a constant 5-sec delay occurred on all trials, followed by 20 sessions in which the delay was 10 sec on all trials. A correction procedure was reinstated during this phase, but, again, only the choices made on noncorrection trials were used to determine accuracy levels.

Phase 2b: Delay testing with 10-sec baseline. This phase immediately followed the 10-sec delay training and consisted of 20 test sessions in which the 10-sec baseline delay occurred on a randomly determined 75% of the trials. Within each session, two test delays occurred equally often on the remaining 25% of the trials, one shorter and one longer than the 10-sec baseline delay. On odd-numbered days, these test delays were 0 and 20 sec; on even-numbered days, the test delays were 5 and 15 sec. The correction procedure was not in effect during this phase.

Phase 3a: 20-sec delay training. During this phase, the birds were given 20 sessions in which a constant 20-sec delay occurred on all trials, and the correction procedure was in effect.

Phase 3b: Delay testing with 20-sec baseline. This phase, consisting of 20 sessions, was identical to Phase 2b except that the baseline delay was 20 sec and the test delays were 0 and 30 sec on odd-numbered days and 5 and 25 sec on even-numbered days.

RESULTS

0-sec Delay Training

The birds did not display any systematic tendency to make one type of choice more often than the other during 0-sec delay training. Averaged across all sessions in this phase, the percentage of correct short choices (mean = 81%, *SEM* = 1.6%) did not differ significantly [$t(5) = 1.16, p > .1$] from the percentage of correct long choices (mean = 79%, *SEM* = 1.9%).

Delay Testing with 0-sec Baseline

Figure 1 shows the birds' mean accuracy scores on short-sample and long-sample trials as a function of delay, averaged over the 20 test sessions of Phase 1b. Data for the individual subjects are shown in the top panel of

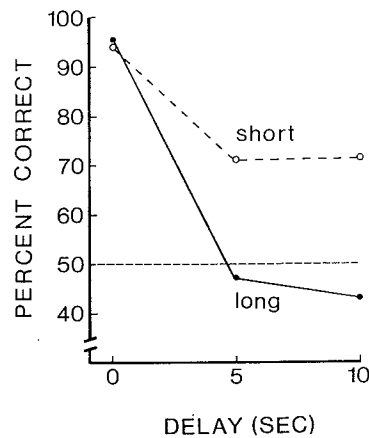


Figure 1. Mean accuracy on short-sample trials (○) and long-sample trials (●) as a function of test delay. Data points are averaged across the 20 test sessions of Phase 1b, in which the training delay was 0 sec.

choices were made. A two-way repeated measures analysis of variance revealed significant main effects of delay [$F(2,10) = 99.59, p < .001$] and sample duration [$F(1,5) = 12.76, p < .02$], as well as a significant delay \times sample duration interaction [$F(2,10) = 25.85, p < .001$].

Training with 5- and 10-sec Delays

The constant 5-sec delay training served to rapidly (i.e., within four sessions) eliminate the predominance of choose-short errors and resulted in a steady increase in overall accuracy. During the last block of two sessions, mean accuracy was 95% ($SEM = 1.6\%$) on short-sample trials and 93% ($SEM = 1.5\%$) on long-sample trials. A tendency to make choose-short errors reappeared with the shift to a 10-sec delay, but again rapidly disappeared with the training at this delay. By the last block of two 10-sec delay training sessions, mean accuracy was 93% ($SEM = 2.2\%$) with short samples and 91% ($SEM = 1.2\%$) with long samples.

Table 1. At the 0-sec training delay, accuracy was very high and was similar on short- and long-sample trials. However, choose-short errors appeared at the 5- and 10-sec delays: The lower accuracy on long-sample trials than on short-sample trials indicates that more incorrect short

Delay Testing with a 10-sec Baseline

Figure 2 shows the birds' mean accuracy on short and long trials as a function of test delay when the baseline delay was 10 sec; individual subject data are shown in the middle panel of Table 1. Accuracy was high and simi-

Table 1
Percentage of Correct Choices on Short and Long Trials as a Function of Test Delay for Each of the Six Pigeons During the 20 Test Sessions of Phases 1, 2, and 3

| Pigeon | Phase 1b (0-sec Baseline Delay) | | | | | |
|--------|---------------------------------|-------------|------------------|-------------|-------------------|-------------|
| | 0-sec Test Delay | | 5-sec Test Delay | | 10-sec Test Delay | |
| | Short Sample | Long Sample | Short Sample | Long Sample | Short Sample | Long Sample |
| S5B0 | 92 | 97 | 65 | 54 | 63 | 49 |
| S5B1 | 95 | 98 | 77 | 69 | 82 | 62 |
| SB80 | 95 | 93 | 73 | 29 | 81 | 34 |
| SB81 | 98 | 94 | 76 | 33 | 71 | 23 |
| SB82 | 97 | 96 | 69 | 40 | 61 | 41 |
| SB83 | 88 | 98 | 66 | 57 | 73 | 51 |

| Pigeon | Phase 2b (10-sec Baseline Delay) | | | | | | | | | |
|--------|----------------------------------|-------------|------------------|-------------|-------------------|-------------|-------------------|-------------|-------------------|-------------|
| | 0-sec Test Delay | | 5-sec Test Delay | | 10-sec Test Delay | | 15-sec Test Delay | | 20-sec Test Delay | |
| | Short Sample | Long Sample | Short Sample | Long Sample | Short Sample | Long Sample | Short Sample | Long Sample | Short Sample | Long Sample |
| S5B0 | 61 | 100 | 69 | 100 | 92 | 90 | 100 | 67 | 93 | 50 |
| S5B1 | 57 | 100 | 96 | 100 | 97 | 95 | 97 | 85 | 100 | 56 |
| SB80 | 85 | 93 | 76 | 100 | 91 | 93 | 86 | 85 | 85 | 72 |
| SB81 | 55 | 100 | 84 | 100 | 87 | 85 | 88 | 59 | 90 | 48 |
| SB82 | 74 | 96 | 71 | 97 | 91 | 82 | 96 | 55 | 100 | 38 |
| SB83 | 85 | 100 | 85 | 100 | 96 | 95 | 96 | 74 | 100 | 74 |

| Pigeon | Phase 3b (20-sec Baseline Delay) | | | | | | | | | |
|--------|----------------------------------|-------------|------------------|-------------|-------------------|-------------|-------------------|-------------|-------------------|-------------|
| | 0-sec Test Delay | | 5-sec Test Delay | | 20-sec Test Delay | | 25-sec Test Delay | | 30-sec Test Delay | |
| | Short Sample | Long Sample | Short Sample | Long Sample | Short Sample | Long Sample | Short Sample | Long Sample | Short Sample | Long Sample |
| S5B0 | 56 | 80 | 38 | 100 | 72 | 81 | 78 | 75 | 82 | 56 |
| S5B1 | 76 | 100 | 78 | 96 | 86 | 86 | 95 | 76 | 79 | 44 |
| SB80 | 56 | 100 | 44 | 100 | 64 | 81 | 59 | 88 | 61 | 77 |
| SB81 | 56 | 95 | 62 | 91 | 84 | 75 | 86 | 75 | 95 | 67 |
| SB82 | 54 | 100 | 58 | 96 | 83 | 78 | 91 | 61 | 86 | 52 |
| SB83 | 86 | 95 | 32 | 100 | 95 | 93 | 100 | 83 | 100 | 68 |

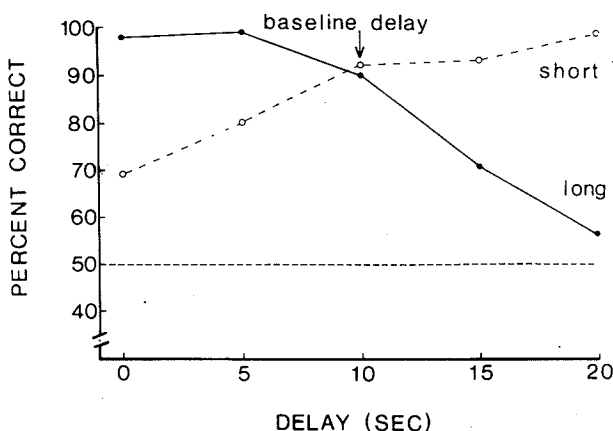


Figure 2. Mean accuracy on short- (o) and long-sample (●) trials as a function of test delay during Phase 2b. The baseline (training) delay is 10 sec. Data points are averaged across all test sessions of Phase 2b.

lar on both types of trials at the 10-sec delay. However, the fact that, with delays shorter than the baseline delay, accuracy was higher on long-sample trials than on short-sample trials indicates that choose-long errors predominated, whereas the opposite was true for delays longer than the baseline delay. A two-way repeated measures analysis of variance indicated a significant effect of delay [$F(4,20) = 13.24, p < .001$], but not of sample duration [$F(1,5) = 1.76, p > .1$]. The interaction between delay and sample duration was again highly significant [$F(4,20) = 26.23, p < .001$].

20-sec Delay Training

The higher mean accuracy on short-sample trials disappeared within four sessions of training with the constant 20-sec delay. By the last block of two sessions, accuracy levels were reasonably high with both short samples (mean = 79%, SEM = 1.9%) and long samples (mean = 83%, SEM = 2.9%).

Delay Testing with a 20-sec Baseline

The birds' mean accuracy on short- and long-sample trials during Phase 3b is shown as a function of delay in Figure 3, and the data for individual subjects are shown in the bottom panel of Table 1. Once again, accuracy was similar on both types of trials at the baseline delay (20 sec), but was higher on long-sample trials at delays shorter than baseline and higher on short-sample trials at delays longer than baseline. A two-way repeated measures analysis of variance on these data revealed no significant main effect of delay [$F(4,20) = 2.33, p > .5$] or of sample duration [$F(1,5) = 3.33, p > .1$], but again a highly significant interaction between delay and sample duration [$F(4,20) = 26.17, p < .001$].

Systematic Choice Errors

Figure 4 shows the mean difference between choose-long errors and choose-short errors as a function of test delay for each of the three baseline delays. As Figure 4

indicates, the difference between the two types of errors was very close to 0 for each point at which the test delay matched the baseline delay. However, when the test delay was shorter than the baseline delay, more choose-long errors were made, and when the test delay was longer than the baseline delay, more choose-short errors were made. Moreover, with one exception, the deviation from 0 was more extreme with test delays that were farther from the baseline delay. The correlation coefficient between the absolute magnitude of the error difference and the difference between the baseline and test delay (on the basis of the data from all phases and for each bird) was positive and highly significant [$r(70) = .627, p < .001$].

DISCUSSION

The present results extend the previous demonstration of choose-short and choose-long errors (Spetch & Wilkie, 1983) in several ways. First, the occurrence of systematic choose-long errors was shown to be a replicable phenomenon. Like choose-short errors, these errors were found to persist over test sessions when a variable delay procedure was used. Second, the present study demonstrated the occurrence of both types of systematic errors within the same test sessions. Third, and most important, the direction of these errors was shown to depend on the relationship between the training delay and the test delay: A tendency to make choose-short errors occurred when the test delay was longer than the training delay, whereas more choose-long errors occurred when the test delay was shorter than the training delay. Furthermore, the selectivity of these errors tended to become more extreme as the absolute discrepancy between the training and test delay increased.

The systematic nature of these selective errors and their dependence on the discrepancy between training and test delays indicate that they should not be considered per-

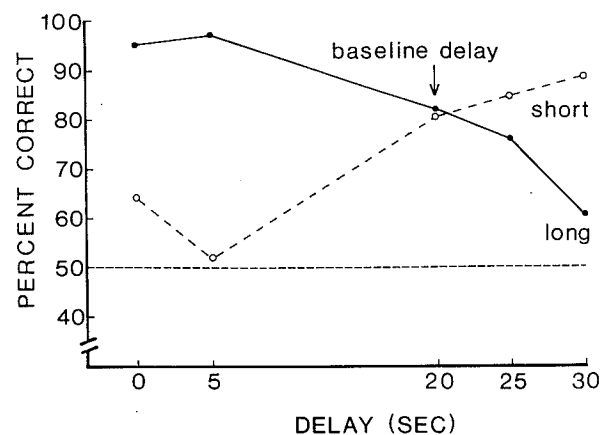


Figure 3. Mean accuracy on short- (o) and long-sample (●) trials as a function of test delay during Phase 3b, with the 20-sec baseline (training) delay. Data points are averaged across all Phase 3b test sessions.

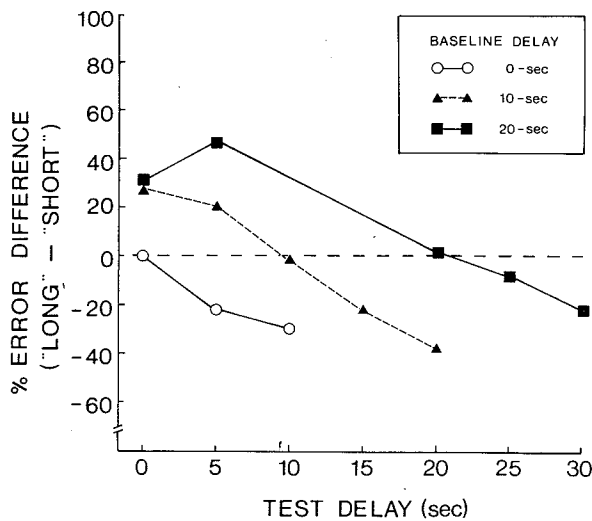


Figure 4. Mean difference between choose-long and choose-short errors as a function of test delay for the three baseline delay conditions.

formance biased. The present results, for example, cannot be explained in terms of matching to obtained reinforcement rates (cf. McCarthy, 1983). The errors also cannot be attributed to a simple response preference that emerges in any disrupting, novel, or uncertain situation. Instead, they appear to be perceptual or decision errors and seem to arise from processes underlying the pigeon's memory for sample duration. In this respect, they may be similar to the time-order errors often observed in studies of human time perception (see Hellstrom, 1985).

The nature of the processes that give rise to these systematic errors in pigeon memory for duration remains in question. Two proposed models seem able to account for the present data. The subjective shortening model (Spetch & Wilkie, 1983) directly predicts these results. According to this model, choose-short errors arise when subjects retrospectively remember a sample duration that has become foreshortened over the delay interval. Training at a given long delay is assumed to result in the learning of new associations between the foreshortened duration values and the correct responses. Choose-long errors are thought to arise on short-delay test trials because the samples on these trials would not be foreshortened and therefore should seem longer than those that had become associated with the responses during the long-delay training.

More recently, Kraemer et al. (1985) offered an alternative, coding interpretation of choose-short and choose-

long errors. Their model suggests that the sample durations are coded into prospective representations. They suggest that the choose-short effect represents a bias that occurs whenever the code is forgotten. With the assumption that codes for long samples can be remembered better than codes for short samples, they account for choose-long errors by assuming that, during training with a long delay, the pigeons learn to respond "long" whenever they remember any sample code and to respond "short" if the code is forgotten. Then, because all codes would be remembered well on short-delay test trials, a tendency to respond "long" after all samples would emerge.

These data strongly suggest that choose-short and choose-long effects arise from memory processes, but do not differentiate between the subjective shortening model and the coding model. Nevertheless, the two models make somewhat divergent predictions with respect to other manipulations, and research is currently under way to test these alternative predictions.

REFERENCES

- CHURCH, R. M. (1980). Short-term memory for time intervals. *Learning & Motivation*, 11, 208-219.
- COWLES, J. T., & FINAN, J. L. (1941). An improved method for establishing temporal discrimination in white rats. *Journal of Psychology*, 11, 335-342.
- HELLSTROM, A. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin*, 97, 35-61.
- KRAEMER, P. J., MAZMANIAN, D. S., & ROBERTS, W. A. (1985). The choose-short effect in pigeon memory for stimulus duration: Subjective shortening versus coding models. *Animal Learning & Behavior*, 13, 349-354.
- MCCARTHY, D. (1983). Measures of response bias at minimum detectable luminance levels in the pigeon. *Journal of the Experimental Analysis of Behavior*, 39, 87-106.
- REYNOLDS, G. S., & CATANIA, A. C. (1962). Temporal discrimination in pigeons. *Science*, 135, 314-315.
- SPETCH, M. L., & TREIT, D. (1984). The effect of d-amphetamine on short-term memory for time in pigeons. *Pharmacology, Biochemistry & Behavior*, 21, 663-666.
- SPETCH, M. L., & WILKIE, D. M. (1981). Duration discrimination is better with food access as the signal than with light as the signal. *Learning & Motivation*, 12, 40-64.
- 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.
- WOODROW, H. (1928). Temporal discrimination in the monkey. *Journal of Comparative and Physiological Psychology*, 8, 395-427.

(Manuscript received February 24, 1986;
revision accepted for publication November 25, 1986.)