

Discriminal distance analysis supports the hypothesis that pigeons retrospectively encode event duration

DONALD M. WILKIE and ROBERT J. WILLSON

University of British Columbia, Vancouver, British Columbia, Canada

A discriminial distance analysis procedure similar to that used by Roitblat (1980) was employed to test the hypotheses that animals either retrospectively (Spetch & Wilkie, 1983) or prospectively (Kraemer, Mazmanian, & Roberts, 1985) encode durations of events. Pigeons were required to discriminate 2-, 8-, and 10-sec presentations of light. Choices of red, orange, and green keys were correct after 2, 8, and 10 sec, respectively. The key elements in this design were (1) that some samples (8 and 10 sec) and some choice stimuli (red and orange) were more difficult to discriminate than were others, and (2) that an easy sample discrimination (2 vs. 8 sec) was mapped onto a difficult choice discrimination (red vs. orange), and vice versa. An examination of raw error scores and calculated confusion indexes in three experiments supported the hypothesis that subjects retrospectively, rather than prospectively, encode event duration.

When subjects such as pigeons are tested for delayed recall of the duration of a preceding event such as the illumination of a houselight, a systematic error occurs—namely, subjects tend to report that the event was short, regardless of whether the light was on for a short or long period. This *choose-short* effect has been observed on numerous occasions (Church, 1980; Kraemer, Mazmanian, & Roberts, 1985; S. Roberts, 1982; Spetch, 1987; Spetch & Rusak, 1989; Spetch & Sinha, 1989; Spetch & Treit, 1984; Spetch & Wilkie, 1982). [A similar effect occurs when animals remember the rate of alternation of two stimuli (Honig & Spetch, 1988) or the number of events (Fetterman & MacEwen, 1989).] The choose-short effect occurs only when the test for recall is delayed; at short retention intervals, subjects remember both short and long events about equally well.

Two models have been proposed to account for the choose-short effect. Both models assume an interaction of working and reference memory (Honig, 1978). In the *subjective shortening* model (Spetch & Wilkie, 1982, 1983), rules such as “choose the red alternative after a short light” and “choose the green alternative after a long light” are stored in reference memory. Event duration is stored in working memory. In this model, the choose-short effect is accounted for by assuming that the remembered duration of the event in working memory systematically shrinks or shortens with the passage of time. Thus,

after a long sample and a long retention interval, the contents of working memory would be more similar to a short sample; shortening of the remembered duration of the long sample produces an increased tendency for the subject to choose short. Because this model assumes that subjects attempt to recall the actual duration of the target event, it may be classified as a retrospective encoding model.

The second model, the *prospective coding* model (Kraemer, Mazmanian, & Roberts, 1985), assumes that, during the event being timed, subjects form a code or instruction about what to do on the subsequent test phase of a trial (e.g., when the event is short, a code of “peck red” might be formed, whereas if the stimulus is long, a code of “peck green” would be formed). Reference memory contains rules relating these codes to response choices. It is also assumed that these codes are retained in an all-or-none manner, but that long codes are better retained than are short codes, perhaps because of increased opportunities for rehearsal on long-sample trials. Finally, it is assumed that in the absence of a code (i.e., when a code is forgotten), subjects choose short. In this model, the choose-short effect is accounted for by assuming that codes are more likely to be forgotten with longer retention intervals.

A procedure adapted by Roitblat (1980) from one used by Conrad (1964) has the potential for distinguishing between retrospective and prospective encoding of information. Consequently, we have adapted this procedure to the event duration discrimination procedure with the intention of producing evidence to support one or the other of the two models of event-duration memory.

EXPERIMENT 1

The procedure used by Roitblat was a variant of delayed matching-to-sample. On each trial, presentation of one

This research was supported by the Natural Sciences and Engineering Research Council of Canada. Suzanne MacDonald provided helpful comments. Shayne Kardal and Kevin Lee assisted with the experiments. Christina Von Borman, Rachel McGhee, and Jacalyn Hill assisted with data analysis. Correspondence regarding this article may be sent to Donald M. Wilkie, Department of Psychology, University of British Columbia, #2509-2136 West Mall, Vancouver, British Columbia V6T 1Y7, Canada, or to the following e-mail address: Donald.M._Wilkie@mts.ubc.ca

of three samples (blue, orange, or red pecking keys) was followed by three choice stimuli (0° , 12.5° , and 90° line stimuli presented on three pecking keys). Choices of the 0° , 12.5° , and 90° lines were correct and consequently rewarded after presentation of the blue, orange, and red samples, respectively. The key element in the design is the fact that some samples and some choice stimuli are harder to discriminate than are others. Red and orange samples are more difficult to discriminate than are red and blue or orange and blue. The 0° and 12.5° lines are more difficult to discriminate than are 0° and 90° or 12.5° and 90° lines. These differences in discriminability allow inferences to be made about whether subjects use retrospective or prospective encoding. If subjects retrospectively remember the sample stimulus, then more errors are likely to be made on trials with similar samples (e.g., red and orange) than on trials with dissimilar samples (e.g., red and blue). On the other hand, if subjects prospectively encode the sample by, for example, coding blue as "peck 0° ," orange as "peck 12.5° ," and red as "peck 90° ," subjects should have most difficulty in discriminating on blue versus orange trials, trials that would be easy if subjects retrospectively coded samples. Consequently, an examination of the patterns of errors permits inferences about the type of encoding employed by the subjects.

In the research reported here, we adapted this procedure to the event-duration discrimination paradigm. Rather than using blue, orange, and red, we used durations of house-light presentations (2, 8, and 10 sec) as samples. Red, orange, and green pecking keys were used as choice stimuli. Choices of red, orange, and green keys were correct and consequently rewarded following presentations of 2, 8, and 10 sec of house-light, respectively. As in Roitblat's procedure, some samples (8 and 10 sec of house-light) and some choices (red vs. orange) are more difficult than are others. By examining errors on the various types of trials, we were able to make inferences about the nature of the subject's encoding of event duration.

Method

Subjects. Six King pigeons, maintained at about 90% of normal body weight, served as subjects. These subjects had served in previous timing experiments (e.g., Wilkie, 1988). The pigeons were housed individually in large plastic-coated mesh cages, with free access to water, health grit, and oyster shells. The colony was maintained on a light:dark cycle matched to natural sunrise and sunset times. The subjects were tested 5 days a week, a few hours after colony light onset.

Apparatus. Each pigeon was tested in one of six test chambers containing a white house-light (intensity of roughly 20 cd/m^2), a grain feeder, and three pecking keys, which could be lit with red (Kodak Wratten gelatin filter No. 29), orange (No. 21), or green (No. 53) light, roughly 0.75 cd/m^2 . A Data General NOVA 3 computer and MANX state notation software (Gilbert & Rice, 1979) controlled the experimental equipment and collected the responses.

Procedure. Each subject received 50 sessions of training in the first phase of Experiment 1. Data were collected during the final 20 of these sessions. In these sessions, 18 different trial types occurred three times, with the order of trial types randomized. Trials were separated by a 15-sec interval of darkness. All trials started

with the illumination of the house-light. On one third of the trials, the house-light remained on for 2 sec. On other trials, the house-light remained on for 8 or 10 sec. House-light presentation was followed immediately by presentation of three choice stimuli: red, orange, and green pecking keys. The location (left, center, or right) of the three colors was counterbalanced over trials, resulting in six trial types. Red was designated as correct after 2 sec of house-light. Orange was correct after 8 sec; green was correct after 10 sec. Correct choices were rewarded with 4-sec presentations of mixed grain.

The first phase was followed by three other phases, each lasting for 12 sessions. In the second phase, sessions were identical to those of the first phase, except that the 8-sec house-light presentations were shortened to 6 sec. In the third phase, the 6-sec house-light presentations were lengthened to 8 sec. In the fourth and final phase, the middle sample duration was again 6 sec. The rationale for manipulating the duration of the middle sample was to make discrimination of duration on the different trial types more or less difficult. When task difficulty is varied, certain indices of confusion may be calculated (see Roitblat, 1980, who varied task difficulty by manipulating the duration of sample presentation and the duration of the retention interval; see also Experiment 2 below). These confusion scores may be used to make inferences about the encoding strategies used by the subjects.

Data Collection and Analyses. In each session, a record was kept of the number of times each subject chose red, orange, and green keys on trials when the house-light was presented for 2, 8, and 10 sec. The first type of analysis consisted of an examination of the frequency of the different types of errors: orange and green choices on 2-sec trials, red and green choices on 8-sec trials, and red and orange choices on 10-sec trials. The second analysis consisted of an examination of discriminational distance scores. These scores were calculated in exactly the same manner as in Roitblat's (1980) study. For each subject, a 3×3 matrix of choices was constructed (i.e., the number of red, orange, and green choices after 2-, 8-, and 10-sec samples). If the number of correct red choices after 2 sec is designated as R , the number of incorrect orange choices after 2 sec is designated as S , and the number of other choices is designated as shown in Table 1, then it is possible to calculate from the frequency of incorrect choices the following discriminational distance scores:

$$AB = 1 - (S + U)/N$$

$$BC = 1 - (W + Y)/N$$

$$AC = 1 - (T + X)/N$$

where S , U , W , Y , T , and X are the choice frequencies appearing in Table 1 and N is the total number of choices.

These discriminational distance scores become larger as errors are less frequently made and become smaller as more errors are made. The AB score indexes orange errors on 2-sec trials and red errors on 8-sec trials. The BC score indexes green errors on 8-sec trials and orange errors on 10-sec trials. The AC score indexes green errors on 2-sec trials and red errors on 10-sec trials. Because highly trained subjects should easily discriminate 2- and 10-sec samples and red and green choices, the AC discriminational distance score should be much larger than either the AB or the BC scores. The values of the latter scores are of particular interest. The AB score reflects red-orange

Table 1
Design for Discriminational Distance Analysis Experiment

Sample	Choice		
	Red	Orange	Green
2 sec	R	S	T
8 sec	U	V	W
10 sec	X	Y	Z

choice errors on the retrospectively relatively easy, but prospectively more difficult, 2- versus 8-sec trials. That is, a subject employing a prospective encoding strategy may mistakenly encode "peck red" as "peck orange," or vice versa. The BC score reflects green-orange errors on the prospectively relatively easy (two easily discriminated colors), but retrospectively difficult, 8- versus 10-sec trials. If subjects coded sample duration retrospectively, they should make frequent green-orange errors. That is, BC should have a relatively small value. On the other hand, if subjects encoded sample duration prospectively (by coding 2 sec as red, 8 sec as orange, and 10 sec as green), they should have relatively small AB scores. Thus, a basic question concerning discriminial distance scores was whether AB scores were larger or smaller than were BC scores.

Roitblat (1980) calculated the ratio of AB scores to BC scores and examined this ratio as task difficulty was varied through manipulation of sample and retention interval duration. We did similar calculations as our task was made more or less difficult by varying the duration of the middle sample from 8 to 6 sec. The magnitude of this ratio will change in different ways as task difficulty varies, depending upon whether subjects employ a retrospective or prospective encoding strategy. If subjects code prospectively, this ratio will increase as the task is made easier. On the other hand, the ratio will become smaller as the task becomes easier, if subjects employ a retrospective coding strategy. Consequently, the final analysis we performed was to examine the AB/BC ratio as the middle sample was varied between 8 and 6 sec.

Results

Error patterns. Figure 1 shows the number of times the subjects chose the red, orange, and green keys when samples were 2, 8, and 10 sec during the final 20 sessions of the first phase of Experiment 1. When samples were 2 sec, the subjects chose the correct red key most frequently, orange was sometimes chosen, and green was rarely chosen. When samples were 10 sec, a mirror-image pattern was observed: The correct green key was chosen most frequently, orange was sometimes chosen, and red was rarely chosen. When the sample was 8 sec, a different pattern of choices was observed. Red was rarely chosen, but orange and green were chosen nearly equally often. These patterns of choices were clearly expected of subjects who had retrospectively encoded houselight du-

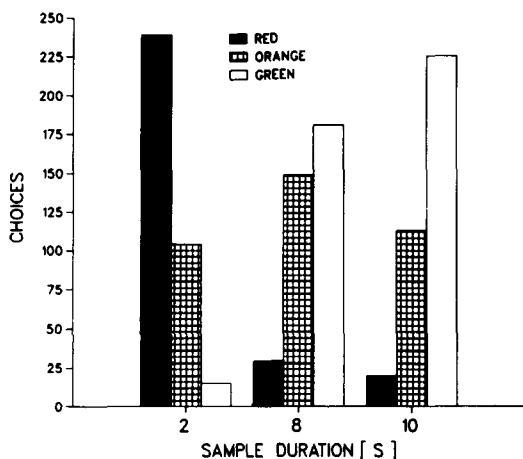


Figure 1. Choices of red, orange, and green keys after presentation of 2-, 8-, and 10-sec samples.

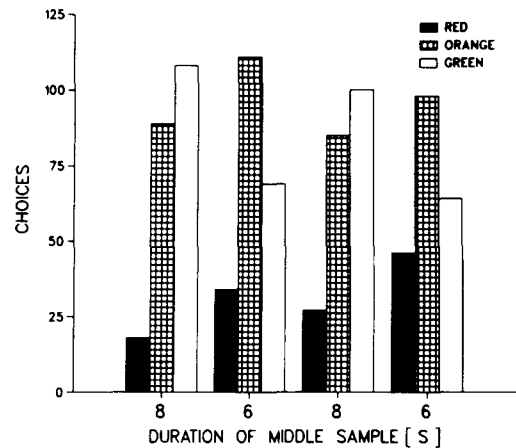


Figure 2. Choices of red, orange, and green keys on middle duration trials as the duration of the sample on these trials was varied from 8 to 6 sec.

ration. Our subjects were more likely to confuse dissimilar choices (orange and green) that were mapped onto similar durations (8 and 10 sec) than similar choices (red and orange) that were mapped onto dissimilar durations (2 and 8 sec).

Choice scores in Figure 1 were subjected to a 3 (sample) \times 3 (choice) repeated measures analysis of variance. A significant interaction [$F(4,20) = 37.67, p < .001$] confirmed that pattern of choices differed for the different sample durations. Subsequent Newman-Keuls tests revealed the following significant differences: On 2-sec sample trials, red choices were more frequent than were both orange and green choices. Orange choices were more frequent than were green choices. On 8-sec trials, red choices were less frequent than were choices of both orange and green. Orange and green choices did not differ. On 10-sec trials, green choices were higher than were both orange and red choices. Red choices were lower than were orange choices.

Figure 2 shows the number of choices of red, orange, and green on the middle sample duration trials as the duration of this sample was varied during the various phases of the experiment. (Data from the first 8-sec condition are the same as shown in Figure 1, except that a correction has been applied to take into account the fact that there were different numbers of sessions in the first and subsequent phases.) The number of errors on middle duration sample trials decreased as the duration of this sample was changed from 8 to 6 sec. This change in the pattern of errors is consistent with the hypothesis that subjects code houselight duration retrospectively. Choices between dissimilar keys (orange and green) changed as samples were made more or less similar.

The choice data in Figure 2 were subjected to a key color \times phase repeated measures analysis of variance. Key choices were significantly different in the different phases [$F(6,30) = 7.42, p = .0001$]. Subsequent Newman-Keuls tests revealed no change in red choices over the phases, but did reveal differences in orange and

Table 2
Discriminal Distance Data from Experiment 1

	Pigeon					
	1	2	3	4	5	6
	Middle Sample = 8 sec					
AB	.45	.44	.93	.55	.72	.64
BC	.18	.13	.18	.14	.28	.17
AC	.86	.90	.97	.87	.94	.85
AB/BC	2.45	3.38	5.17	3.86	2.53	3.77
	Middle Sample = 6 sec					
AB	.41	.63	.89	.49	.69	.60
BC	.27	.31	.59	.27	.44	.28
AC	.83	.90	.97	.75	.96	.86
AB/BC	1.53	1.99	1.51	1.77	1.55	2.15
	Middle Sample = 8 sec					
AB	.49	.68	.94	.45	.86	.58
BC	.21	.22	.25	.31	.15	.30
AC	.71	.92	.99	.74	.94	.70
AB/BC	2.25	3.02	3.63	1.75	5.52	1.91
	Middle Sample = 6 sec					
AB	.35	.68	.90	.56	.62	.52
BC	.37	.44	.53	.34	.39	.58
AC	.74	.92	.49	.81	.89	.70
AB/BC	0.96	1.53	1.70	1.61	1.57	0.89

green choices. Specifically, orange was chosen significantly more frequently than was green when the sample was 6 sec, but not when the sample was 8 sec.

Discriminal distance analysis. Table 2 shows discriminational distance scores for each subject for each phase of Experiment 1. The AC scores, which reflect red-green confusions, were, as expected, higher than both the AB scores, which reflect orange-red confusions, and the BC scores, which reflect green-orange confusions. As predicted by a retrospective coding hypothesis, AB scores were larger than BC scores. This difference was confirmed by an AB versus BC \times phase repeated measures analysis of variance. The AB scores ($M = .632$, $s = .173$) were significantly larger than BC scores ($M = .351$, $s = .200$), $F(1,5) = 49.11$, $p = .0013$. The differences between these scores changed in the different phases, $F(3,15) = 15.07$, $p = .001$; the AB scores were much larger than the BC scores during the phases in which the middle sample was 8 sec rather than 6 sec.

Figure 3 shows the ratio of AB scores to BC scores as the duration of the middle sample was varied from 8 to 6 sec in the different phases of Experiment 1. As expected from the retrospective coding hypothesis, this ratio became smaller as the duration of the middle sample was changed from 8 to 6 sec. These changes were significant when tested in a repeated measures analysis of variance [$F(3,15) = 7.61$, $p = .0026$]. Newman-Keuls tests revealed that the ratio was higher during the 8-sec sample phases.

Discussion

There were two key features in the design of Experiment 1. The first was that the three samples and the three

choice stimuli were chosen such that some pairs of samples and some pairs of choice stimuli were more difficult to discriminate than were others. The second feature was the manner in which the samples were mapped onto the choice stimuli. More specifically, samples that were relatively easy to discriminate (i.e., 2 vs. 8 sec) were mapped onto choices that were more difficult to discriminate (i.e., red vs. orange keys), whereas samples that were more difficult to discriminate (i.e., 8 vs. 10 sec) were mapped onto relatively easy choices (i.e., orange vs. green). This design permits inferences to be made about the nature of the sample encoding strategies used by subjects. If subjects encode the actual duration of a sample event and attempt to recall it, they will have difficulty on similar sample (8 vs. 10 sec) trials, which will be manifested in frequent orange-green confusions. If, on the other hand, subjects do not encode the actual duration of the sample event but rather transform the event into some instruction about what to do during a subsequent test, then they will tend to confuse similar (red and orange) choice instructions. Our results unambiguously support the former assumption about the nature of encoding processes in event-duration discrimination paradigms.

Two types of data support the hypothesis that subjects attempt to remember the actual durations of the sample events rather than some transformation of these into choice instructions. The first is the raw number of various sorts of errors made by the subjects on different trial types. On 2-sec sample trials on which red was correct, the subjects sometimes chose orange, but they chose red significantly more often. On 8-sec sample trials on which orange was correct, the subjects chose red infrequently and, again, at a significantly different frequency from orange. On 10-sec sample trials on which green was correct, the subjects again chose red significantly less frequently than they chose orange. Thus, on all trial types, the subjects tended to chose red and orange at significantly different rates. The pattern of choices was different for green and orange. With 2-sec samples, orange and green were not

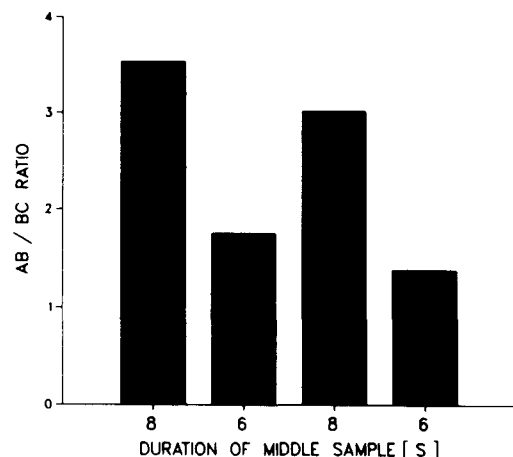


Figure 3. The ratio of AB to BC discriminational distance scores as a function of the duration of the middle sample.

confused, showing that the subjects could significantly discriminate these two colors. On 10-sec sample trials, orange was incorrectly chosen fairly often, but again at a significantly different frequency than was green. With the 8-sec sample, on the other hand, orange and green were not chosen at significantly different frequencies. Thus, these raw choice data demonstrate clearly that the subjects had more difficulty on prospectively easy choices (green-orange) than on prospectively hard choices (red-orange).

Calculated confusion indexes (i.e., discriminial distance scores) also support the hypothesis that subjects attempt to recall the actual durations of the sample events. Three main indexes were calculated: AB (which reflects red-orange errors), BC (which reflects orange-green errors), and the ratio of these (i.e., AB/BC). More frequent confusions decrease the values of AB and BC. The AB scores were larger than BC scores, again showing that the subjects had more difficulty in choosing between easy choice stimuli (green-orange) than more difficult ones (red-orange). The AB to BC ratios were typically larger than 1.00, reflecting the fact that AB was larger than BC.

The results of manipulating the duration of the middle duration sample event provided important corroborative evidence. Because this manipulation affected the ease with which sample duration could be recalled (i.e., subjects should not confuse 6 and 10 sec as readily) without affecting which prospective codes would have to be formed (i.e., subjects could continue to code "peck red," "peck orange," and "peck green"), the obtained effect of this manipulation on green-orange errors and the AB, BC, and AB/BC ratio also supports the hypothesis that subjects retrospectively encode event duration.

EXPERIMENT 2

The choose-short effect, for which the retrospective subjective shortening model was proposed, occurs when a delay (i.e., retention interval) is imposed between sam-

ple presentation and the choice period. The purpose of Experiment 2 was to examine the tendency to choose the short comparison key in the discriminial distance experimental design employed in Experiment 1. In particular, we examined error patterns when a delay was imposed between 2-, 8-, and 10-sec samples and the presentation of the red, orange, and green choice keys. Delays of 0, 5, and 10 sec were used.

Method

Subjects and Apparatus. The same subjects and apparatus that were used in Experiment 1 were employed in Experiment 2.

Procedure. Each subject received 26 sessions. Each session consisted of 54 trials. On one third of the trials, the red, orange, and green choice keys were lit immediately (0-sec delay) after the offset of the sample houselight. On another one third of the trials, a 5-sec delay separated the offset of the sample and the onset of the choice keys. On the remaining trials, the delay was 10 sec. The test chambers were dark during the delays. The delays occurred in a random order in each session, as did the three sample durations (i.e., 2, 8, and 10 sec) of houselight. In each session, a record was kept of the number of times that the subjects chose the red, orange, and green alternatives after 2-, 8-, and 10-sec samples, at each of the three delays (0, 5, and 10 sec). Data analyses were based on these scores, summed over the 26 sessions.

Results

Error patterns. Figure 4 shows the number of times subjects chose red, orange, and green keys when samples were 2, 8, and 10 sec, at each of the three delays (0, 5, and 10 sec). When there was no delay, the pattern of choice was similar to that observed in Experiment 1. On 2-sec sample trials, the subjects chose red frequently, orange occasionally, and green rarely. On 8-sec sample trials, orange was chosen most often, followed in frequency by green and red. After 10-sec samples, green was chosen most frequently, followed in frequency by orange and red. As in Experiment 1, orange-green errors were more frequent than were red-orange errors. With delays of 5 and 10 sec, there were several changes in the patterns of key choices. Choices on 2-sec sample

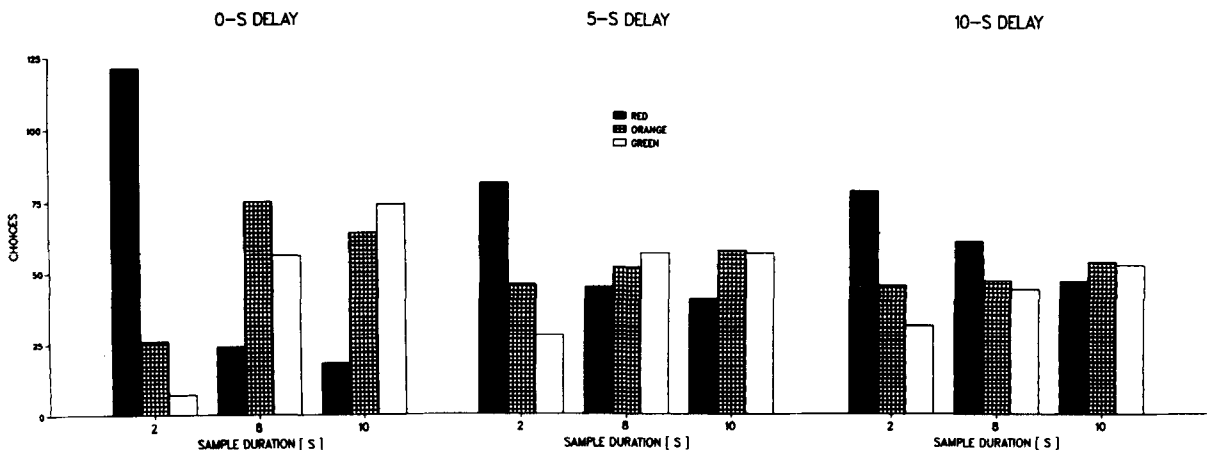


Figure 4. Choices of red, orange, and green keys on 2-, 8-, and 10-sec sample trials when the delay was 0, 5, and 10 sec.

trials were the least affected by the imposition of the delays: errors (orange and green choices) increased slightly on these trials. Choices on 8- and 10-sec sample trials were more affected. On these trials, correct and incorrect choices of orange and green occurred equally often. There was a large increase in red errors on both 8- and 10-sec trials. A sample duration \times key color \times delay analysis of variance confirmed that the pattern of key choices for the three sample durations was affected by the delay [$F(8,60) = 12.5, p < .001$].

These data clearly illustrate the nature of the choose-short effect. Accuracy on the short-sample trials decreased as the delay interval was increased from 0 to 5 and 10 sec, but remained above chance levels. Accuracy on longer sample trials decreased to chance or below chance levels as the delay was increased. The lowered accuracy on 8- and 10-sec sample trials was due in part to increased incorrect choices of orange and green but was primarily the result of the increased incorrect choices of the short-sample (i.e., red) key. Post hoc Newman-Keuls tests confirmed that there was a significant increase in incorrect red choices on both 8- and 10-sec sample trials at the 10-sec delay.

These data are consistent with a retrospective, but not a prospective, model of how pigeons remember event duration. At the 0-sec delay, most errors were the result of incorrect choices between the prospectively easy, but retrospectively difficult, orange and green keys. At 5- and 10-sec delays, the subjects chose orange and green about equally often, as was expected from the retrospective model that assumes that longer delays make it increasingly difficult to distinguish between 8- and 10-sec samples. Red-orange errors also increased at the 5- and 10-sec delays, and actually increased more than did orange-green errors. At first glance, this increase would seem to support a prospective model that would predict that, when the task is made harder (i.e., when delays increase from 0 to 5 and 10 sec), subjects increasingly confuse instructions to peck red or peck orange on the choice phase of the trial. Two aspects of the data, however, are inconsistent with this interpretation. First, the increase in red-orange errors occurred primarily on longer sample (8 and 10 sec) trials as the delay was increased. The subjects correctly chose red on 2-sec trials more frequently than they incorrectly chose orange. On 8-sec trials, the subjects made fewer correct orange choices and many more red errors. From a prospective model, there is no a priori reason for the selective confusion of red and orange codes on some (long sample) but not other (short sample) trials. The second point that is important to note is that there is a ceiling effect in the frequency with which errors can be made. Because red-orange errors occurred less frequently than did orange-green errors during the 0-sec delay condition, there is more of an opportunity for red-orange errors to increase relative to orange-green errors. This factor will also be an important consideration in interpreting the discriminial distance data.

Table 3
Discriminal Distance Data from Experiment 2

	Pigeon					
	1	2	3	4	5	6
Delay = 0 sec						
AB	.83	.91	.96	.84	.93	.89
BC	.75	.72	.73	.72	.75	.77
AC	.85	.95	.99	.96	.98	.91
AB/BC	1.11	1.26	1.32	1.17	1.24	1.16
Delay = 5 sec						
AB	.79	.81	.82	.84	.82	.74
BC	.76	.76	.70	.74	.75	.81
AC	.79	.88	.88	.88	.88	.81
AB/BC	1.04	1.07	1.17	1.14	1.09	0.91
Delay = 10 sec						
AB	.80	.78	.79	.78	.72	.73
BC	.77	.79	.73	.78	.81	.82
AC	.78	.841	.84	.84	.83	.83
AB/BC	1.04	0.99	1.08	1.00	0.89	0.89

Discriminal distance analysis. Table 3 shows discriminial distance scores for each subject for each of the three delays.

At the 0-sec delay, the relative magnitudes of the AB, BC, and AC values and the direction of the AB to BC ratio were similar to those found in Experiment 1. The AC scores were typically large, reflecting the fact that red-green confusions were infrequent. The AB scores were larger than the BC scores, reflecting the fact that red-orange confusions were less frequent than were orange-green confusions. This difference replicates the results of Experiment 1 and supports a retrospective model of pigeons' memory for event duration.

At longer delays, the AB scores decreased, reflecting the fact that red-orange confusions increased as the delay increased. Changes in the discriminial distance scores over the delays were statistically significant [$F(4,20) = 17.9, p < .001$]. Figure 5 shows the ratio of AB scores to BC scores over the delays averaged across the 6 subjects.

As can be readily seen in this figure, the AB/BC ratio decreased as the delay increased. This decrease was statistically significant [$F(2,10) = 22.4, p = .0002$]. This decrease, which reflects the fact that red-orange errors increased more than did orange-green errors as the delay increased, suggests that the pigeons may have used prospective coding, coding the short houselight sample as "peck red" and the 8-sec sample as "peck orange," confusing these more frequently as the delay increased. There are, however, problems with this interpretation and, moreover, this pattern of results is also consistent with the retrospective subjective shortening hypothesis.

The problem with interpreting these results as supportive of a prospective coding model of memory for event duration is that the red-orange confusions tended to happen most often on certain types of trials. As is apparent in Figure 4, red and orange were confused most on 8-

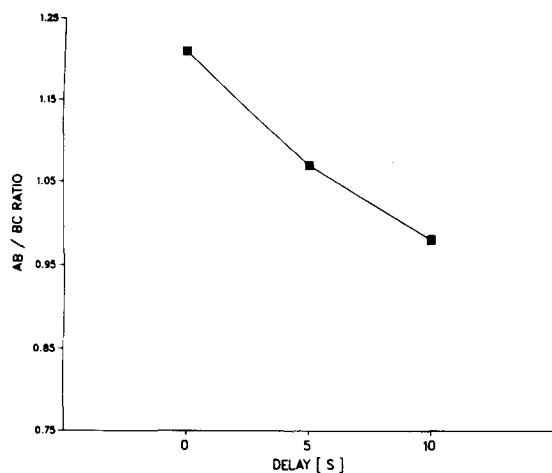


Figure 5. The ratio of AB to BC discriminial distance scores as a function of the delay interval.

and 10-sec sample trials, less so on 2-sec sample trials. If subjects code 2 sec of houselight as "peck red" and 8 sec as "peck orange," there is no a priori reason to expect more red-orange confusion on 8-sec trials than on 2-sec trials, because, in both cases, there should be equal opportunity to confuse the two codes during the delay.

The decrease in the AB/BC ratio is predicted by the retrospective subjective shortening hypothesis. The increase in red choices on long-sample trials as the delay is increased is the choose-short effect that has been observed in numerous experiments, and not simply confusion of red and orange. This explains why red-orange "confusions" are more frequent on 8-sec trials than on 2-sec trials.

Discussion

When there was no delay between sample presentation and the opportunity to choose between red, orange, and green keys, the pigeons often chose green on 8-sec trials and orange on 10-sec trials. Choices of orange on 2-sec trials and red on 8-sec trials were less frequent. This pattern, which is the same as that observed in Experiment 1, is consistent with a retrospective coding strategy of attempting to recall the actual duration of the 2-, 8-, and 10-sec samples, and of finding it difficult to discriminate between 8 and 10 sec. When delays were imposed between the sample presentations and the choice period, choices of orange on 2-sec trials and red on 8-sec trials increased. Both the pattern of errors and the calculated discriminial distance indexes on the delayed trials matched those predicted by the retrospective subjective shortening hypothesis.

EXPERIMENT 3

Taken together, the results of Experiments 1 and 2 support the retrospective subjective shortening model of memory for duration of events. The purpose of Experiment 3 was to further test retrospective and prospective

models. We adapted the multiple-sample procedure used by Spetch and Sinha (1989) to the discriminial distance experimental design. Retrospective and prospective models make explicit, and clearly different, predictions in this paradigm.

Briefly, the procedure we used was as follows. During baseline sessions, a procedure similar to the 0-sec delay condition in Experiment 2 was used. Red, orange, and green keys were correct after 2-, 8-, and 10-sec samples of houselight, respectively. During test sessions, the 2- and 10-sec sample trials were unchanged. The 8-sec trials, however, were different from baseline in that they were always preceded by a 2-sec sample. That is, on these trials the houselight was turned on for 2 sec, off for 1 sec, and then on again for 8 sec. Of key interest in Experiment 3 is pattern of choices on this type of trial. Prospective models predict increased red-orange confusions on these trials. From this view, subjects should code the first sample as red and the second as orange, and they should be more likely to confuse red and orange because of proactive interference. Retrospective models, on the other hand, predict the opposite type of error (that is, more green choices) because of the possibility of "summing across" the two sample presentations.

Method

Subjects and Apparatus. Four of the subjects used in Experiments 1 and 2 were employed in Experiment 3, as was the same apparatus.

Procedure. The subjects received approximately 14 sessions of baseline training identical to that of Experiment 1. This training was followed by three test sessions. Test sessions were identical to training sessions in all respects but one: Each 8-sec sample was preceded by a 2-sec sample. The interval between the two sample presentations was 1 sec. The data of primary interest were the number of times the subjects chose orange and green on 8-sec sample trials during baseline and test sessions.

Experiment 3 was replicated. The three test sessions were followed by another 12 baseline sessions and then another two test sessions.

Results

Figure 6 shows key choice data from the last three baseline sessions and the three test sessions. The baseline data resembled those of Experiments 1 and 2. On most trials, the pigeons correctly chose red, orange, and green after 2-, 8-, and 10-sec samples. The most frequent errors were green on 8-sec trials and orange on 10-sec trials. During test sessions, performance on 2- and 10-sec trials was essentially unaffected. On the critical 8-sec trials, which were always preceded by a 2-sec sample, green choices increased, but red choices did not. This finding is clearly at odds with prospective models of memory for event duration.

Choices of orange and green on 8-sec trials during baseline and test sessions were analyzed in a two-way repeated measures analysis of variance in which choice (orange or green) and phase (baseline or test) were the factors. In this analysis, the interaction was highly significant [$F(1,3) = 67.21, p = .0065$].

2 - 8 TEST

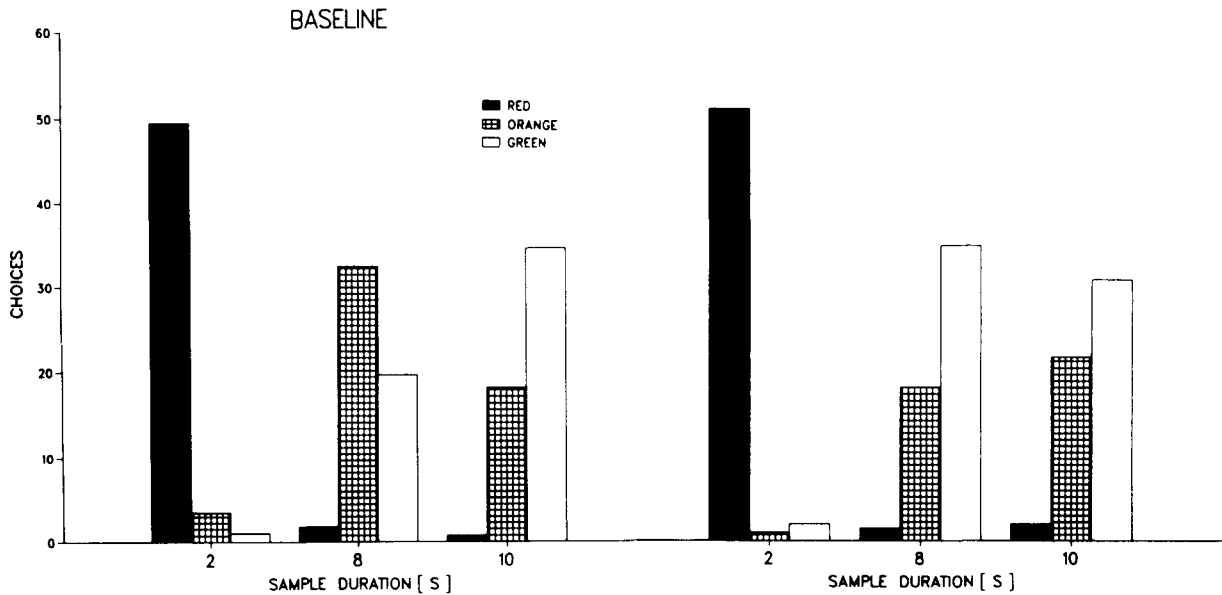


Figure 6. Choices of red, orange, and green keys on 2-, 8-, and 10-sec trials during baseline trials and during test sessions in which the 8-sec sample was preceded by a 2-sec sample.

The same pattern of results was seen during the replication. During the last two baseline sessions, the subjects chose orange an average of 20.5 times and green an average of 12 times on the 8-sec trials. During the two test sessions, they chose orange an average of 10.75 times and green an average of 24.25 times. Again, this change in choices was statistically significant [$F(1,3) = 18.49, p = .0261$].

Discussion

Both the results and conclusions of our multiple-sample experiment are similar to those reported by Spetch and Sinha (1989). In their experiment, short (2 sec) and long (10 sec) samples were preceded on test trials by either short or long samples to yield four trial types: short-short, short-long, long-long, and long-short. From a prospective coding hypothesis, two of these trial types should facilitate performance (short-short and long-long, called *consistent* trials), whereas the other trial types (short-long and long-short, called *inconsistent* trials), should produce interference. From this view, consistent trials create the opportunity for rehearsal of a code such as "peck red," which should facilitate correct key choice. Contrary to the prospective coding hypothesis, accuracy was not related to consistency. Correct performance was reduced on short-short and long-short trials and increased on long-long and short-long trials.

Although there is good evidence that pigeons use prospective coding in some paradigms (e.g., Spetch, Wilkie, & Skelton, 1981), the results of this and the preceding experiments, as well as other recent research (e.g., Spetch & Sinha, 1989), suggest that coding seems to be retrospective in event-duration discrimination tasks.

In future research, it will prove interesting to attempt to determine why different coding strategies are used in different tasks. It will also be interesting to attempt to specify in more detail the way in which temporal durations are represented and the mechanisms that underly the subjective shortening process. We still do not know, for example, whether shortening is the result of a trace decay type of process, the loss of "clock ticks" stored in an internal clock's working memory, or other processes.

REFERENCES

- CONRAD, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, *55*, 75-83.
- CHURCH, R. M. (1980). Short-term memory for time intervals. *Learning & Motivation*, *11*, 208-219.
- FETTERMAN, J. G., & MACEWEN, D. (1989). Short-term memory for responses: The "choose-small" effect. *Journal of the Experimental Analysis of Behavior*, *52*, 311-324.
- GILBERT, S. G., & RICE, D. C. (1979). NOVA SKED II: A behavioral notation language utilizing the Data General Corporation real-time disk operating system. *Behavior Research Methods & Instrumentation*, *11*, 71-73.
- HONIG, W. K. (1978). Studies of working memory in the pigeon. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 211-248). Hillsdale, NJ: Erlbaum.
- HONIG, W. K., & SPETCH, M. (1988). Short-term memory for rate of alternation in the pigeon. *Bulletin of the Psychonomic Society*, *26*, 152-154.
- 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.
- ROBERTS, S. (1982). Cross-modal use of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes*, *8*, 2-22.
- ROITBLAT, H. L. (1980). Codes and coding processes in pigeon short-term memory. *Animal Learning & Behavior*, *8*, 341-351.
- SPETCH, M. L. (1987). Systematic errors in pigeons' memory for event

- duration: Interaction between training and test delay. *Animal Learning & Behavior*, **15**, 1-5.
- SPETCH, M. L., & RUSAK, B. (1989). Pigeons' memory for event duration: Intertrial interval and delay effects. *Animal Learning & Behavior*, **17**, 147-156.
- SPETCH, M. L., & SINHA, S. S. (1989). Proactive effects in pigeons' memory for event duration: Evidence against the coding model. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 347-357.
- 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. (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.
- SPETCH, M. L., WILKIE, D. M., & SKELTON, R. W. (1981). Control of pigeons' keypecking topography by a schedule of alternating food and water reward. *Animal Learning & Behavior*, **9**, 223-229.
- WILKIE, D. M. (1988). Proactive effects in pigeons' timing behavior: Implications for an internal-clock model. *Animal Learning & Behavior*, **16**, 132-136.

(Manuscript received August 10, 1989;
revision accepted for publication November 25, 1989.)