

Temporal Context Effects in Pigeons' Memory for Event Duration

MARCIA L. SPETCH

The University of Alberta, Edmonton, Alberta, Canada

AND

BENJAMIN RUSAK

Dalhousie University, Halifax, Nova Scotia, Canada, and McMaster University, Hamilton, Ontario, Canada

These experiments extend previous studies of how pigeons' memory for the duration of events is affected by the length of the delay and intertrial intervals. Pigeons were trained using a delayed matching-to-sample procedure to peck one color following a long sample and another color following a short sample. Experiments 1a and 1b tested a prediction, derived from the relative duration hypothesis, that systematic errors in response to delay manipulations should be greater in the context of a short intertrial interval than in the context of a long one. Pigeons were given delay tests in the context of either 10-s or 60-s intertrial intervals. Systematic errors in response to delay manipulations were of greater magnitude in the context of the 10-s intertrial interval, as predicted by the relative duration hypothesis. Experiments 2a and 2b examined whether stable delay and intertrial intervals during training are important for the development of a temporal reference memory. Variability in the intertrial or delay interval during training had only minor effects on acquisition and did not affect pigeons' tendency to make systematic errors in response to subsequent manipulations of these intervals. Moreover, even during acquisition, birds trained with variable delays showed a systematic increase in the tendency to make choices appropriate to the short sample as a function of the training delay, with the least bias occurring at an intermediate delay. Thus, reference memory appeared to be based on an average of the delays experienced during training. © 1992 Academic Press, Inc.

This research was supported by Natural Sciences and Engineering Research Council of Canada operating grants (OGP0038861 to M.S. and A0305 to B.R.). We thank E. Buckle, R. Willson, and C. Gagne for assistance, and Dallas Treit for helpful comments. The results of Experiment 1a were presented at the conference on Cognitive Aspects of Stimulus Control, Halifax, N.S., in June 1989 and are summarized in the conference proceedings (Spetch & Rusak, 1992). Reprints may be obtained from Marcia Spetch, Department of Psychology, University of Alberta, Edmonton, Alberta, Canada T6G 2E9.

Pigeons that have been trained to make choices based on the duration of a preceding event often make systematic errors when the delay between the event and the opportunity to choose is varied (e.g., Spetch & Wilkie, 1982; Spetch, 1987). A tendency to make more choices appropriate to the short sample as the delay is lengthened (the "choose-short effect," Spetch & Wilkie, 1982) has been found for rats (Church, 1980) and pigeons (e.g., Spetch & Wilkie, 1982, 1983; Spetch, 1987). Pigeons have also been found to display a tendency to make choices appropriate to the long sample (the "choose-long effect") when the delay is shortened relative to the training value (Spetch & Wilkie, 1983; Spetch, 1987). Choice biases analogous to the choose-short effect have also been observed in pigeons' memory for number of responses (the "choose-small effect," Fetterman & MacEwen, 1989), and in pigeons' memory for rate of alternation (the "choose-fast effect," Honig & Spetch, 1988).

Several interpretations have been offered for why variations in the delay interval cause systematic choice errors in delayed choice tasks using event durations as the sample stimuli. Spetch and Wilkie (1983) proposed a subjective shortening model, in which increases in the delay were assumed to decrease the remembered duration of the sample stimulus. According to this account, when tested with a delay longer than the training delay, the current samples would seem shorter than the reference memory values based on the training delay; when tested with a delay shorter than the training values, the samples would seem longer than the reference memory values. This view assumes that the duration of the sample is remembered in an analogical form and is assessed retrospectively at the time of choice. Others (Church, 1980; Kraemer, Mazmanian, & Roberts, 1985) have suggested that the choose-short effect reflects a choice bias that occurs when the sample duration or some non-temporal code for the sample duration (e.g., a prospective response decision) has been forgotten. Although many results have favored the subjective shortening model (e.g., Spetch & Sinha, 1989; Wilkie & Willson, 1990), some results have favored models based on categorical/prospective coding (Kraemer *et al.*, 1985). Still other investigators (Killeen & Fetterman, 1988) have interpreted the choose-short effect in terms of nonmemorial processes. Although the behavioral theory of timing proposed by Killeen and Fetterman (1988) can account for the choose-short effect, as presently formulated, it does not appear to account for the choose-long effect.

Recently, we (Spetch & Rusak, 1989) reported some results that were not directly predicted by any of the above accounts of the systematic choice errors in memory for duration. Specifically, we found that similar effects occur when pigeons are tested with intertrial intervals (ITI) that are longer or shorter than the training value: On trials preceded by a longer-than-normal ITI, pigeons are more likely to make choices appropriate to the short sample event, whereas on trials preceded by a shorter

than normal ITI, they are more likely to make long choices (Spetch & Rusak, 1989). Although some quantitative differences were observed, the qualitative similarity of the effects of delay and ITI manipulations suggested that a common mechanism might be operating.

To account for this similarity we proposed a modification of the subjective shortening model, in which remembered event durations are assumed to depend upon the temporal context in which they occur (Spetch & Rusak, 1989, 1992). Specifically, we suggested that the ITI and the delay might form a common "temporal background" against which the remembered duration of the sample is evaluated. According to this "relative duration hypothesis," manipulations that decrease the temporal background (i.e., decreases in the delay or ITI) should make the sample seem longer, whereas manipulations that increase the temporal background (increasing the delay or ITI) should make the sample seem shorter. Although this hypothesis did not account for all of the details of the results we reported (Spetch & Rusak, 1989), the fit was qualitatively good and suggested that further tests of the hypothesis were warranted.

The studies reported here further explored the role of temporal context in regulating memory for duration. Experiment 1 was designed to test a direct prediction of the relative duration hypothesis, namely, that the duration of the ITI should influence the effect of varying delay intervals. In Experiment 2, we examined whether variability in the temporal context (ITI or delay) during acquisition would influence the formation or nature of a temporal reference memory.

GENERAL METHODS

Apparatus

The experimental environments consisted of standard two-key and three-key pigeon chambers. Pecking keys required a force of about .25N to operate. Choice stimuli were always presented on the two side keys when three-key chambers were used. A grain feeder was centered below the pecking keys, and grain presentations were accompanied by illumination of a lamp in the feeder. The houselight was centered at the top of the response panel. Exhaust fans provided ventilation and background masking noise.

In the chambers used for Experiments 1a and 2a, stimulus projectors mounted behind each key were used to transilluminate them with white, red, or green light. For the chambers used in Experiments 1b and 2b, a color monitor was used instead of stimulus projectors to present stimuli on the pecking keys. Patches of color (red and green) displayed on the monitor were projected onto the keys through lenses mounted at the back of the chamber.

Experimental contingencies and data recording were controlled by a computer located in an adjacent room.

Procedure

All experiments employed variations of the delayed symbolic matching-to-sample (DSMTS) task. Trials began with the presentation of a sample stimulus, which lasted for either a short or a long duration. Sample termination was followed by a delay interval of a specified duration, and then red and green lights were presented on two horizontally aligned pecking keys to provide choice stimuli. One color was designated as correct following short samples, the other color was correct following long samples. A peck at the correct choice stimulus was reinforced with access to grain for 4 s. A peck at the incorrect choice stimulus resulted in termination of the trial without food. Trials were separated by an ITI. In Experiments 1a and 2a, the chamber was dark during both the ITI and the delay interval. In Experiments 1b and 2b, the houselight was continuously illuminated at a low level throughout each session; this procedure was necessary because the colors projected onto the keys from the monitor did not provide as much illumination as standard projector lights, and some birds would not respond reliably without a minimum level of illumination.

Within each session, the short and long sample durations occurred equally often in a randomly determined order. The right-left location of the correct comparison stimulus was varied across trials and was always balanced within sample duration and any other manipulated variable (e.g., delay or ITI). The designation of the colors as correct for short or long was balanced to the extent possible over subjects in all groups.

A correction procedure was used during all baseline training phases. Following an error, the current trial conditions (e.g., sample duration, arrangement of choice stimuli) were duplicated on the next trial. All reported accuracy scores were derived from performance on non-correction trials only. The correction procedure was not in effect during any of the test phases.

Other procedural details varied across experiments and are described in the individual methods sections.

Data Analysis

Choice behavior was assessed in terms of several measures. Accuracy in each condition was assessed by the overall percentage of correct choices made. Systematic error tendencies were assessed in terms of the percentage of short choices (i.e., choices of the comparison stimulus that was correct for short samples) in each condition. Because an equal number of short and long sample trials were scheduled in each condition, this measure generates a value of 50% when no systematic tendency to choose one stimulus over the other exists. A value above 50% indicates a disproportionate number of short choices, whereas a value below 50% in-

icates a disproportionate number of long choices. These two measures provided the primary data on which the statistical analyses were based. However, we also calculated two additional measures of accuracy and bias based on signal detection theory (Davison & Tustin, 1978). Log d , a bias-free measure of discriminability is calculated by

$$\log d = 0.5 \log(B_w B_z / B_x B_y), \quad (1)$$

where B_w and B_z are the number of correct choices following short and long samples, respectively, and B_y B_x are incorrect choices of the short and long alternatives, respectively. For this measure, positive values indicate above-chance performance, whereas negative values indicate below-chance performance. Log b , a measure of response bias, is calculated by

$$\log b = 0.5 \log(B_w B_y / B_x B_z), \quad (2)$$

where B_w , B_y , B_x , and B_z are defined as in Eq. (1). For this measure, positive values indicate a bias toward the short choice alternative, whereas negative values indicate a long choice bias. Although we present these measures in the tables, we did not elect to use them as our primary measure of analysis because they can be problematic whenever error rates are extremely low. First, the values cannot be calculated when no errors are made for one of the sample types, because this results in a value of zero in the denominator. Second, the bias measures can be quite variable and may not be appropriate when overall error rates are extremely low (see Fetterman & MacEwen, 1989). Nevertheless, for the values that could be calculated, these measures revealed a pattern of results similar to those provided by our two primary measures. Where possible, we provided some statistical analysis based on these signal detection measures to corroborate our conclusions.

Unless otherwise noted, all data presented in the tables and figures are based on the average of all test sessions in each condition.

EXPERIMENT 1

In this experiment, we asked whether the duration of the baseline ITI would affect pigeons' responses to delay manipulations. According to the relative duration hypothesis, event durations are remembered relative to the duration of the background (i.e., ITI plus delay) in which they appeared. Variations in the delay interval thus would have proportionally larger effects on the total background duration when the ITI is short than when the ITI is long. A direct prediction of the relative duration hypothesis is that pigeons' systematic errors in response to delay manipulations should be of greater magnitude in the context of a short ITI than in the context of a long ITI. Experiments 1a and 1b are replications using slightly different procedures to test this prediction.

Experiment 1a

Method

Subjects. The subjects were four experimentally naive adult Silver King pigeons. Mixed grain obtained primarily during experimental sessions maintained the birds at 85–90% of their free-feeding weights. The birds were housed in individual wire-mesh cages; water and grit were freely available.

Procedure. Preliminary training. Each bird received one or two sessions of magazine training, followed by a few sessions with an autoshaping procedure to establish pecking at each of the side keys when illuminated with either red or green light.

Training and testing. The basic DSMTS procedure described in the General Methods section was used during training and testing. The samples were 2-s and 10-s houselight presentations. The delay interval was 0 s during training. During test sessions, 24 of the 48 trials contained a 0-s delay, 12 contained a 5-s delay, and 12 contained a 10-s delay. Each bird was exposed to two conditions of training and testing (ITI-10 and ITI-60) that differed only with respect to the duration of the ITI (10 s on every trial or 60 s on every trial). Each bird was trained initially for 25 sessions and then given 10 delay testing sessions under one ITI condition (ITI-10 for Birds 2 and 4; ITI-60 for Birds 1 and 3). Then, the ITI conditions were reversed and the birds were given 10 additional sessions of training and 10 sessions of delay testing with the other ITI.

Results

Table 1 shows, for each bird, the percentage of correct choices and the percentage of short choices, as well as the signal detection measures of accuracy (log *d*) and bias (log *b*) at each delay in each ITI condition. Missing values for the two signal detection measures indicate that one of the cells contained a value of zero, and hence the values could not be calculated. Figure 1 shows the mean percentage of short choices as a function of delay for each ITI condition.

Overall accuracy declined as a function of delay but did not appear to differ across the two ITI conditions. A repeated-measures ANOVA on the percentage of correct choices revealed a significant effect of delay, $F(2, 6) = 198.24, p < .001$, but not of ITI, $F(1, 3) = 2.31, p > .2$. The interaction between delay and ITI was not significant, $F(2, 6) = 4.32, p > .05$.

As revealed in Fig. 1, the tendency to make short choices increased as a function of delay (the choose-short effect), and the magnitude of this increase was greater when the ITI was 10 s than when it was 60 s. These observations were confirmed by a repeated measures ANOVA on the percentage of short choices, which revealed a significant main effect of

TABLE 1
Measures of Accuracy (Percent Correct and Log *d*) and Bias (Percent Short and Log *b*) for All Birds At Each Delay under the Two ITI Conditions in Experiment 1a

Delay	ITI-10			ITI-60		
	0	5	10	0	5	10
<i>Bird 1</i>						
% correct	85.8	65.9	65.0	91.7	77.5	64.2
Log <i>d</i>	.809	.420	.301	—	.537	.254
% short	45.0	65.0	65.0	41.6	50.8	52.5
Log <i>b</i>	-.185	.333	.301	—	.020	.047
<i>Bird 2</i>						
% correct	93.3	59.2	56.7	95.9	80.0	75.0
Log <i>d</i>	1.176	.186	.144	1.373	.626	.485
% short	47.5	67.5	71.7	49.2	56.7	55.0
Log <i>b</i>	-.182	.331	.413	-.094	.188	.117
<i>Bird 3</i>						
% correct	84.6	59.2	57.5	82.2	62.9	53.6
Log <i>d</i>	.762	.191	.220	.666	.318	.073
% short	46.3	69.1	80.8	48.2	74.6	69.6
Log <i>b</i>	-.127	.367	.658	-.054	.527	.363
<i>Bird 4</i>						
% correct	85.0	55.0	49.2	80.0	61.3	51.7
Log <i>d</i>	.759	.101	-.019	.604	.201	.032
% short	47.5	68.3	74.1	47.6	53.1	65.0
Log <i>b</i>	-.086	.338	.458	-.067	.057	.269

delay, $F(2, 6) = 34.27, p < .001$, and ITI, $F(1, 3) = 12.82, p < .05$, as well as a significant interaction between delay and ITI, $F(2, 6) = 5.16, p < .05$.

The effect of the ITI on the magnitude of the choose-short effect is also apparent from the log *b* values presented in Table 1. For example, at the 10-s-delay, log *b* reveals a greater short bias in the 10-s ITI condition (mean = .458) than in the 60-s ITI condition (mean = .199), $t(3) = 10.28, p < .005$. However, log *d*, the measure of accuracy, does not differ significantly between the 10-s ITI (mean = .162) and the 60-s ITI (mean = .211) at the 10-s delay, $t = 0.47, p > .5$.

Experiment 1b

Method

Subjects. The subjects were three White King pigeons, which had previously served in another experiment. Their previous experience is reported in Experiment 2b (Group C-Del). Housing and deprivation conditions were as described in Experiment 1a.

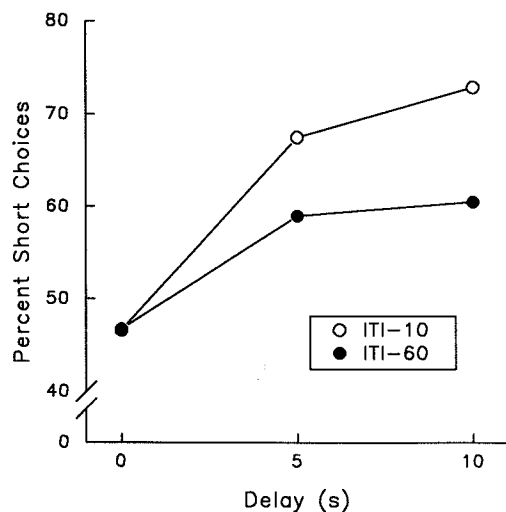


FIG. 1. Mean percentage of short choices as a function of delay interval in the context of a 10-s ITI and a 60-s ITI in Experiment 1a.

Procedure. The birds did not require any preliminary training. The basic DSMTS procedure was as described in the General Methods section, and the samples were 2-s and 8-s food presentations. The birds had previously been trained with a 5-s delay, and therefore the baseline delay for this experiment was 5 s rather than 0 s.

Each bird received 10 sessions of training and 10 sessions of delay testing under each of two successively presented ITI conditions (ITI-10 and ITI-60). Birds 6-2 and 6-4 were exposed to the ITI-10 condition first, and Bird 6-3 was exposed to the ITI-60 condition first. During delay testing, 24 of the 48 trials in each session contained the baseline 5-s delay, 12 contained a 0-s test delay, and 12 contained a 20-s test delay.

Results

Table 2 shows the percentage of correct choices, log d scores, percentage of short choices, and log b scores for each bird at each delay under the two ITI conditions. Figure 2 shows the mean percentage of short choices as a function of delay for the two ITI conditions.

Overall accuracy declined as a function of delay in both ITI conditions and appeared to be slightly lower in the 10-s ITI condition than in the 60-s ITI condition. An ANOVA on the percentage of correct choices revealed a significant effect of delay, $F(2, 4) = 54.83$, $p < .01$, and the effect of ITI was marginally significant, $F(1, 2) = 18.55$, $p = 0.5$. The interaction between delay and ITI was not significant, $F(2, 4) = 2.61$, $p > .1$.

TABLE 2
Measures of Accuracy (Percent Correct and Log d) and Bias (Percent Short and Log b) for All Birds at Each Delay under the Two ITI Conditions in Experiment 1b

Delay	ITI-10			ITI-60		
	0	5	20	0	5	20
<i>Bird 6-2</i>						
% correct	86.7	78.8	49.2	94.7	83.6	51.0
Log d	.846	.680	—	1.292	.709	.028
% short	45.0	63.7	99.1	46.7	50.3	81.0
Log b	-.197	.411	—	-.298	.008	.630
<i>Bird 6-3</i>						
% correct	85.9	81.2	50.0	97.9	92.6	61.6
Log d	.972	.668	.000	—	1.104	.226
% short	39.2	57.1	98.3	47.9	49.0	64.2
Log b	-.495	.209	1.762	—	-.061	.270
<i>Bird 6-4</i>						
% correct	72.5	70.4	54.2	87.4	84.2	56.4
Log d	.443	.409	.099	.888	.801	.117
% short	40.8	37.9	75.8	44.4	41.4	60.4
Log b	-.206	-.264	.503	-.235	-.309	.187

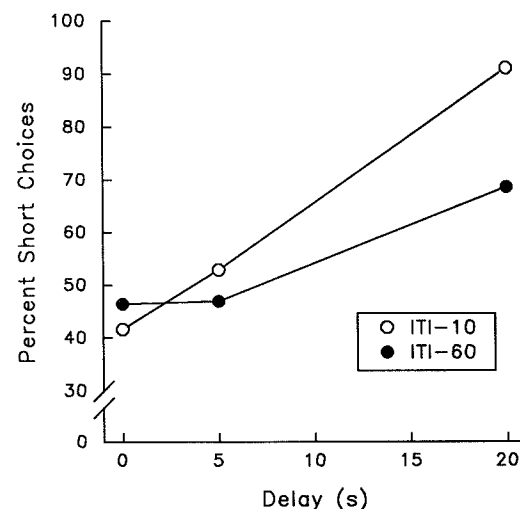


FIG. 2. Mean percentage of short choices as a function of delay interval in the context of a 10-s ITI and a 60-s ITI in Experiment 1b.

As in Experiment 1a, the tendency to make short choices increased as a function of delay, and the magnitude of this increase was greater when the ITI was 10 s than when it was 60 s. An ANOVA on the percentage of short choices revealed a significant main effect of delay, $F(2, 4) = 42.23, p < .01$, and a significant interaction between delay and ITI, $F(2, 4) = 8.79, p < .05$. The main effect of ITI failed to reach significance, $F(1, 2) = 9.23, p > .05$. The same pattern of results is apparent from the log b scores, although the missing values precluded any meaningful statistical analyses on this measure.

Discussion

These results indicate that the absolute duration of the ITI influences the magnitude of the choose-short effect generated by manipulations of the delay interval. Pigeons displayed a larger increase in the tendency to make short choices at the long test delays when the ITI was 10 s than when it was 60 s. This effect is in the direction predicted by the relative duration hypothesis. The generality of this effect is suggested by its appearance in two experiments that differed in a number of experimental details, including type of sample event and baseline delay interval.

EXPERIMENT 2

It is often assumed that timing behavior in animals entails a comparison between the duration value on the current trial (working memory) and a value stored in reference memory on the basis of previous trials (e.g., Gibbon & Church, 1984; Spetch & Wilkie, 1983). For example, both the original subjective shortening model (Spetch & Wilkie, 1983) and the relative duration modification of this model (Spetch & Rusak, 1989, 1992) assume that systematic errors arise because of discrepancies between working memory and reference memory durations. It is not yet clear, however, which conditions are important for the formation of a stable temporal reference memory.

When the duration of the delay interval and the ITI are held constant during training, the formation of a reference memory for the duration of each stimulus seems straightforward. For example, according to the subjective shortening model, during training with a constant delay, animals are assumed to learn and retain in a reference memory the durations of the short and long samples, as they are perceived at that training delay, together with the choices associated with these perceived durations. This reference memory is assumed to remain reasonably stable during testing provided that many of the test trials continue to use the training delay on which the reference memory is based. However, when the delay is shifted to a new value and this new value is presented on all trials of the session, the reference memory is assumed to gradually change to include

a representation of the sample durations as they are perceived at the new delay (Spetch & Wilkie, 1983).

Similar assumptions are made within the relative duration framework, except that the values retained in reference memory are assumed to consist of values for the sample duration that are relative to the temporal background provided by the ITI and delay. In this case, it is assumed that the subjective duration of the short sample, relative to the background duration of the training ITI and delay, is associated with one choice and the subjective duration of the long sample, relative to the training ITI and delay, is associated with the other choice. These stable values provide the referents against which the values on any test trial are assessed. Systematic errors emerge when alterations in the delay or ITI make the samples appear shorter or longer than these learned reference values.

These descriptions are based on the assumption of constancy in the temporal context provided during training. It is not obvious from either the subjective shortening model or the relative duration hypothesis what values would be used as the referents if the duration of the delay or ITI varied during training. According to both of these models, trial-by-trial alterations in the temporal context during the training sessions should alter the subjective durations of the short and long samples on different trials. Such variation might have any of several effects.

The first possibility we considered is that variation in the temporal context during training might interfere with pigeons' ability to form a stable reference memory and lead to performance deficits on the baseline task, either in rate of acquisition or quality of asymptotic performance. Although this possibility seems reasonable on the basis of the models outlined above, there appears to be little suggestion in either the timing or the memory literature that acquisition or asymptotic performance is disrupted by variability of the ITI. For example, although we are not aware of any studies directly comparing the effects of variable versus constant ITI values on the acquisition of a temporal discrimination, several studies have used variable ITIs and obtained accurate temporal discrimination performance (e.g., Platt & Davis, 1983; Roberts, Cheng, & Cohen, 1989). In delayed matching procedures with non-temporal samples, memory performance has been found to vary as a function of the mean ITI duration, but it is not disrupted by variability in the ITI (Roberts & Kraemer, 1982). Finally, classical conditioning, which may share some mechanisms with timing (e.g., Holder & Roberts, 1985; Roberts & Holder, 1985), is affected by the mean duration of the ITI but not by variability of the ITI (e.g., Jenkins, 1984). Nevertheless, these results would not preclude an effect of variability of the ITI on the acquisition of a temporal reference memory, because some variables may have different effects in autoshaping and timing procedures. For example, autoshaped responding is more sensitive to the average duration of the ITI

than to the duration of the immediately preceding interval (see Jenkins, 1984; Lucas & Wasserman, 1982), whereas the duration of the immediately preceding ITI can produce large choice biases in our timing task (Spetch & Rusak, 1989). Similarly, intertrial reinforcers appear to disrupt auto-shaped responding and temporal discrimination performance in subtly different ways (Wilkie, Symons, & Tees, 1988).

The second possible effect of variation in the temporal context we considered was derived from the relative duration framework: Because the background time is not constant, the temporal background might cease to modulate judgments of the sample duration. That is, during training with a variable temporal background, pigeons might learn to rely exclusively, or more heavily, on the absolute duration of the sample event. As a result, subsequent manipulations of the delay or ITI would not be expected to cause systematic errors because these values would not form an essential part of the assessment of the sample duration. Thus, the prediction derived from this possibility is that variability in the ITI or delay during training would reduce the systematic choice biases that normally emerge when the duration of either the delay or ITI is manipulated. Depending on the degree of flexibility in the encoding of temporal information, this might occur with or without a retardation of acquisition.

The third possibility we considered is that when the temporal context is variable, the average of the durations experienced during training would serve as the reference value. That is, the reference memory duration for "long" might be the subjective duration of the long sample as it appears in the context of the average of the training ITIs and the average of the training delays. In this case, one would expect that manipulations of the delay or ITI should still generate systematic errors. Specifically, delay or ITI values that are longer than the average of the training range should generate choose-short effects, whereas delay or ITI values that are shorter than the average value should generate a choose-long tendency.

This third possibility is perhaps most congruent with models of timing behavior that assume some means of averaging across variable individual signal durations to generate a mean reference memory value (e.g., Gibbon, Church, & Meck, 1984; Gallistel, 1990). Within the relative duration framework, trial-by-trial variations in temporal context may result in greater variation in the individual signal durations encoded as short or long. However, if the average duration of the temporal context remains the same over sessions, a stable mean reference memory value for short and long signal durations should develop. Performance on individual trials should still reflect a comparison of the current working memory value with the mean reference memory value and, consequently, systematic errors should still occur when the working memory values deviate from the mean.

We explored these possibilities by manipulating the constancy of the

ITI (Experiment 2a) or the delay (Experiment 2b) during training. Experiment 2a was conducted before the third possibility was considered and the data collected did not permit a direct assessment of that possibility. Experiment 2b was conducted later and included tests appropriate for assessing all three possibilities.

Experiment 2a

In this experiment, birds were trained with either a constant 45-s ITI (Group C-ITI) or with an ITI that varied across trials but was 45 s on average (Group V-ITI). The performance of these two groups was compared during acquisition and subsequently during tests in which the delay or ITI was varied.

Method

Subjects. The subjects were 13 adult White King pigeons. Six randomly selected pigeons were assigned to a constant ITI group (Group C-ITI); the remaining birds were assigned to a variable ITI group (Group V-ITI). Three subjects in each group had served briefly in an autoshaping experiment prior to this experiment; the remaining birds were experimentally naive. All were maintained at 85 to 90% of their free-feeding weight and housed as described in Experiment 1.

Procedure. Preliminary training. This phase consisted of one session of magazine training for the naive birds and a few sessions of autoshaping for all subjects to establish reliable pecking to red and green illuminations of each key.

0-s delay training. The basic DSMTS procedure used for this experiment was as described in the General Methods section, with one exception: Trials began with the presentation of white light on the center pecking key and a single peck to this trial-initiating stimulus was required to produce the sample. The samples were 2- and 8-s presentations of the illuminated grain feeder, and the baseline training delay was 0 s. The constancy of the ITI duration was manipulated during training. For Group C-ITI, the ITI was 45 s on every trial. For Group V-ITI, the ITI varied across trials, with the values being 5, 15, 30, 45, 60, 75, and 85 s; these were scheduled to occur equally often but in a randomly determined order within each session, so that the arithmetic mean of the ITIs experienced would be 45 s. Baseline training continued until each subject reached an accuracy criterion of at least 85% correct for five consecutive sessions.

Variable delay testing. The delay testing procedure lasted for 10 sessions and was identical for the two groups. Each session contained 24 0-s delay (baseline) trials, 12 5-s delay trials, and 12 10-s delay trials; these occurred in a randomly determined order. The ITI was constant at 45 s throughout this test phase for both groups.

5-s and 10-s delay training. The procedure used during this phase was

TABLE 3

Measures of Accuracy (Percent Correct and Log d) and Bias (Percent Short and Log b) for Groups C-ITI and V-ITI during the First Phase of Variable Delay Testing in Experiment 2a

Delay	C-ITI			V-ITI		
	0	5	10	0	5	10
% correct						
Mean	92.2	58.2	57.9	91.1	58.1	54.7
SEM	1.7	5.1	3.4	2.4	3.0	0.9
Log d						
Mean	1.211	.163	.163	1.150	.171	.100
SEM	.139	.107	.079	.140	.058	.019
% short						
Mean	49.4	55.8	59.9	48.4	60.9	63.6
SEM	1.8	2.3	2.6	1.0	5.4	5.8
Log b						
Mean	-.005	.111	.199	-.156	.225	.266
SEM	.104	.041	.061	.060	.120	.115

identical to that used during 0-s delay training except that delays of 5 s and 10 s were used in successive blocks of 20 sessions each. As during the 0-s delay training phase, the ITI was constant for Group C-ITI and varied for Group V-ITI.

Variable delay and ITI testing. During this phase, both groups were first given 10 sessions of variable delay testing. The ITI was constant at 45 s for both groups. The baseline 10-s delay occurred on a randomly selected 75% of the trials and two test delays occurred equally often on the remaining trials. On alternate sessions, these test delays were 0 s and 20 s or 5 s and 15 s. Following delay testing, birds in both groups were given five baseline sessions, during which their previous 10-s delay training conditions were reinstated (i.e., the ITI was constant for Group C-ITI and variable for Group V-ITI). Next, both groups were given 10 sessions of ITI manipulation tests, during which the delay was held constant at 10 s. The 45-s ITI preceded a randomly selected 75% of the trials and longer or shorter ITIs preceded the remaining trials. On alternate sessions, the test ITIs were 5 s and 85 s or 15 s and 75 s.

Results

0-s delay training. Subjects in Group C-ITI required a mean of 16.5 sessions to reach the accuracy criterion, while those in Group V-ITI required a mean of 19.7 sessions. Although the difference was in the

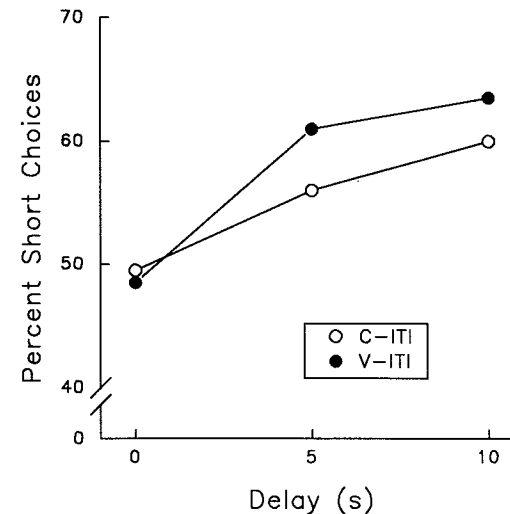


FIG. 3. Mean percentage of short choices as a function of delay interval for birds trained with a constant ITI (C-ITI) or a variable ITI (V-ITI) in Experiment 2a.

appropriate direction according to the first hypothesis, it was only marginally significant, $t(11) = 1.90$, $p < .05$, one-tailed.

Variable delay testing. Table 3 shows the percentage of correct choices, percentage of short choices, and the log b and log d scores for each subject. The mean percentage of short choices at the three delays is shown in Fig. 3.

Overall accuracy declined as a function of delay but did not differ substantially for the two groups. A between-within ANOVA on the percentage of correct choices revealed a significant main effect of delay, $F(2, 22) = 148.76$, $p < .001$, but not of group, $F(1, 11) = 0.21$, $p > .2$, and no significant interaction between delay and group, $F(2, 22) = 0.24$, $p > .2$. A similar ANOVA on the log d scores also revealed a significant main effect of delay, $F(2, 22) = 102.16$, $p < .001$, and no other significant effects ($ps > .2$).

Both groups displayed a large choose-short tendency at the long delays. A between-within ANOVA on the percentage of short choices revealed a significant main effect of delay, $F(2, 22) = 10.12$, $p < .01$, but not of group, $F(1, 11) = 0.38$, $p > .2$, and no significant interaction between delay and group, $F(2, 22) = 0.58$, $p > .2$. A similar ANOVA on the log b scores also revealed a significant effect of delay, $F(2, 22) = 9.30$, $p < .01$, and no other significant effects ($ps > .2$).

5-s and 10-s delay training. Pigeons in both groups displayed a steady improvement in performance over training sessions with the constant 5-s delay and subsequently with the constant 10-s delay (Fig. 4). Between-

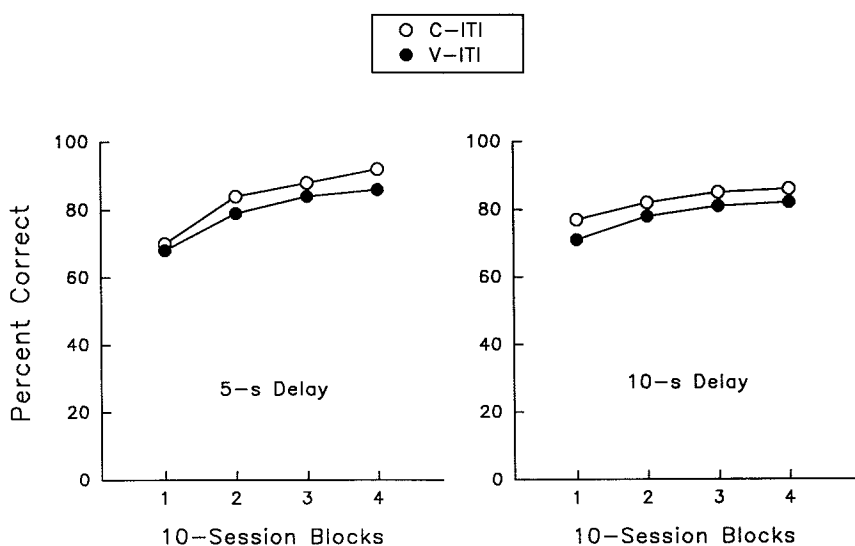


FIG. 4. Mean percentage of correct choices during 5-s and 10-s delay training for birds trained with a constant ITI (C-ITI) or a variable ITI (V-ITI) in Experiment 2a.

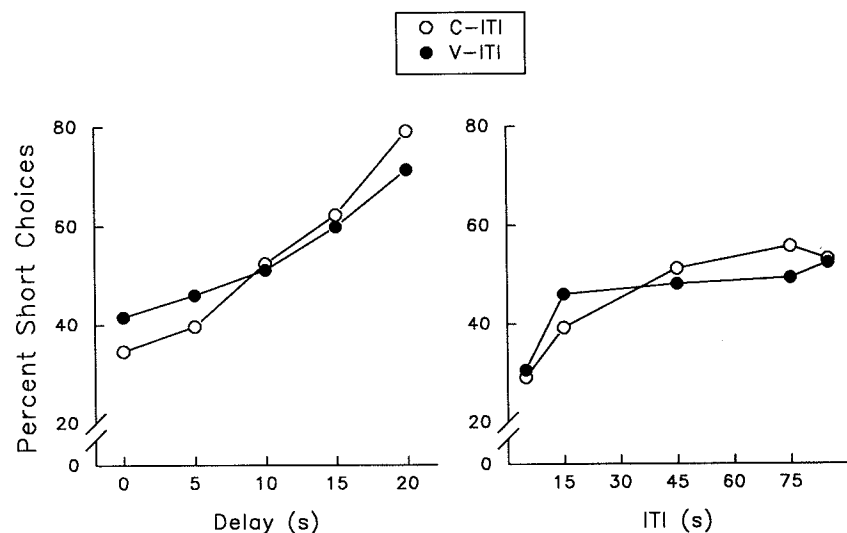


FIG. 5. Mean percentage of short choices during the delay manipulation tests (left panel) and ITI manipulation tests (right panel) for birds trained with a constant ITI (C-ITI) or a variable ITI (V-ITI) in Experiment 2a.

TABLE 4
Mean Percentage of Correct Choices and Percentage of Short Choices for Groups C-ITI and V-ITI during the Second Two Test Sessions of Experiment 2a

Variable delay tests with 10 s baseline delay										
Delay	C-ITI					V-ITI				
	0	5	10	15	20	0	5	10	15	20
% correct										
Mean	74.8	81.7	86.4	77.7	65.8	79.3	89.3	80.9	74.1	66.9
SEM	2.7	1.6	4.1	2.4	1.1	3.3	1.7	3.9	5.4	4.0
% short										
Mean	34.7	39.5	52.4	62.3	79.0	41.6	46.0	51.0	59.9	71.4
SEM	3.6	3.2	1.3	2.5	3.6	5.7	2.2	1.3	4.6	5.0
Variable ITI Tests										
ITI	C-ITI					V-ITI				
	5	15	45	75	85	5	15	45	75	85
% correct										
Mean	71.1	80.6	90.0	90.5	91.2	72.6	84.9	87.7	91.1	89.4
SEM	5.4	5.6	1.8	2.3	2.0	2.5	3.3	3.3	3.1	3.0
% short										
Mean	28.9	39.8	51.3	54.8	53.0	30.6	46.0	48.1	49.4	52.3
SEM	5.7	2.6	0.7	2.2	2.4	2.7	2.9	1.5	2.3	1.5

within analyses of variance on overall accuracy scores at each delay revealed significant effects of training block (5-s delay, $F(3, 33) = 36.8$, $p < .001$; 10-s delay, $F(3, 33) = 11.9$, $p < .001$) but no other significant effects.

Variable delay and ITI testing. Figure 5 shows the mean percentage of short choices for the two groups during the delay and ITI manipulation tests. Table 4 shows the mean and standard error scores of the percentage of correct choices and the percentage of short choices for each group. The values for $\log d$ and $\log b$ are not reported or analyzed because there were a large number of instances in which these measures could not be calculated. The responses of Group V-ITI to both the delay manipulations and to the ITI manipulations were similar to those of Group C-ITI.

For the delay tests, a between-within ANOVA on the percentage of correct responses showed a significant main effect of delay, $F(4, 44) = 14.07$, $p < .001$, but not of group, and no significant interaction between delay and group (both $ps > .1$). A similar ANOVA on the percentage of short choices also revealed a significant main effect of delay, $F(4, 44) =$

38.32, $p < .01$ but not of group, and no significant interaction between delay and group (both $ps > .1$).

For the ITI tests, a between-within ANOVA on the percentage of correct choices revealed a significant main effect of ITI, $F(4, 44) = 16.70$, $p < .001$, but not of group, and no significant interaction between ITI and group (both $ps > .2$). The ANOVA on the percentage of short choices also revealed a significant effect of ITI, $F(4, 44) = 26.11$, $p < .001$, but not of group and no significant interaction between ITI and group (both $ps > .2$).

Discussion

Constancy of the intertrial interval during training does not appear to be a critical factor in pigeons' memory for duration. It had only minor effects on the speed of acquisition and did not significantly alter the birds' tendency to make choose-short errors in response to increases in the delay interval. Thus, only very weak support was found for the first possibility considered: namely, that variability in background time would interfere with the formation of a stable reference memory. The second possibility considered received no support. That is, no evidence was found to suggest that the birds experiencing variable background times during training ceased to attend to background times and instead developed control by the absolute sample durations: these birds responded to subsequent manipulations of the delay or the ITI similarly to birds trained with constant background times.

The results of the ITI manipulation tests are perhaps the most interesting. During these tests, birds that had received extensive prior training with each of the test ITIs responded in a similar fashion to birds that had previously been exposed only to the middle ITI value. Specifically, the tendency to make short choices increased as a function of ITI in both groups of birds, and the magnitude of this change was comparable across the two groups. Thus, the duration of the ITI preceding a sample stimulus apparently continued to modulate the remembered duration of the sample, even after training with a variable ITI. This finding led us to consider the third possibility outlined in the introduction, namely, that the pigeons' reference memory was based on the average of the temporal backgrounds experienced during training.

Experiment 2b

The results of Experiment 2a indicated that constancy of the ITI during training is not a critical factor in the emergence of systematic errors in response to delay manipulations. Moreover, the systematic errors that emerged during ITI manipulation tests suggested that when temporal parameters vary during training, the average of the range of values may determine the reference memory. This conclusion, is however, weakened

by the fact that accuracy was not assessed separately at each ITI value during training. It is possible, for example, that dominance by the middle ITI values developed only after the birds had been exposed to the delay tests, during which the middle ITI was constantly presented. The present experiment extends Experiment 2a in two ways. First, the temporal parameter varied or held constant during training was the delay interval rather than the ITI. Second, accuracy at each training delay was recorded separately for short-sample and long-sample trials during acquisition.

Method

Subjects. The subjects were eight adult experimentally naive White King pigeons. They were housed and maintained as described in the previous experiments. Four randomly selected birds were assigned to a constant delay group (Group C-Del); the remaining birds were assigned to a variable delay group (Group V-Del).

Procedure. Preliminary training. All birds received a few sessions of magazine training and then autoshaping, as described in the previous experiments.

Baseline training. The basic DSMTS procedure described in the General Methods section was used. The samples were 2-s and 8-s food presentations. The ITI remained constant at 45 s for all birds. For birds in Group C-Del, the delay was 5 s on every training trial. For birds in Group V-Del, delays of 2, 4, 6, and 8 s occurred equally often, in a randomly determined order within each training session. All birds received 40 sessions of baseline training.

Delay and ITI testing. In this phase, all birds had sessions in which the delay was manipulated as well as sessions in which the ITI was manipulated. These delay and ITI test sessions alternated with each other and were separated by a baseline session. The procedure during baseline sessions was identical to that used during the training phase (i.e., the delay was constant at 5 s for Group C-Del, but varied from 2 s to 8 s for Group V-Del; the ITI was 45 s on every trial for both groups). During delay and ITI tests, the two groups were exposed to identical procedures. For the delay test sessions, 24 of the 48 trials contained a 5-s delay, 12 contained a 0-s delay, and 12 contained a 20-s delay; the ITI was 45 s on every trial. During ITI test sessions, the delay was 5 s on every trial, but the ITI varied: 24 trials were preceded by the 45-s ITI, 12 by a 5-s ITI, and 12 by an 85-s ITI. All birds received a total of 20 delay test sessions and 20 ITI test sessions, alternating with 40 baseline sessions.

Results

Baseline training. Acquisition of the task was not significantly affected by variation in the delay during training. During the four successive 10-session blocks of training Group C-Del produced a mean of 60%, 72%,

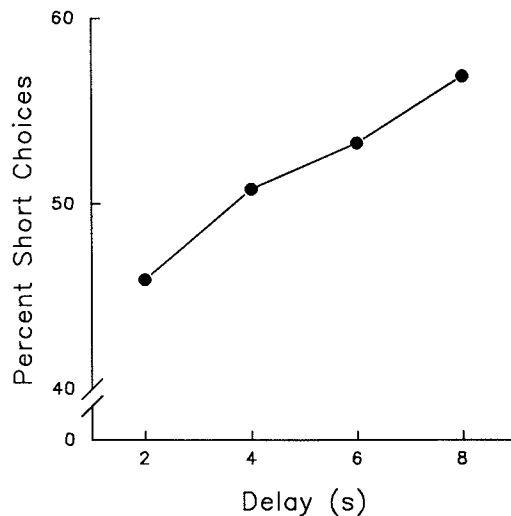


FIG. 6. Mean percentage of short choices as a function of delay during training with variable delays by Group V-Del in Experiment 2b.

83%, and 84% correct choices, and Group V-Del produced a mean of 52%, 69%, 79%, and 82% correct choices. A between-within ANOVA revealed a significant effect of blocks, $F(3, 18) = 37.38, p < .001$, but not of groups, $F(1, 6) = 0.80$, and no significant interaction between blocks and groups, $F(3, 18) = 0.47$.

TABLE 5

Measures of Accuracy (Percent Correct and Log d) and Bias (Percent Short and Log b) for Group V-Del at Each of the Four Training Delays, Averaged over the 40 Training Sessions

Delay	2	4	6	8
% correct				
Mean	73.9	71.5	71.0	66.1
SEM	3.1	3.0	2.5	3.0
Log d				
Mean	.467	.408	.399	.303
SEM	.067	.063	.056	.061
% short				
Mean	45.9	50.8	53.3	56.9
SEM	2.1	1.6	1.0	1.5
Log b				
Mean	-.093	.016	.075	.142
SEM	.045	.036	.027	.035

Figure 6 shows the mean percentage of short choices made by birds in Group V-Del at each of the four training delays, averaged over the 40 training sessions. Table 5 shows the results of the four measures (percent correct, log d , percent short, and log b). Even though each delay was represented equally often during each training session and even though the range of delays was quite small, choice tendencies varied systematically across these delays.

Within-subject ANOVAs on the data shown in Table 5 revealed a significant effect of delay for each of the four measures (percent correct: $F(3, 9) = 29.25, p < .001$; log d : $F(3, 9) = 22.93, p < .001$, percent short: $F(3, 9) = 10.13, p < .01$; log b : $F(3, 9) = 8.30, p < .01$). Inspection of Fig. 6 suggests that the 4-s delay generated the least bias; the mean percentage of short choices at this delay was very close to 50. To support this observation, a regression function relating delay to short choices was calculated for each bird and the point of bisection (delay associated with 50% short choices) was estimated. The points of bisection for Birds 1 to 4 were at delays of 4.32, 3.47, 3.27, and 5.42 s. Although there is considerable variability among the four birds, it is interesting to note that the average of these values is 4.12 s, which is closer to the geometric (4.43 s) and harmonic (3.84 s) means of the four training delays than to their arithmetic mean (5.0 s). A similar regression analysis on the log b scores generated very similar bisection points: 4.48, 3.44, 3.43, and 5.52 s. Again the average of these values (4.22 s) is closer to the geometric mean than to the arithmetic mean.

Delay and ITI tests. Figure 7 shows the mean percentage of short choices made during delay tests (left panel) and ITI tests (right panel) by the two groups. Mean scores for the four measures (percent correct, log d , percent short, and log b) are shown in Table 6. During delay tests, both groups showed a large choose-short effect at the 20-s delay and weak choose-long tendency at the 0-s delay. During ITI tests, both groups showed a strong choose-long tendency with a 5-s ITI and little bias at the other ITIs.

For the delay test data, between-within ANOVAs revealed a significant effect of delay for each of the four measures (percent correct: $F(2, 12) = 81.17, p < .001$; log d : $F(2, 12) = 49.97, p < .001$; percent short: $F(2, 12) = 47.80, p < .001$; log b : $F(2, 12) = 49.97, p < .001$), but none of the measures showed a significant effect of groups (all $ps > .2$), nor a significant interaction between delay and group (all $ps > .2$). Thus, accuracy declined and the tendency to choose short increased as a function of delay, but these effects were very similar across the two groups.

Similar ANOVAs on the ITI test data revealed a significant main effect of ITI for each of the four measures shown in Table 6 (percent correct, $F(2, 12) = 36.21, p < .001$; log d : $F(2, 12) = 20.59, p < .001$; percent short: $F(2, 12) = 34.77, p < .001$; log b : $F(2, 12) = 20.15, p < .001$),

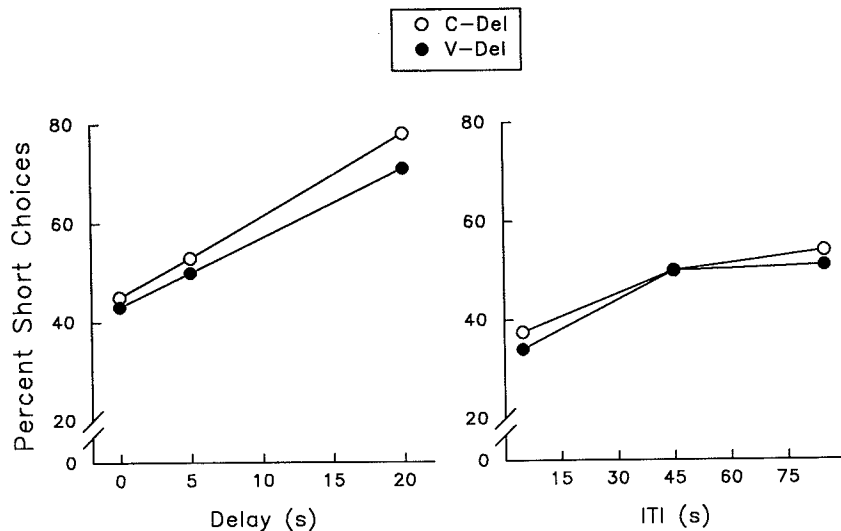


FIG. 7. Mean percentage of short choices during the delay manipulation tests (left panel) and ITI manipulation tests (right panel) for birds trained with a constant delay (C-Del) or a variable delay (V-Del) in Experiment 2b.

but again none of the measures showed a significant effect of group (all $ps > .2$), nor a significant interaction between ITI and groups (all $ps > .2$). Thus, the two groups responded in a similar fashion to manipulation of the ITI.

Discussion

Variability in the training delay had little effect on acquisition or on performance during delay or ITI manipulation tests. The training data for birds in Group V-Del suggests a reason for this. Despite the fact that each delay occurred equally often during training and despite the use of a correction procedure, opposite response tendencies emerged at the two extreme training delays; that is, the birds were more likely to make long choices at the shortest delay and short choices at the longest delay. Moreover, on average, the pigeons appeared to treat a delay of 4.12 s (or 4.22 s according to the signal detection analysis of bias) as the standard delay, which is reasonably close to the geometric mean (4.43 s) of the delays presented. Thus, the reference memory appeared to be based on some average of the training delays, perhaps on their geometric mean.

GENERAL DISCUSSION

The studies reported here explored the role of temporal context in pigeons' memory for event duration. In Experiment 1 we tested a pre-

TABLE 6

Measures of Accuracy (Percent Correct and Log d) and Bias (Percent Short and Log b) for Groups C-Del and V-Del during Variable Delay and Variable ITI Tests in Experiment 2b

Delay	Variable Delay Tests					
	C-Del			V-Del		
	0	5	20	0	5	20
% correct						
Mean	81.6	82.1	52.0	80.4	82.3	57.8
SEM	1.2	1.5	2.1	3.2	4.0	2.6
Log d						
Mean	.689	.678	.047	.684	.705	.185
SEM	.048	.045	.048	.079	.110	.069
% short						
Mean	44.9	52.9	77.5	43.0	50.1	71.0
SEM	2.7	1.6	2.6	4.4	1.8	5.0
Log b						
Mean	-.174	.082	.551	-.200	.010	.434
SEM	.085	.050	.068	.120	.065	.120
ITI	Variable ITI Tests					
	C-Del			V-Del		
	5	45	85	5	45	85
% correct						
Mean	73.8	87.0	86.4	70.0	85.5	84.9
SEM	3.2	2.7	2.2	1.4	3.6	4.5
Log d						
Mean	.519	.857	.845	.478	.825	.830
SEM	.092	.107	.096	.089	.147	.182
% short						
Mean	37.5	50.2	53.9	34.5	50.0	51.1
SEM	2.0	0.9	0.9	4.2	1.0	1.0
Log b						
Mean	-.316	-.006	.155	-.397	-.030	.068
SEM	.054	.033	.033	.142	.048	.053

diction, derived from the relative duration hypothesis (Spetch & Rusak, 1989), that the absolute length of one element of the background time surrounding the sample event (ITI and delay) would modulate the effect of manipulating the length of the other element. The underlying premise is that the two elements are summed to form a temporal framework against

which sample durations are judged. The impact of manipulating one element should, therefore, depend on the relative contribution of that element to the sum.

The results of Experiments 1a and 1b are consistent with this prediction. The impact of manipulating the delay interval in the range of 0 to 20 s depended on the absolute length of the concurrent ITI. When ITIs were short (and the delay interval therefore contributed proportionally more to the background), lengthening the delay produced stronger choose-short effects than when the ITI was long. Taking the extreme example of the change from 0-s to 20-s delay (Experiment 1b), the mean percentages of short choices increased from 41.7 to 91.1 (118%) in the 10-s ITI condition and from 46.3 to 68.5 (48%) in the 60-s ITI condition.

Another way of viewing these results is to note that in comparing the 20-s delay trials to the 0-s delay trials, the total background time increased from 10 s to 30 s (200%) in the case of a 10-s ITI, whereas it increased from 60 s to 80 s (33%) in the case of the 60-s ITI. The failure to match the increase in total background with a proportional increase in short choices for the 10-s ITI condition may reflect in part a ceiling effect (a 200% increase is not possible). But, the fact that the change from 0-s to 20-s delays had a larger impact on short choices (48% increase) than predicted by the increase in background duration (33%) implicates another process.

The lack of perfect proportionality in these increases probably reflects the fact that the two intervals are not summed in simple algebraic fashion. A unit of delay time contributes more to the total perceived background time than an equal unit of ITI time, probably because the delay interval is closer to the choice time. Assuming that these intervals are assessed retrospectively (Spetch & Sinha, 1989; Wilkie & Willson, 1990), perceived durations are probably affected by foreshortening, which strengthens the impact of more recent events (Spetch & Rusak, 1992; Spetch & Wilkie, 1983; Staddon, 1984). This conclusion is consistent with our earlier finding that manipulations of the delay interval generally produce larger systematic errors than manipulations of the ITI, even when the latter are varied over a much wider range of intervals (Spetch & Rusak, 1989).

The qualitative support provided from Experiments 1a and 1b for the prediction derived by the relative duration hypothesis strengthens our confidence that this notion has merit. However, the inability of this hypothesis to account for all details of the results we have collected to date suggests that other factors such as foreshortening are also operating. In addition, our previous finding (Spetch & Rusak, 1989) that ITI manipulations do not have maximal effects in the context of a 0-s delay (i.e., when ITI is the exclusive contributor to the background time) suggests some control by absolute sample duration. One possibility (Spetch & Rusak, in press) is that control by the absolute duration of the sample

(or by stimuli or behaviors correlated with sample duration, see Killeen & Fetterman, 1988) may be stronger immediately following the sample than when the sample is temporally more remote. That is, when the delay is very short, control by the perceived relative duration of the sample may be less powerful than control by the absolute sample duration. However, as the delay is lengthened, the relationship between the sample duration and the background duration may become more salient and exert greater control over choice performance.

Despite the necessity of invoking other processes to account for all details of the results collected to date, the accumulation of evidence in support of the hypothesis that pigeons evaluate sample durations relative to a temporal background indicates that this relative duration hypothesis merits serious consideration and should be tested further. It is also worth noting that this hypothesis is compatible with recent evidence that pigeons are capable of making discriminations on the basis of the ratio of two stimulus durations (Fetterman, Dreyfus, & Stubbs, 1989) and with evidence that the value or behavioral effects of a temporal signal for a significant event (e.g., food) depend not only on its absolute temporal proximity to the event, but also on its relative proximity to the event (e.g., Balsam, 1984; Fantino, 1981). On a more general level, our evidence for temporal context effects in the judgment of event duration is also congruent with evidence of pervasive context effects on magnitude estimations made by people (Foley, Cross, & O'Reilly, 1990). Reliance on relative rather than absolute judgments of magnitude may be common to many situations and species.

Although Experiment 1 supported the predictions of the relative duration hypothesis, it is possible to interpret the effects within other theoretical frameworks. For example, the results are quite consistent with Killeen and Fetterman's (1988) behavioral theory of timing. According to this theory, time discrimination is based on transitions between different adjunctive behavioral states, which are precipitated by pulses from a pacemaker. Because the rate of the pacemaker is assumed to vary directly with the rate of reinforcement in the experiment context, the pacemaker should run faster in the 10-s ITI condition than in the 60-s ITI condition. Killeen and Fetterman suggested that the choose-short effect reflects biased guessing, which occurs when animals have exited from the signal state to a no-food state during the delay. It follows that if pacemaker speed is faster in the 10-s ITI condition, transitions from the signal state to the no-food state should occur earlier. Consequently, the magnitude of the choose-short effect should be larger at a given delay because more of the responses will be occurring while in a no-food state and hence will be biased toward short. This behavioral theory of timing may, however, require modification in order to account for other features of pigeons' memory for temporal events, such as the occurrence of a choose-long

effect in response to decreases in the delay (e.g., Spetch, 1987) and the occurrence of systematic errors in response to local (i.e., trial-by-trial) variations in the ITI (Spetch & Rusak, 1989).

Experiment 2 concerned the role of temporal context in the formation of reference memories for stimulus duration. In formulating the relative duration hypothesis, we assumed that pigeons form a reference memory for short and long stimulus durations based on the ratios of stimulus to background duration experienced during short-sample and long-sample training trials. We expected that varying these ratios during training might prevent or at least slow the acquisition of stable reference values to which test trials would be compared. Moreover, we expected that pigeons with variable stimulus:background ratios during training might fail to show systematic errors in response to varying the delay or the ITI during testing. In fact, we found that pigeons which experienced varying ITIs or delays during training showed systematic errors that were indistinguishable from those shown by pigeons trained with a standard constant ITI or delay. These results imply that pigeons had acquired stable reference memories by the time of testing.

In addition, the length of time needed to achieve the criterion for discrimination between the short and long sample events differed only slightly between pigeons trained with constant or variable values for the ITI or delay. An examination of the errors generated by the pigeons exposed to variable delays during acquisition indicated that, even during the training trials, pigeons were showing systematic errors in response to the extreme delay values presented (Experiment 2b). Within the framework of the relative duration hypothesis, these results imply that pigeons rapidly extract reference values from the range of stimulus:background ratios they experience, and they evaluate the stimulus on each trial with respect to these extracted reference values. The delay value that generated the least error on average was nearer to the geometric and harmonic means of the training delays presented than to their arithmetic mean. Although between-subject variability precluded any strong conclusions regarding this finding, it is interesting to note the consistency of this result with evidence from a number of psychophysical studies of timing in animals that the subjective midpoint between long and short temporal values falls approximately at the geometric mean (e.g., Church & Deluty, 1977; Platt & Davis, 1983; Stubbs, 1968, 1976).

In Experiment 2b, the four different delays were equally spaced on a linear scale and occurred with equal frequency. It remains to be determined to what extent the extracted reference values would be influenced by variations in the distribution of values or by presentation of unequal numbers of training trials for different values. Varying the range, distribution, and frequency of delay or ITI values during training should provide

further insight into how temporal reference values are acquired and subsequently used in the evaluations of stimulus durations.

In summary, the present results suggest a rather complex role for temporal context in memory for event duration. First, as demonstrated in Experiment 2 and in earlier studies (e.g., Spetch & Rusak, 1989), local variations in either the time before the event (ITI) or time after the event (delay) systematically alter the tendency to respond with a bias for short choices, suggesting that the perceived or remembered duration of the event is altered by these changes in the temporal context. Second, Experiment 1 showed that ITI duration modulates the effects of delay on choice bias. This modulation occurs even though the pigeons have received training sessions with the ITI and appeared to be fully adjusted to the ITI prior to the delay tests. This result confirmed a prediction of the relative-duration hypothesis. Finally, the results of Experiment 2 suggest that, although working memory is markedly altered by local variations in temporal context, both the formation and characteristics of reference memory are impervious to variability in the temporal context during training.

REFERENCES

- Balsam, P. (1984). Relative time in trace conditioning. *Annals of the New York Academy of Sciences*, **423**, 211-227.
- Church, R. M. (1980). Short-term memory for time intervals. *Learning and Motivation*, **11**, 208-219.
- Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes*, **3**, 216-228.
- Davison, M. C., & Tustin, R. D. (1978). *Journal of the Experimental Analysis of Behavior*, **29**, 311-336.
- Fantino, E. (1981). Contiguity, response strength, and the delay-reduction hypothesis. In P. Harzem & M. D. Zeiler (Eds.), *Advances in analysis of behavior: Vol. 2. Predictability, correlation, and contiguity* (pp. 168-201). Chichester, England: Wiley.
- 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.
- Fetterman, J. G., Dreyfus, L., & Stubbs, D. (1989). Discrimination of duration ratios. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 253-263.
- Foley, H. J., Cross, D. V., & O'Reilly, J. A. (1990). Pervasiveness and magnitude of context effects: Evidence for the relativity of absolute magnitude estimation. *Perception and Psychophysics*, **48**, 551-558.
- Gallistel, C. R. (1990) *The organization of learning*. Cambridge, MA: MIT.
- Gibbon, J., & Church, R. M. (1984). Source 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.
- Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences*, **423**, 322-334.
- Holder, M. D., & Roberts, S. (1985). Comparison of timing and classical conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, **11**, 172-193.

- Honig, W. K., & Spetch, M. L. (1988). Short-term memory for rate of alternation in the pigeon. *Bulletin of the Psychonomic Society*, **26**, 152-154.
- Jenkins, H. M. (1984). Time and memory. *Annals of the New York Academy of Sciences*, **423**, 322-334.
- Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological Review*, **95**, 274-285.
- 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.
- Lucas, G. A., & Wasserman, E. A. (1982). US duration and local trial spacing affects autoshaped responding. *Animal Learning & Behavior*, **10**, 490-498.
- Platt, J. R., & Davis, E. R. (1983). Bisection of temporal intervals by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **9**, 160-170.
- Roberts, S., & Holder, M. D. (1985). Effects of classical conditioning on an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes*, **11**, 194-214.
- Roberts, W. A., Cheng, K., & Cohen, J. S. (1989). Timing light and tone signals in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 23-35.
- Roberts, W. A., & Kraemer, P. J. (1982). Some observations of the effects of intertrial interval and delay on delayed matching to sample in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **8**, 342-353.
- 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., & Rusak, B. (1992). Time present and time past. In W. K. Honig and J. G. Fetterman (Eds.) *Cognitive aspects of stimulus control* (pp. 47-67). Hillsdale, NJ: Lawrence Erlbaum.
- Spetch, M. L., & Sinha, S. S. (1989). Proactive effects in pigeons' memory for event duration: Evidence for analogical retention. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 347-357.
- 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 durations. *Journal of Experimental Psychology: Animal Behavior Processes*, **9**, 14-30.
- Staddon, J. E. R. (1984). Time and memory. *Annals of the New York Academy of Sciences*, **423**, 322-334.
- Stubbs, A. (1968). The discrimination of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior*, **11**, 223-238.
- Stubbs, D. A. (1976). Scaling of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior*, **26**, 15-25.
- Wilkie, D. M., & Willson, R. J. (1990). Discriminal distance analysis supports the hypothesis that pigeons retrospectively encode event duration. *Animal Learning & Behavior*, **18**, 124-132.
- Wilkie, D. M., Symons, L. A., & Tees, R. C. (1988). Effects of intertrial reinforcers on rats' timing behavior. *Behavioural Processes*, **17**, 229-238.

Received December 11, 1990

Revised June 15, 1991