The role of trial tracking in rats' working memory

J. S. COHEN and M. NJEGOVAN
University of Windsor, Windsor, Ontario, Canada

The experiments reported in the present study tested whether decreasing intertrial intervals (ITIs) intensifies the disruptive effects of increasing retention intervals (RIs) in a delayed conditional discrimination by decreasing the animal's trial tracking accuracy (Cohen & Armstrong, 1996; Cohen & Roberts, 1996). Rats responded on a fixed ratio (FR) 1 or fixed interval (FI) 10-sec reinforcement schedule at a second light or tone stimulus, S2, when the first light or tone stimulus, S1, had signaled an FI 10-sec or FR 1 schedule, respectively. RIs between S1 and S2 were increased from 3 to 24 sec and never exceeded ITIs that were reduced from 24 to 6 sec. For some rats, the trials were separated from each other by extending the lever at S1 and retracting it at the end of S2 (ITI lever-retracted group). For other, control rats, the lever remained extended throughout the session (lever-extended group, Experiment 1) or was extended and retracted with the onset and offset of each stimulus (RI/ITI lever-retracted group, Experiment 2). The rats under all trial conditions learned to delay leverpressing on the FI 10-sec schedule. Latency to begin leverpressing on the FI 10-sec schedule declined as RIs were increased, but this effect was attenuated in the ITI lever-retracted groups in both experiments, as would be predicted by the *trial tracking* hypothesis. Decreasing ITIs from 24 to 6 sec intensified the disruptive effects of increasing RIs from 3 to 6 sec in the RI/ITI lever-retracted group (Experiment 2), as would be predicted by the trial tracking hypothesis.

Increasing retention intervals (RIs) and reducing intertrial intervals (ITIs) typically reduce animals' accuracy in delayed two-choice and successive matching-to-sample tasks. These effects have been observed in pigeons (Grant, 1975; Maki, Moe, & Bierley, 1977; Nelson & Wasserman, 1978; Roberts, 1980), monkeys (Jarrard & Moise, 1971), and a dolphin (Herman, 1975). These effects are particularly reliable in free-operant preparations containing only two sample stimuli and were originally accounted for by a decline in the working memory for sample events as RIs become longer and by a buildup of proactive interference (PI) between trials as ITIs become shorter (Grant, 1975; Maki et al., 1977; Nelson & Wasserman, 1978). Direct evidence of intertrial PI is indicated by the *intertrial disagreement* effect, as can be seen in the poorer matching accuracy in cases in which the incorrect choice on one trial (*n*) was the correct rather than the incorrect choice on the previous trial (Maki et al., 1977). Although some studies found that massing trials increased such intertrial disagreement effects in two-choice delayed color (Grant, 1975) or line pattern matching (Roberts, 1980), others failed to replicate these effects (Edhouse & White, 1988; Roitblat & Scopatz, 1983). Furthermore, Roberts (1980)

found that massing trials continued to reduce pigeons' delayed matching accuracy, even on sessions in which all the samples were the same, thus removing any possibility of an intertrial disagreement effect.

Failure to find reliable direct evidence that massing trials increases intertrial PI has prompted alternative explanations of the effects of varying both RIs and ITIs in delayed conditional discriminations. One idea is that massing trials causes a general decrement in attention to or processing of sample stimuli or events, independent of any specific intertrial PI (Edhouse & White, 1988; Roitblat & Scopatz, 1983). Another notion, the *task instruction confusion* hypothesis (Zentall, 1997), suggests that the reduced matching accuracy that accompanies increased RIs and decreased ITIs may reflect, in part, some decrement in the animal's reference memory about the instructional rules of the task. Zentall notes that animals require considerable training with brief RIs and relatively long ITIs to initially acquire difficult conditional discriminations. Therefore, the basic rules of the task in their reference memory may be specific to these training RI and ITI levels, so that even slight variations in these intervals during initial testing may produce considerable generalization decrements. A third related theoretical approach, based on Gibbon's (1977) scalar expectancy theory, assumes that an animal's expectancy of being reinforced for correctly responding at the end of a trial is inversely related to the ratio of its RI to the ITIs within a session (Roberts & Kraemer, 1982). Thus, the smaller the ratio, the greater will be an animal's expectancy of reinforcement at the end of the trial for responding on the basis of the conditional rule. On the other hand, increasing this ratio will promote reinforcement expectancies based

more on responding over similar delays within the session. Possible evidence of this expectancy model has been seen in pigeons' equivalent color matching accuracy at the same RI:ITI ratios as RIs were varied from 0.5 to 4 sec within sessions while ITIs were varied from 4 to 32 sec over sessions (Roberts & Kraemer, 1982).

A more recent account of the disruptive effects of massing trials, and a possible explanation for the RI:ITI effect, is the *trial tracking confusion* hypothesis (Cohen & Armstrong, 1996). According to this notion, an animal may become more confused about whether it is still within a trial or between trials during an interval between stimulus events as RIs and ITIs become more similar. As it becomes more confused about where it is between stimuli, it may stop processing information about the sample stimulus and become more likely to make incorrect matching responses. In other words, either increasing RIs to ITI levels or decreasing ITIs to RI levels decreases the animal's ability to accurately track trials within a session.

Evidence in support of the trial tracking hypothesis comes from recent studies on the effects of varying RIs and ITIs on rats' performance on successive delayed modality (onset of chamber lights or a tone) matching tasks (Cohen & Armstrong, 1996; Cohen & Roberts, 1996). One study (Cohen & Roberts, 1996) found that decreasing ITIs from 24 to 6 sec disrupted matching accuracy for both short 1-sec and longer 5-sec RIs only when reinforcement occurred for responding to each sample, S1, and to its matching test stimulus, S2, but not when reinforcement was only contingent on responding to matching S2s. When rats could always obtain reinforcement for responding at S1 as well as for responding at S2 when it matched the S1 (Cohen & Armstrong, 1996), the ITI effect failed to occur with 2-sec S1s and 10-sec S2s but emerged when S1's duration was later increased to 10 sec. These results suggest that trial tracking was more difficult with short ITIs when reinforcement was available at S1 and S2 of the same or similar durations than when either reinforcement was only available at the end of a trial (Cohen & Roberts, 1996) or when a short stimulus always began a trial and a longer one always ended it (Cohen & Armstrong, 1996). The independence between ITI and intertrial disagreement effects was also observed for matching accuracy to a 10-sec tone S1 (Cohen & Armstrong, 1996).

In the present study, the trial tracking hypothesis was tested by investigating the effects of allowing rats to differentiate RIs from ITIs by means other than the duration of the intervals. According to the trial tracking hypothesis, performance should be less disrupted by increases in RIs and decreases in ITIs when these intervals can be distinguished from each other by means other than differences in duration. Previous attempts to differentiate ITIs from RIs by illuminating the operant chamber during ITIs merely resulted in greater declines in pigeons' color matching accuracy, however (Edhouse & White, 1988, Experiment 2; Santi, 1984). Perhaps ITI illumination perceptually interferes with animals' processing of the visual attributes of samples. Therefore, ITIs were differentiated from RIs by the retraction of the lever between trials in the present study. The present study consisted of two experiments. In both experiments, the rats in the experimental group experienced the extension of the lever at the onset on the sample stimulus, S1, and its retraction at the offset of the test stimulus, S2 (the ITI lever-retracted group). In Experiment 1, the rats in the control group were exposed to the lever's being continuously extended throughout the session (the lever-extended group), whereas in Experiment 2, the control rats experienced the lever's being extended at the onset and retracted at the offset of S1 and S2 (the RI/ITI lever-retracted group). Rats in an ITI lever-retracted group should be able to distinguish trials from intertrial intervals by noticing the presence of the lever during RIs and its absence during ITIs. The trial tracking hypothesis predicts that rats in an ITI lever-retracted group should be less disrupted by increases in RI or decreases in ITI than rats in either of the other two control groups, which cannot use the presence and absence of the lever to determine whether they are within or between trials.

EXPERIMENT 1

Our original intention, to investigate the effects of varying lever presentation on successive delayed modality matching performance (Cohen & Armstrong, 1996; Cohen & Roberts, 1996), proved to be unfeasible, as too few rats could acquire this task when the lever was retracted during ITIs in a pilot experiment. Therefore, we devised the following, easier conditional task. The rats were also exposed to pairs of stimuli presented successively, but each stimulus indicated that the rat could begin leverpressing under one of two possible reinforcement schedules, fixed-ratio (FR) 1 or fixed-interval (FI) 10 sec. The schedule at S2 was conditional on that at S1. FI 10 sec was only in effect at S2 when FR 1 had been in effect at S1, and FR 1 was only in effect at S2 when FI 10 sec had been in effect at S1. Thus, this task can be considered to be a type of successive delayed conditional mismatching task. Following acquisition with short 3-sec RIs and long 24-sec ITIs, the rats were tested with RIs being lengthened to 24 sec while ITIs were being shortened to 6 sec over sessions. These variations ensured that the rats would be exposed to three RI:ITI ratios, each derived from two different pairs of intervals: 0.25 from 6-sec RIs:24-sec ITIs and 3-sec RIs:12-sec ITIs; 0.5 from 12-sec RIs:24-sec ITIs and 6-sec RIs:12-sec ITIs; 1 from 24-sec RIs: 24-sec ITIs and 12-sec RIs:12-sec ITIs.

A between- and a within-subjects independent variable was used to test the trial tracking hypothesis. Lever presentation was varied between subjects so that the experimental rats were exposed to the lever's being extended at S1 onset and retracted at S2 offset (ITI lever-retracted group), whereas control rats were exposed to a continuously extended lever (lever-extended group). Modality of the S1 and S2 (onset of chamber lights or a tone) was

also varied for all the rats, so that each stimulus was the same on every trial in some sessions (same modality condition) and consistently alternated between S1 and S2 on others (consistent alternating modality condition).

All the rats were expected to learn to respond quickly on the FR 1 schedule but delay responding on the FI 10 sec at S2 during initial training. According to the trial tracking hypothesis, these response differences at S2 were expected to be better maintained over variations in RI and ITI in the ITI lever-retracted group than in the lever-extended group, for reasons already explained. Superior performance was also expected to occur when S1 and S2 were consistently signaled by a different, rather than the same, modality within a session. Furthermore, equivalent performances at the same RI:ITI ratios, regardless of the actual durations of each interval, were expected to more likely occur in the lever-extended group, especially when S1 and S2 were the same modality. Finally, any intertrial PI, as shown by a poorer performance on Trial n S2 when its S1 schedule differed from that on Trial $n-1$ S2 than when it was the same (an intertrial disagreement effect), was expected to be independent of variations in RI and ITI.

Method

Subjects. Sixteen experimentally naive, Wistar male albino rats from our breeding colony were used. Their ages and weights ranged between 90 and 150 days and 340 and 360 g, respectively, at the beginning of the experiment. The rats were housed in large group cages, 4 per cage, between experimental sessions. Each rat received a 30-min free access to water in its individual holding cage after each session. Dry food (Purina Lab Chow) was always available in the group cages. A 12-h dark:light cycle was in effect in the colony room, and the rats were always run at the beginning of the dark period. The experiment was interrupted every 2 weeks to allow rats a 48-h free access to water in their home cages.

Apparatus. Four standard operant chambers, each enclosed in a light- and sound-attenuating isolation chamber (Tech-Serv), were used. Each operant chamber contained a 0.05-ml water dipper and a retractable lever on one wall and five 24-V 6-W incandescent lights and a 2.8-kHz tone generator (Malloy Sonalert) in the ceiling. The tone generator was dampened to emit a 75-dB tone. Each operant chamber's ventilation fan produced 55-dB white noise. Apple II+ computers, in an adjoining room, programmed the presentation of each chamber's lights, tone, lever, and water dipper and monitored the rat's leverpresses. A reinforcement consisted of a 0.05-ml droplet of artificially sweetened water (0.10% sodium cyclamate, w/v).

Procedure. The experiment consisted of a preexperimental and the experimental phase. The first phase was designed to acquaint the rats with the basic procedures of the task, their respective lever presentation conditions, and variations in RI and ITI similar to those to be used in the actual experiment in Phase 2. The rats were trained on a simpler version of the task, consisting of single alternations of FR 1 and FI 10-sec schedules over stimuli. The rats were placed into either the experimental group, in which the lever was extended at S1 and retracted at S2, or the control group, in which the lever remained extended throughout each session. We took into account Zentall's (1997) task instruction confusion hypothesis and extensively trained and tested rats in this phase over 122 sessions, hoping to prevent generalization decrements that were due to changes in RI and ITI between training and testing conditions.

During the experimental phase, FR 1 and FI 10-sec reinforcement schedules at S1 occurred equally over trials within each session in a random order, with the restriction that the same schedule

at S1 was never repeated on more than 2 trials in a row. Whenever FR 1 occurred at S1, FI 10 sec occurred at S2, and whenever FI 10 sec occurred at S1, FR 1 occurred at S2. Each session was programmed to contain up to 96 S1–S2 trials. A rat had up to 20 sec to make a reinforced response to each stimulus. Failure to respond within this period to S1 terminated the trial, and the next programmed S1 occurred after a 4-sec delay. Failure to respond within this period to S2 merely terminated that stimulus and initiated the appropriate ITI.

This phase consisted of a training stage of 16 sessions, followed by a testing stage of 36 sessions. During the training stage, RIs between S1 and S2 were kept at 3 sec and ITIs between S2 and the next S1 were kept at 24 sec. The first 16 sessions of the test stage contained 24-sec ITIs while RIs were increased from 3 to 6 sec, then to 12 sec, and finally to 24 sec over four successive 4-session blocks. The next 12 test sessions contained 12-sec ITIs while RIs were increased from 3 to 6 sec, and finally to 12 sec over three successive 4-session blocks. The final 8 sessions contained 6-sec ITIs while RIs were increased from 3 to 6 sec over two successive 4-session blocks. Thus, RIs were never increased beyond the ITIs within any of the three respective blocks of test sessions. The two kinds of stimulus sequences (S, same modality at S1 and S2 on each trial; A, consistent alternating modalities between S1 and S2 on each trial) were varied over each block of 4 sessions within the training and testing stages in an ASSA and SAAS order. Within each block of sessions, one S session contained only light stimuli, and the other only tone stimuli, whereas one A session contained a light S1 and tone S2 on every trial, and the other contained the opposite sequence. The order of these S and A sessions was randomized over successive blocks.

Dependent measures. Each operant chamber's computer partitioned each stimulus into ten 1-sec bins over its first 10 sec and the 11th bin for the remaining period (maximum of 10 sec). It monitored the latency of each initial leverpress to S1 and S2 and recorded the frequency of these initial responses in each bin for each stimulus in a session. Although we calculated frequency distributions of these initial responses to each reinforcement schedule in both phases of the experiment, we describe only the procedures involved in calculating the dependent measure in the experimental phase. The frequency distributions for each reinforcement schedule were separately determined at S1 and S2. These distributions were summed over blocks of four training sessions or 192 trials per schedule at each stimulus and over the two test sessions or 96 trials per schedule within each RI/ITI block at each stimulus modality sequence condition. Only data from trials in which the rat obtained reinforcement for responding at each stimulus were used to construct these distributions in either phase.

The extent to which FR 1 and FI 10-sec distributions departed from each other was measured by the A index, a rating scale sensitivity measure from signal detection theory (McNicol, 1972). A represents the proportion of the area under the receiver operating curve (ROC) and ranges from .50 when the distributions of initial responses to FR 1 and to FI 10-sec schedules completely overlap to 1.00 when they never overlap. Procedures for calculating A scores were adapted from those used by Green, Terman, and Terman (1982). It should be noted that using initial leverpress latencies to derive A scores, rather than other measures of differential responding to each schedule at S2 (e.g., relative rates of leverpressing), avoids a potential confound between the discriminative control by the S1 schedule and the reinforcement control by the S2 schedule.

The A index, a distribution-free measure, is preferable to other signal detection measures of sensitivity (e.g., d' or log d') when distributions are not Gaussian. Visual inspection of each rat's response distributions convinced us that their responding to FR 1 and FI 10 sec were neither normally distributed nor of similar variability. The rats typically developed steep, skewed response latency distributions over the first few bins to FR 1 and broad, flat distributions to FI 10 sec (see Table 1A for a typical example).

An example illustrating the procedures for calculating A scores is shown in Table 1. This example shows a control rat's frequency

Table 1
Example of Calculating A Scores at S2 on a Conditional FR 1/FI 10-sec
Schedule Task at 3-sec and 6-sec RIs at 12-sec ITIs in
Phase 2 From Experiment 1 for Rat 2

A. Proportions of Number of Initial Responses Made Within Each 1-sec Bin

	1-sec Bins										
	1	2	3	4	5	6	7	8	9	10	11+
3-sec RIs											
FR 1	0	.50	.46	.04	0	0	0	0	0	0	0
FI 10 sec	.02	.04	.11	.08	.12	.08	.06	.13	.11	.08	.17
12-sec RIs											
FR 1	0	.45	.40	.09	.03	.01	0	0	0	0	.02
FI 10 sec	.02	.28	.27	.14	.08	.02	.02	.05	.04	.03	.05

B. Cumulative Proportions

	1-sec Bins										
	1	2	3	4	5	6	7	8	9	10	11+
3-sec RIs											
FR 1	0	.50	.96	1	1	1	1	1	1	1	1
FI 10-sec	.02	.06	.17	.25	.37	.45	.51	.64	.75	.83	1
12-sec RIs											
FR 1	0	.45	.85	.94	.97	.98	.98	.98	.98	.98	1
FI 10-sec	.02	.30	.57	.71	.79	.81	.83	.88	.92	.95	1

C. Algebraic Calculations of A scores (a = FR 1; b = FR 10 sec)

1. Basic Formula:
$$A = .5(a_1)(b_1)+a_1(b_2-b_1)+.5(a_2-a_1)(b_2-b_1)+ \\ \ldots +a_{10}(b_{11}-b_{10})+.5(a_{11}-a_{10})(b_{11}-b_{10})$$

2. Calculation of the A score at 3-sec RIs:
$$A = .5(0)(.02) +0(.04)+.5(.50)(.04)+ \ldots +1(.17)+.5(0)(.17) = .92$$

3. Calculation of the A score at 12-sec RIs:
$$A = .5(0)(.02)+0(.28)+.5(.45)(.28)+ \ldots +.95(.05)+.5(.02)(.05) = .65$$

distributions for each schedule at S2 of the conditional task during tests with 3-sec RIs and 6-sec RIs at 12-sec ITIs when stimuli were always either lights or tones. The frequency distributions to the FI 10-sec and FR 1 schedules at 3-sec RIs were each derived from 96 trials and at 6-sec RIs from 94 trials. As can be seen in Table 1, each distribution was transformed into proportions of initial presses in each temporal bin (A), and then into cumulative proportions (B), from which A scores were calculated by the algebraic formula (C).

Figure 1 graphically illustrates the ROCs from which each A score is derived. The cumulative proportions from each schedule (Table 1B) are plotted to produce two ROCs in a two-dimensional space described by scales of the FI 10-sec schedule along the abscissa and the FR 1 schedule along the ordinate. The area under each ROC as the proportion of the total area in this two-dimensional space defines the A score. Each A score is greater than the area below the dotted diagonal line that defines $A = .50$, and the area under the ROC for the 12-sec RI is less than that for the 3-sec RI.

Intertrial disagreement effects were possible in the conditional task. That is, intertrial PI would be shown by poorer performance on Trial n S2 when its S1 schedule differed from that on Trial $n-1$ S2 than when it was the same. Too few trials occurred within each RI/ITI combination to compute A scores at each Trial n S2 as a function of Trial $n-1$ S2 schedule. Therefore, reciprocals of the latencies to the first press on each Trial n S2 schedule following each Trial $n-1$ S2 schedule were calculated to reduce heterogeneity of variance between schedules and to generate more normal-like distributions within each schedule (Winer, 1971) for parametric statistical (analysis of variance [ANOVA]) analysis.

Results and Discussion

As was expected, all the rats made most of their first presses within 2 sec of S1's onset, regardless of the spe-

cific schedule in effect on the conditional task, to generate A scores close to .50. Therefore only A scores from latency distributions to each schedule at S2 are reported. One rat in the control lever-extended group failed to develop A scores above .50 at S2, and its data are disregarded
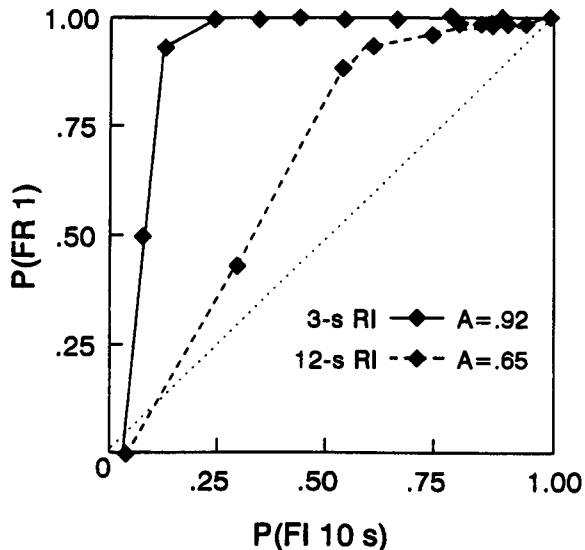
Figure 1. Example of the generation of A scores from cumulative proportions of initial response latencies in 1-sec bins (Table 1B) for each reinforcement schedule.

in the following analyses. All the other rats developed *A* scores above .80 during training in the conditional task. Therefore, 7 control rats in the lever-extended group were compared with the 8 experimental rats in the ITI lever-retracted group.

Figure 2 shows the effects of changes in RIs as a function of ITI (top panel) and the effects of changes in ITIs as a function of RI (bottom panel). As can be seen in this figure, *A* scores generally declined as RIs were lengthened (top panel) but not as ITIs were shortened (bottom
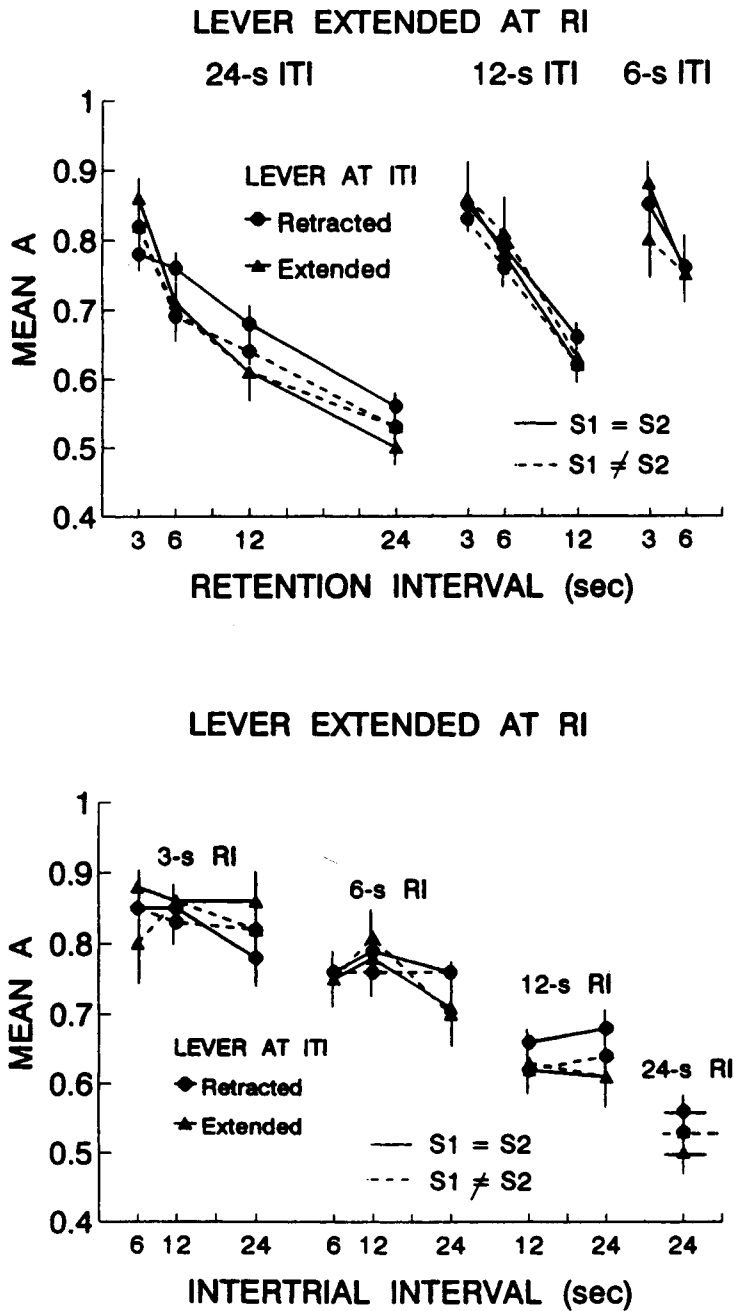

Figure 2. Mean *A* scores as a function of changes in retention interval (RI) within each intertrial interval (ITI) block of sessions (top panel) and as a function of changes in ITI at each RI (bottom panel) for each group and each stimulus modality presentation sequence in Experiment 1 (Phase 2). S1 = S2, same modality for each stimulus within a session; S1 ≠ S2, modality consistently alternated between stimuli. Vertical lines represent standard errors of means. The lever was extended at RI for all the animals, but at ITI it was retracted in the experimental group and extended in the control group.

panel). Also evident in the top panel was the fact that increasing RIs tended to disrupt performance slightly more in the lever-extended group than in the ITI lever-retracted group only when both stimuli were the same on each trial. The same RI:ITI ratios derived from both shorter and longer intervals did not appear to generate similar performance under any of the lever or stimulus modality conditions. Rather, lower *A* scores occurred at the same RI:ITI ratios when they were derived from the longer rather than from the shorter intervals.

The statistical reliability of these observations was determined by a series of mixed factorial ANOVAs. Due to logistical considerations, we did not use a completely crossed design—that is, we did not combine all RIs with all ITIs. Rather, RIs were only lengthened to equal the specific ITI within each specific series of test sessions. Hence, these data could not be analyzed with a single ANOVA but had to be partitioned into separate ANOVAs in the following manner. The data in the top panel were analyzed by two separate four-way ANOVAs: lever presentation group × stimulus modality sequence (same vs. alternating S1/S2 modality sequences) × RI (3 sec and 6 sec in the first ANOVA) or RI:ITI (.5, 1 in the second ANOVA) × ITI (6, 12, and 24 sec), with repeated measures on the last three factors. The first ANOVA assessed the effects of increasing RIs from 3 to 6 sec as ITIs were reduced from 24 to 12 and finally to 6 sec. The second ANOVA assessed the effects of increasing RIs from half to equal that of their respective ITI—that is, from 12- to 24-sec RIs at 24-sec ITIs, from 6- to 12-sec RIs at 12-sec ITIs, and finally from 3- to 6-sec RIs at 6-sec ITIs. The data in the bottom panel were also analyzed by two separate ANOVAs. The first four-way ANOVA of data in the top panel was also used to assess the effects of decreasing ITI s from 24 to 12 sec and then to 6 sec at 3-sec and at 6-sec RIs. The second analysis assessed the effects of decreasing ITIs from 24 to 12 sec at 12-sec RIs by a three-way ANOVA—group × stimulus modality × ITI. As RIs never exceeded ITIs, no analysis for decreasing ITIs at 24-sec RIs was possible. Newman–Keuls tests of individual comparisons were carried out when appropriate. The significance of effects from ANOVAs and individual comparisons was maintained at *p* < .05.

A main effect for RI [$F(1,13) = 21.51$] in the first four-way ANOVA confirmed that *A* scores declined as RIs were increased to 6 sec. Although an expected decline in *A* scores as ITIs were shortened did not materialize, a main effect for this factor [$F(2,26) = 5.47$] resulted from *A* scores pooled over 3- and 6-sec RIs being slightly lower at 24-sec ITIs (.77) than at 12- (.81) or 6-sec ITIs (.80) . Reducing ITIs at 12-sec RIs had no apparent or significant effects. The observation that increasing RIs at 24-sec ITIs disrupted performance more in the lever-extended than in the ITI lever-retracted group only when both stimuli were always the same on every trial was supported by a triple interaction in the first ANOVA among group, stimulus modality sequence, and RI [$F(1,13) = 5.03$] and a triple interaction in the second ANOVA among group,

stimulus modality sequence, and ITI [$F(2,26) = 4.03$]. This effect is in accordance with the trial tracking hypothesis. The prediction that consistent alternation of stimulus modality would also reduce any disruptive effects from changes in RI and ITI was not confirmed, however.

Rather than the same RI:ITI ratios generating equivalent performances in the lever-extended group, *A* scores continued to decline as RIs were lengthened within each RI:ITI ratio. This effect was confirmed by a significant difference between 6-sec RIs:24-sec ITIs and 3-sec RIs:12-sec ITIs in post hoc tests from the first ANOVA for the .25 RI:ITI ratio and by a main effect for ITI in the second ANOVA [$F(2,26) = 71.60$] for the .5 and 1 ratios. Thus, of the two types of ISIs, only variations in RI influenced performance.

The effects of Trial *n*−1 S2 schedule on speed of responding (reciprocal latencies) on the Trial *n* S2 schedule were analyzed by a four-way ANOVA (group × trial *n*−1 S2 schedule × trial *n* S2 schedule × stimulus modality presentation, with repeated measures on the last three factors) at each RI/ITI combination in this and the next experiment. No Trial *n*−1 S2/Trial *n* S1 schedule sequence effects were found that even remotely approached significance in either experiment (*ps* > .30). Thus, no evidence of intertrial PI was obtained in this study.

EXPERIMENT 2

Although retracting the lever between trials reduced the disruptive influence of increasing RIs in Experiment 1, as predicted by the trial tracking hypothesis, this effect was modest and limited to sessions in which all the stimuli were of the same modality. Furthermore, alternating modalities did not further reduce the effects of lengthening RIs, nor did massing trials generate poorer performance in the lever-extended control group. The second experiment was designed to investigate a possible alternative explanation of the rather modest effects of the lever presentation variable from Experiment 1.

Rather than allowing the rats to better differentiate RIs from ITIs, lever extensions at S1 and retractions at S2 may have merely helped the experimental rats maintain attention to an otherwise highly repetitive task. By the same token, the control rats would have more easily reduced their attention to this task in a more static environment in which the lever remained always extended. According to this *attention decrement* hypothesis, where the lever is extended or retracted in the session should have little effect on performance, other than to maintain attention to the task.

We tested this hypothesis in Experiment 2 by retraining and testing some of the same rats from Experiment 1 with the lever being extended at the onset and being retracted at the offset of each stimulus (RI/ITI lever-retracted group). For the other rats from that experiment, the lever was retracted only during ITI, as in Experiment 1 (ITI lever-retracted group). According to the attention decrement hypothesis, extending the lever at the onset

and retracting it at the offset of each stimulus should promote similar attention to the task and generate comparable performance as extending it at the onset of S1 and retracting it at the offset of S2. According to the trial tracking hypothesis, however, the rats in the ITI lever-retracted group should be less disrupted by changes in RI, if not by changes in ITI, than those in the RI/ITI lever-retracted group, because only in the former can the rat determine that a trial has not ended during long RIs by noticing the lever's presence.

Method

Subjects and Apparatus. The same rats and apparatus as those from the previous experiment were used in this experiment. The rats from the two lever presentation groups in Experiment 1 were randomly and equally divided into the experimental ITI lever-retracted group and a control RI/ITI lever-retracted group.

Procedure. The same basic training and testing procedures as those from Phase 2 of Experiment 1 were employed, with the following modifications. As already described, rather than the lever being continuously extended for control rats, it was extended at the onset and retracted at the offset of each stimulus (the RI/ITI lever-retracted group). The rats received initial training with 3-sec RIs and 24-sec ITIs for 32 sessions. The same testing sequence of 36 sessions as that used in Experiment 1 was employed in this experiment. The rats continued to receive up to 96 S1–S2 trials per session, and their behavior was analyzed in the same manner as that in Experiment 1.

Results and Discussion

Despite increased training, the control rat that had failed to acquire the conditional task in Experiment 1 was still unable to learn it under the RI/ITI lever retraction condition. The remaining rats developed A scores above .80. Consequently, only data from the 8 rats in the reconstituted ITI lever-retracted experimental group were compared with those of the remaining 7 rats in the new RI/ITI lever-retracted control group.

Figure 3 shows the effects of RI and ITI variations, pooled over stimulus modality sequences in the same manner as that in Figure 2. Stimulus modality sequences were disregarded, because they did not interact with the other independent variables or generally influence performance. Two effects are evident in the top panel. First, A scores declined to lower levels in the RI/ITI lever-retracted group than in the ITI lever-retracted group as RIs were lengthened. This observation was supported by an interaction between group and either RI (3-sec and 6-sec) or RI/ITI ratio [$Fs(1,13) = 23.24$ and 34.45]. Second, differences between groups at 6-sec RIs were greater at 6-sec ITIs than at the other two ITI durations. This effect resulted from both a steeper decline in A scores in the RI/ITI lever-retracted group and the disappearance of any such reductions in the ITI lever-retracted group as RIs were lengthened from 3 to 6 sec. A triple interaction among group, RI, and ITI [$F(2,26) = 6.58$] supported these observations. Individual comparisons within each group revealed that A scores in the experimental group only significantly declined as RIs were increased to 6 sec

within the 24-sec ITI block of sessions. While A scores in the control group significantly declined within each ITI block, they were significantly steeper within the 6-sec ITI block than within the 12- or 24-sec ITI blocks. These results conform more to the trial tracking than to the attention decrement account of the effects of lever retractions.

An examination of the effects of shortening ITIs at each RI in the bottom panel shows an *ITI* effect for the RI/ITI lever-retracted group at 6- and 12-sec RIs. A scores declined in the control group when ITIs were shortened from 12 to 6 sec at 6-sec RIs or from 24 to 12 sec at 12-sec RIs. On the other hand, A scores increased in the experimental group as ITIs were shortened from 24 to 6 sec at 6-sec RIs and from 24 to 12 sec at 12-sec RIs. The triple interaction from the first ANOVA already noted confirmed the opposite effects of shortening ITIs at 6-sec RIs between groups. An interaction between group and ITI [$F(1,13) = 14.2$] at the 12-sec RI was also found, but individual comparisons revealed only a significant decrease in A scores as ITIs were shortened in the ITI lever-retracted experimental group. Although, shortening ITIs increased the disruptive effects of lengthening RIs in the RI/ITI lever-retracted group, as expected, these rats did not display equivalent performance at the same RI/ITI ratios. Rather, A scores increased as ITIs were shortened within each RI:ITI ratio, as was confirmed by an individual comparison between 6-sec RIs:24-sec and 3-sec RIs:12-sec ITIs for the .25 ratio and by a main effect for ITI [$F(2,26) = 104.14$] for .5 and 1 ratios.

Results from this experiment offer stronger evidence for the trial tracking hypothesis than do those of Experiment 1. Although performance was still more affected by changes in RI than by changes in ITI, shortening the latter increased the disruptive effects of lengthening the former in the RI/ITI lever-retracted group. Predictions of the trial tracking hypothesis that alternating modalities between S1 and S2 would reduce such performance decrements and that equivalent performance would occur at the same RI:ITI ratio still were not confirmed. Perhaps animals found it difficult to learn to use consistent alternations in modality to signal when trials start and end in some sessions when other sessions with the same modality were interspersed among them. It may be that consistent alternations in S1 and S2 modality would have generated better performance if this condition had occurred over several successive sessions without being interrupted by sessions with the S1 and S2 of the same modality.

To account for the lack of any equivalent performance at the same RI:ITI ratios in the present study, we note that this RI:ITI ratio effect broke down at 4-sec RIs in pigeons (Roberts & Kraemer, 1982, Experiment 2). Roberts and Kraemer attributed the elimination of this effect to a decline in working memory of the sample stimulus over this relatively longer RI. As RIs in the present study were longer than those at which the RI:ITI effect was observed in pigeons, possible decrements in working memory of each schedule at S1 may have also obscured any RI:ITI ratio effects in our rats.
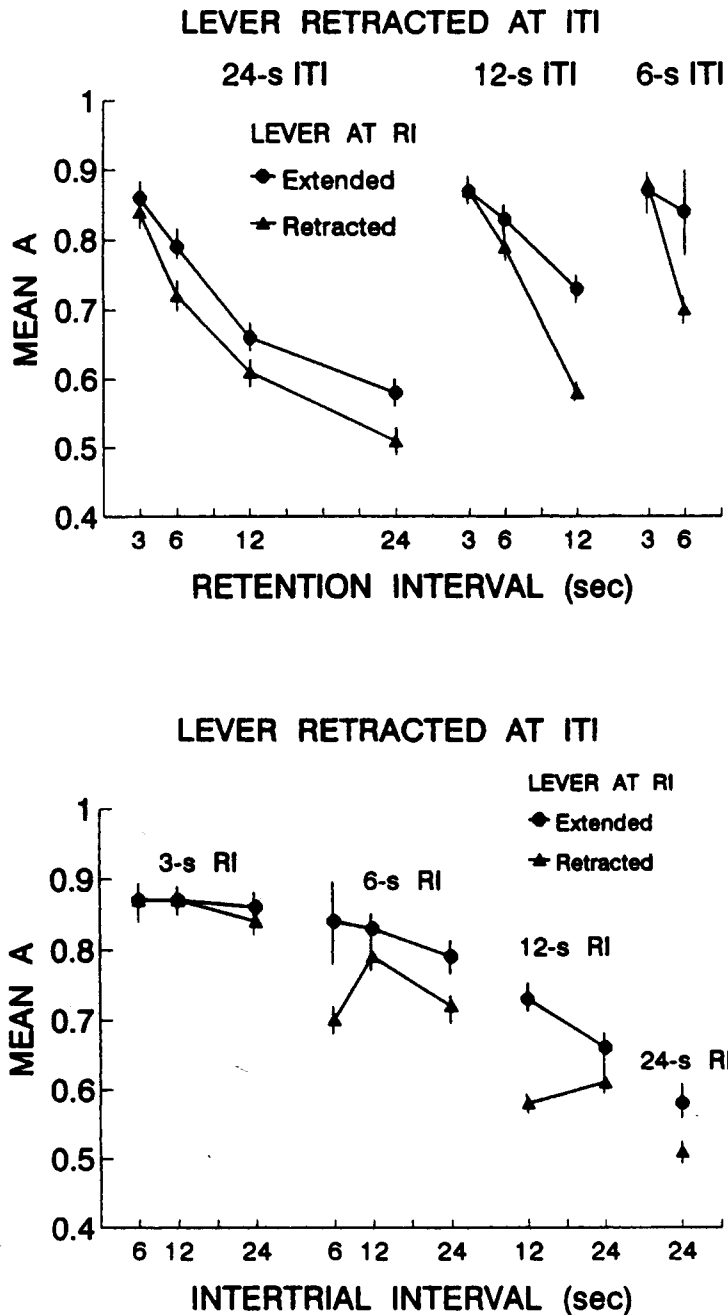
LEVER RETRACTED AT ITI

Figure 3. Mean *A* scores as a function of changes in retention interval (RI) within each intertrial interval (ITI) block of sessions (top panel) and as a function of changes in ITI at each RI for each group in Experiment 2. Vertical lines represent standard errors of means. The lever was retracted at ITI for all the rats, but at RI was extended in the experimental group and retracted in the control group.

GENERAL DISCUSSION

Findings from the present study add to an accumulating body of evidence that disruption of performance in delayed conditional discriminations cannot be solely attributed to declines in or interference with working memory for sample events. Declines in delayed conditional discriminations have also been attributed to decreased

general attention to or processing of sample events (Edhouse & White, 1988; Roitblat & Scopatz, 1983), changes in temporal expectancies for reinforcement (Roberts & Kraemer, 1982), and confusion about the instructional rules of the task in reference memory (Zentall, 1997). Our results indicate that decreases in trial tracking accuracy must also be added to this list. It would appear that extending the lever at the beginning and retracting it at the

end of trials allows rats to better determine whether they are within or between trials as RIs approach relatively long ITIs. Rats are more likely to lose track of where they are in a session when the lever is always either extended or retracted during each interstimulus interval (ISI).

It is important to note that lengthening RIs seemed to disrupt behavior more than did shortening ITIs in either the lever-extended or the RI/ITI lever-retracted group. Indeed, effects from changes in ITI only emerged in Experiment 2, where massing trials increased the disruptive effects of lengthening RIs in the RI/ITI lever-retracted group. There are two possible explanations for this asymmetrical effect. One is that trial tracking accuracy symmetrically declined as RIs and ITIs approached each other but that declines in working memory for sample events over long RIs obscured this symmetry. An alternative explanation assumes that trial tracking declined more when RIs approached longer rather than shorter ITIs because initial, extensive training with 3-sec RIs and 24-sec ITIs caused the rats in both groups to associate long ISIs with ITIs and short ISIs with RIs. Retracting the lever at ITIs only reduced such confusions. According to this account, continued testing with variations in RI and ITI should break down these associations and make the presence and absence of the lever more salient features for determining the nature of an ISI in experimental rats. Therefore, the disruptive effects from increasing RIs or decreasing ITIs should eventually become more symmetrical in control rats but diminish in experimental rats.

Our trial tracking hypothesis bears some resemblance to Zentall's (1997) task instruction confusion hypothesis, as both maintain that disruption of performance reflects some type of confusion by the animal. This comparability is particularly apparent in our second explanation of the asymmetrical effects of changes in RI and ITI. The rats should have learned that short within-trial intervals alternate with long between-trials intervals during initial training. Therefore, they should have learned to alternate between active rehearsing of the appropriate conditional rule within one ISI (RI) and stopping such activity within the other (ITI). The rats may have simply found it more difficult to initially maintain these alternating modes of activating and deactivating stored instructions between long RIs and long ITIs than between short RIs and short ITIs.

These hypotheses, however, make opposite predictions about the fate of performance over extended testing. According to the task instruction confusion hypothesis, rats ought eventually to learn to more accurately alternate between these two modes of processing information over long intervals and thus improve their performance. The trial tracking hypothesis predicts that rats should develop as much difficulty in alternating between these two modes over short as over long intervals if they cannot otherwise differentiate between RIs and ITIs. The fact that shortening ITIs to RIs eventually led to decreased performance in such rats in Experiment 2 supports the trial tracking notion.

It will be recalled that the trial tracking hypothesis assumes that an animal may stop processing the target event in its working memory when it confuses RIs with ITIs and considers the trial to have ended. The second proposition derived from this initial assumption is that, if animals can differentiate between RIs and ITIs by means other than differences in their durations, they ought to maintain their attempts to process information from the target event during long RIs. The idea that retention of information in working memory is related to a subject's general expectation about whether the trial has ended is central to recent active processing accounts of working memory (Kendrick & Rilling, 1986; Roberts, 1998; Roper & Zentall, 1993). Evidence from several studies on *directed forgetting* in pigeons (Kendrick, 1984, as reported by Kendrick & Rilling, 1986; Kendrick, Rilling, & Stonebraker, 1981; Maki & Hegvik, 1980), for example, supports this idea. Results from these studies reveal that a directed-forgetting cue following a sample stimulus only reliably reduces pigeons' matching accuracy on surprise, probe tests when it has previously signaled the end of a trial rather than a sample-independent discrimination.

Further evidence for the second proposition comes from research on pigeons' biased forgetting in delayed symbolic matching to FR samples (Fetterman & MacEwen, 1989) and to sample durations (Spetch & Rusak, 1992a). Both studies found that pigeons' typical bias to choose the test stimulus associated with the smaller sample FR or the shorter visual sample over long RIs was reduced or eliminated when the lighting in the chamber consistently differed between RIs and ITIs. The trial tracking hypothesis can account for the elimination of these biased-forgetting effects when one considers that *choose short or small* is also the default option on nonsample probe trials with long RIs (Fetterman & MacEwen, 1989; Grant & Spetch, 1991; Santi, Coyle, Copps, & Ross, 1998). As RIs during these probe trials are essentially the same as ITIs that occur between tests and the next sample, *choose short or small* biases within sample-present trials may reflect birds' incorrect decisions that they are between rather than within trials. Differentiating RIs from ITIs by systematic variations in chamber illumination, however, removes this confusion and, therefore, this bias. Unlike results from the present study, such differentiation between RIs from ITIs does not improve overall matching accuracy in these preparations. Therefore, improved trial tracking may not necessarily enhance actual retention of sample events over long RIs but does prevent the termination of processing them within RIs. Finally, it should be noted that other models that account for these biased-forgetting effects, such as the *relative duration* hypothesis (Spetch & Rusak, 1992b) and *scalar timing* theory (Santi et al., 1998) are not incompatible with our trial tracking hypothesis. Rather, they suggest specific mechanisms that animals may use to determine whether they are within or between trials.

REFERENCES

COHEN, J. S., & ARMSTRONG, D. L. (1996). Sources of intertrial proactive interference in rats' short-term memory in a delayed successive matching-to-sample modality discrimination. *Learning & Motivation, 27*, 485-499.

COHEN, J. S., & ROBERTS, R. (1996). The role of trial tracking on rats' successive delayed matching-to-sample modality performance. *Behavioral Processes, 36*, 277-287.

EDHOUSE, W. V., & WHITE, K. G. (1988). Sources of proactive interference in animal memory. *Journal of Experimental Psychology: Animal Behavior Processes, 14*, 56-70.

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.

GIBBON, J. (1977). Scalar expectancy theory and Weber's Law in animal timing. *Psychological Review, 84*, 279-325.

GRANT, D. S. (1975). Proactive interference in pigeon short-term memory. *Journal of Experimental Psychology: Animal Behavior Processes, 104*, 207-220.

GRANT, D. S., & SPETCH, M. L. (1991). Pigeons' memory for event duration: Differences between choice and successive matching tasks. *Learning & Motivation, 22*, 180-199.

GREEN, M., TERMAN, M., & TERMAN, J. S. (1982). Comparison of yes–no and latency measures of auditory intensity discrimination. *Journal of the Experimental Analysis of Behavior, 32*, 363-372.

HERMAN, L. M. (1975). Interference and auditory short-term memory in the bottlenosed dolphin. *Animal Learning & Behavior, 3*, 43-48.

JARRARD, L. E., & MOISE, S. L. (1971). Short-term memory in the monkey. In L. E. Jarrard (Ed.), *Cognitive processes of nonhuman primates* (pp. 1-24). New York: Academic Press.

KENDRICK, D. F., & RILLING M. E. (1986). AIM: A theory of active and inactive memory. In D. F. Kendrick, M. E. Denny, & M. R. Denny (Eds.), *Theories of animal memory* (pp. 129-152). Hillsdale, NJ: Erlbaum.

KENDRICK, D. F., RILLING, M. E., & STONEBRAKER, T. B. (1981). Stimulus control of delayed matching in pigeons: Directed forgetting. *Journal of the Experimental Analysis of Behavior, 36*, 241-251.

MAKI, W. S., & HEGVIK, D. K. (1980). Directed forgetting in pigeons. *Animal Learning & Behavior, 8*, 567-574.

MAKI, W. S., MOE, J. C., & BIERLEY, C. M. (1977). Short-term memory of stimuli, responses, and reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes, 3*, 156-177.

MCNICOL, D. (1972). *A primer of signal detection theory.* London: Allen & Unwin.

NELSON, K. R., & WASSERMAN, E. A. (1978). Temporal factors influencing the pigeon's successive matching-to-sample performance: Sample duration, intertrial interval, and retention interval. *Journal of the Experimental Analysis of Behavior, 30*, 153-162.

ROBERTS, W. A. (1980). Distribution of trials and intertrial retention of delayed matching to sample with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 6*, 217-237.

ROBERTS, W. A. (1998). *Principles of animal cognition.* New York: McGraw-Hill.

ROBERTS, W. A., & KRAEMER, P. J. (1982). Some observations on the effects of intertrial interval and delay on delayed matching-to-sample in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 8*, 342-353.

ROITBLAT, H. L., & SCOPATZ, R. A. (1983). Sequential effects in pigeons' delayed matching-to-sample performance. *Journal of Experimental Psychology: Animal Behavior Processes, 9*, 202-221.

ROPER, K. L., & ZENTALL, T. R. (1993). Directed forgetting in animals. *Psychological Bulletin, 113*, 513-532.

SANTI, A. (1984). The trial spacing effect in delayed matching-to-sample by pigeons is dependent upon the illumination condition during the intertrial interval. *Canadian Journal of Psychology, 38*, 154-165.

SANTI, A., COYLE, J., COPPS, R., & ROSS, L. (1998). The timing of auditory and visual signals by pigeons: Differential sensitivity to intertrial duration. *Learning & Motivation, 29*, 345-365.

SPETCH, M. L., & RUSAK, B. (1992a). Temporal context effects in pigeons' memory for event duration. *Learning & Motivation, 23*, 117-144.

SPETCH, M. L., & RUSAK, B. (1992b). Time present and time past. In W. K. Honig & J. G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 47-67). Hilldale, NJ: Erlbaum.

WINER, B. J. (1971). *Statistical principles in experimental design.* New York: McGraw-Hill.

ZENTALL, T. R. (1997). Animal memory: The role of "instructions." *Learning & Motivation, 28*, 280-308.