

Same-different comparison of duration

J. GREGOR FETTERMAN
Arizona State University, Tempe, Arizona

Pigeons were trained on a task in which a red light of duration t_1 was followed by a green light of duration t_2 and then responses to different keys were reinforced according to whether the durations of the stimuli were the same or different. For Experiment 1, duration pairs consisted of all combinations of 1, 2, 4, and 8 sec. In Experiment 2, *different*-duration pairs included only combinations with $t_1 > t_2$ and, in addition, 2 subjects with extended training involving lesser-greater duration comparisons were transferred to the same-different task. Two of 3 subjects learned the task in Experiment 1 and analyses suggested that choices were based on specific instances, not on a temporal same-different rule. All 5 birds acquired the discrimination in Experiment 2, where it appeared that choices were controlled by a combination of relative and absolute rules. Accuracy decreased following transfer from a lesser-greater to a same-different discrimination, but performance was above chance on the first transfer session. In both experiments, however, accuracy was below that found in earlier work with lesser-greater comparisons of duration. These findings are discussed in relation to prior research with lesser-greater comparisons of duration and same-different tasks involving nontemporal stimuli.

Fetterman and Dreyfus (1986, 1987) introduced a new method for the study of animal timing whereby pigeons were presented with a red light of duration t_1 followed by a green light of duration t_2 and then different choices were reinforced depending upon whether $t_1 > t_2$ or $t_1 < t_2$. The durations of the stimuli t_1 and t_2 changed over trials such that approximately 700 different combinations were presented, and the animals responded correctly about 85% of the time, with accuracy a function of the ratio of a duration pair. Detailed analyses of performance suggested that, in most instances, choices were based on a relational comparison of the two durations.

These findings are somewhat surprising in light of a substantial literature indicating that pigeons have difficulty learning relational concepts such as the same-different rule (e.g., see Carter & Werner, 1978, for a review). For example, pigeons have been trained on tasks (e.g., matching-to-sample—MTS) that would appear to require relational (same-different—S-D) comparisons; yet considerable evidence with MTS and related procedures indicates that pigeons learn stimulus-specific rather than relational response rules (Carter & Werner, 1978; Premack, 1978). The matter is somewhat controversial, and there is considerable debate about what does and does not provide an appropriate assay of relational learning. Transfer tests are typically used as the measure of conceptual behavior, but the available transfer data have not resolved the issue

(e.g., D'Amato, Salmon, & Colombo, 1985; Premack, 1978).

The present experiments represent an extension of prior research with the duration comparison task; the experiments were motivated by the question of whether pigeons could be trained to discriminate other temporal relations between stimuli. The initial work (Fetterman & Dreyfus, 1986, 1987) required pigeons to respond on the basis of which of two stimuli was longer. In subsequent work (Stubbs, Dreyfus, & Fetterman, 1987), different choices were reinforced according to whether a duration pair ratio was less or greater than a criterion ratio, for example whether the first stimulus was more or less than twice the second. Discrimination performance was comparable whether the task involved a simple shorter-longer judgment or the seemingly more complex judgment of whether a duration ratio was less or greater than some arbitrary criterion ratio. The question addressed here was whether pigeons could be trained to discriminate a temporal S-D relationship. Would a temporal S-D discrimination be learned as readily as one involving a discrimination based on the relative durations of the stimuli?

Pigeons were trained to make one response when two successive stimuli were of the same duration and the alternate response when the durations of the stimuli were different. In preliminary work, hundreds of duration combinations were presented, half of them same- and the other half different-duration pairs. The animals showed no evidence of learning the discrimination under these training conditions, so the task was simplified and made comparable to one used in preliminary work with lesser-greater (L-G) comparisons (Fetterman & Dreyfus, 1986). These results are reported in Experiment 1. On the basis of the results of Experiment 1, the task was modified for Experiment 2 in the hope of improving discrimination performance. The results of both experiments indicate that,

This research was conducted at Norwich University, Northfield, VT, and supported by a Dana Research Fellowship from Norwich University to J. G. Fetterman. Preparation of this report was supported by Department of Health and Human Services NRSA Postdoctoral Fellowship 1 F32 MH090306 from NIMH to J. G. Fetterman. Thanks are extended to Leon Dreyfus, Peter Killeen, and Alan Stubbs for comments on an earlier version of this manuscript. Requests for reprints should be addressed to J. G. Fetterman, Department of Psychology, Arizona State University, Tempe, AZ 85287.

although the animals learned the temporal S-D discrimination, the task was more difficult than comparable ones involving judgments about the relative magnitudes of two durations or those involving S-D comparisons of color, form, or pictorial stimuli (e.g., Santiago & Wright, 1984).

EXPERIMENT 1

Method

Subjects. The subjects were 3 adult male Silver King pigeons maintained throughout the experiment at 80% of their free-feeding weights. All had prior experimental histories, none on pair-comparison timing tasks.

Apparatus. The experimental space was a standard three-key (BRS-LVE) pigeon chamber measuring 31 cm wide, 35 cm deep, and 35 cm high. The front wall contained three pecking keys located in a row, with 8 cm between centers, 24 cm above the chamber floor. The feeder opening was located below the center response key and measured 6 cm square. The bottom of this opening was 10 cm above the chamber floor. A houselight was mounted over the center key, 32 cm above the floor. Sessions were conducted in a room in which white noise was played to mask extraneous sounds. Additional masking and ventilation were provided by a fan attached to the chamber wall. An Apple IIe microcomputer and MED interface, located in an adjacent room, controlled the experiment and recorded events.

Procedure. A trials procedure was used whereby each trial began with the center key lit white. A peck to this key changed the key color to red and initiated the first duration. The red light remained on for a given duration, which varied over trials, and then changed to green, demarcating the second duration. The green light remained on for a specified duration, changing over trials, and then went off independently of behavior. The offset of green was followed immediately by illumination of the outside keys by amber light. A response to one of these keys was correct if the duration of red was different from that of green (e.g., 4 sec followed by 1 sec or 1 sec followed by 4 sec), whereas a response to the alternate key was correct if the durations of red and green were the same (e.g., 4 sec followed by 4 sec). Correct responses were reinforced with 4-sec access to mixed grain followed by a 20-sec intertrial interval (ITI) and a new trial period. Incorrect responses initiated the 20-sec ITI directly. The keylights and a houselight that was lit during trials were all off during food deliveries and intertrial intervals. For Pigeons 76 and 80, a left-key response was correct when the duration of red differed from that of green and a right-key response was correct when red and green were of the same duration. This arrangement was reversed for the third subject, Pigeon 84.

Duration pairs were constructed by combining the durations 1, 2, 4, and 8 sec in all possible ways, with the result that there were four equal pairs of durations (1-1, 2-2, 4-4, and 8-8 sec) and 12 unequal pairs (8-1, 1-8, 4-1, 1-4, 2-1, 1-2, 8-2, 2-8, 4-2, 2-4, 8-4, and 4-8 sec). Sessions contained equal numbers of problems from each category (*same* and *different*), and the various duration combinations within each category were presented equally often.

The birds were trained for 17 sessions with a correction procedure and with reinforcement for every correct response. A non-correction procedure was used for the final 21 sessions, and food was scheduled to ensure equal reinforcement of left-key and right-key responses. Reinforcers were assigned to each choice with a probability of .50 and held until the appropriate duration pair (same or different) and response occurred. The alternate correct response could not be reinforced until the assigned reinforcer was collected (Stubbs, 1976). This arrangement resulted in food delivery for approximately 50% of the correct responses. Correct responses that did not result in food illuminated the feeder light for 0.5-sec, followed by the ITI. Sessions were conducted 6 days a week and lasted until 50

reinforcers were obtained. Training continued until performance, by visual inspection, appeared stable for 10 sessions. The results that follow are based upon the final five sessions of training, with each session containing approximately 150 trials. After these final data were taken, the choice contingencies were reversed for a single probe session. All statistical tests were two-tailed.

Results and Discussion

Table 1 presents summary measures for individual pigeons. Columns 1 and 2 show the average and range of A' , a nonparametric index of sensitivity (Grier, 1971). The values of A' for Pigeons 76 and 84 were significantly greater than the chance level of .50 [$t(4) = 7.26, p < .01$, and $t(4) = 5.45, p < .01$]; the performance of Pigeon 80 was not significantly different from chance [$t(4) = 1.57, p > .05$]. The third column shows B'' , a nonparametric index of bias (Grier, 1971), which can range between -1.0 and 1.0 . Positive values indicate a bias to respond "different," negative values to respond "same"; a value of zero indicates no bias. Only Pigeon 76 showed an appreciable bias, toward the "same" choice alternative. The next two columns show percent correct for same- (column 4) and different- (column 5) duration pairs. Pigeons 76 and 84 were more accurate on *same* than on *different* stimulus pairs, whereas Pigeon 80 exhibited the opposite pattern. Columns 6 and 7 show accuracy separately for *different* pairs with $t_1 > t_2$ (column 6) or $t_1 < t_2$ (column 7). All 3 birds were more accurate when the first stimulus was the longer of a duration pair. A matched t test indicated that this difference approached, but did not reach, significance [$t(2) = 3.38, p > .05$].

The top and lower left panels of Figure 1 show percent correct responses for individual pairs of durations. The left side of each panel shows the *different* pairs of durations, and the right side, the *same* pairs. The *different* pairs are arranged such that each set of bars shows performance with the identical pair of durations. The filled and unfilled bars show accuracy when the longer stimulus was, respectively, the first (red) or second (green) stimulus of the pair. The duration pairs are ordered along the abscissa to reflect a decreasing ratio between the stimuli.

In 14 of 18 comparisons of performance with identical sets of durations, accuracy was higher when the longer of two stimuli was presented first. In many instances the difference is pronounced, with accuracy well below chance when the first stimulus was the shorter pair member and well above chance with the opposing order of stimuli (e.g., compare 8 and 1 sec for Pigeons 76 and 80). Figure 1 also shows that accuracy decreased as the ratio

Table 1
Summary Measures of Performance for Experiment 1

Pigeon	A'		B''	Percent Correct			
	Average (1)	Range (2)		Same (4)	Different (5)	$t_1 > t_2$ (6)	$t_1 < t_2$ (7)
76	.70	.67-.76	-.17	77%	46%	50%	42%
84	.61	.53-.68	-.02	60%	53%	60%	45%
80	.46	.43-.53	0.0	45%	52%	64%	40%

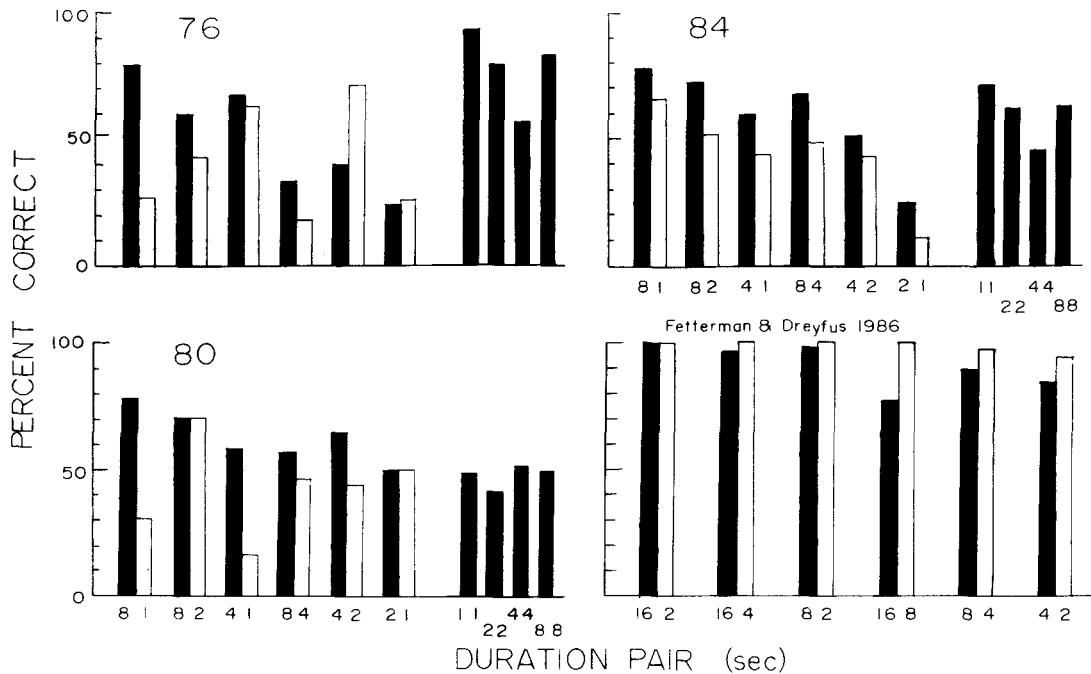


Figure 1. Percent correct responses for individual pairs of durations. Each pair of bars on the left shows accuracy for a given pair of durations; the filled bars show accuracy for pairs with $t_1 > t_2$, and the unfilled bars show accuracy for pairs with $t_1 < t_2$. The filled bars on the right of the top two panels and the lower left panel show accuracy for equal pairs of durations. The panel in the lower right shows data from a study by Fetterman and Dreyfus (1986). See text for details.

of the durations approached 1:1 (i.e., equality), but mainly for pairs with $t_1 > t_2$. The right side of each panel shows that Pigeons 76 and 84 were reasonably accurate on the various *same* pairs, with the exception of the 4- versus 4-sec pair. Pigeon 80 was at or slightly below chance on each of the *same* pairs.

The panel in the lower right of Figure 1 shows data, averaged over 4 pigeons, from Fetterman and Dreyfus (1986), who used a task similar to the one employed in the present experiment. They paired the durations 2, 4, 8, and 16 sec with the requirement that the pigeon make one response when $t_1 > t_2$ and the alternate response when $t_1 < t_2$. Two points of comparison are pertinent. First, pigeons were more accurate on the L-G than on the seemingly comparable S-D discrimination; the animals trained on the L-G task responded appropriately about 90% of the time. Second, the data of Fetterman and Dreyfus indicate a negative time-order error; generally, their pigeons were more accurate when the longer of two stimuli was the second member of a duration pair (unfilled bars), a pattern opposite to that observed in the present experiment.

Table 1 and Figure 1 indicate that 2 of 3 birds "learned" the S-D task by a criterion that identifies learning with a performance that reliably exceeds chance. But this measure of learning does not speak to the issue of whether the pigeons responded on the basis of a S-D concept, or whether they learned a number of specific choice

rules. The data of Figure 1 suggest that choices were based on stimulus-specific rules, not a general S-D concept. Several features of the data are relevant to this point. First, accuracy was well below chance on selected *different* pairs, suggesting the consistent misapplication of a choice rule. For example, accuracy scores for both birds that "learned" the task were below chance when the durations 1 and 2 sec were presented in either order, and above chance when *same* pairs were composed of the durations 1 or 2 sec. This result indicates a strategy of responding "same" when both durations were short. Pigeon 76 evidently employed a second rule that involved just a single pair member. This bird tended to respond "same" whenever the second stimulus of a pair was 8 sec, a strategy that resulted in below-chance performance for the pairs 1 versus 8 sec, 2 versus 8 sec, and 4 versus 8 sec, and reasonably accurate performance for the 8-versus 8-sec pair (see Figure 1). Accuracy intermediate to these extremes may plausibly be attributed to generalization from these exemplars, for example when the sum of the durations was neither "short" nor "long."

Data from the reversal session (see Procedure) provide additional evidence that Pigeons 76 and 84 adopted specific response rules. Figure 2 presents these data showing percent correct responses for the various combinations of stimuli before and after the reversal. For the *different* pairs, circles represent stimuli with $t_1 > t_2$ and triangles those with $t_1 < t_2$. The filled and unfilled points

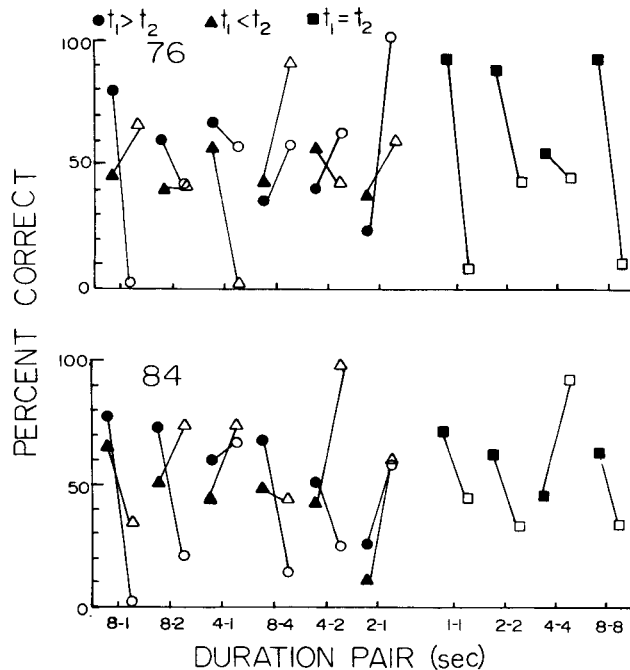


Figure 2. Percent correct responses for individual duration pairs before (filled symbols) and after (unfilled symbols) reversal of the choice contingencies. Circles represent duration pairs with $t_1 > t_2$ and triangles those with $t_1 < t_2$. Squares indicate same duration pairs.

represent, respectively, performance before and after the reversal. The point of these data is to show that, in general, the reversal affected performance on some but not all pairs of stimuli, a result that indicates the learning of rules based on specific instances.

First, sensitivity decreased following the reversal; the decreases in A' were from .76 to .18 for Pigeon 76 and from .65 to .45 for Pigeon 84. But the more important changes involved individual pairs of stimuli. The changes were greatest for pairs with which the birds were well below or above chance prior to the reversal. For example, accuracy decreased markedly for the pairs 8 versus 8 sec and 1 versus 1 sec. These were, presumably, the best exemplars of the choice rules "respond same when long is followed by long or short is followed by short." Conversely, accuracy improved in cases in which the choice strategy did not coincide with the experimental dependencies of the training phase, for example when the durations 1 and 2 sec were presented in either order. These represent instances in which, during training, the birds misapplied the "short followed by short" response rule.

It seems likely then that the birds "solved" the task by grouping pairs of stimuli according to specific rules and that they did not acquire the temporal S-D concept. These results are opposed to those obtained by Fetterman and Dreyfus (1986, 1987) with L-G comparisons of duration. Although their birds appeared to use specific choice rules in some instances, Fetterman and Dreyfus demonstrated that relational comparisons were made in others.

Thus, the temporal S-D task appears more difficult for pigeons that seem unable to "learn the concept." These points are related. The pigeons, apparently unable to extract the S-D rule, learned choice rules based on specific instances, a factor that likely resulted in the poorer discrimination performance on the S-D task.

Why was the S-D task more difficult than a seemingly comparable L-G task? There are several related possibilities. From the standpoint of signal detection theory (Green & Swets, 1974), the discrimination can be viewed as involving three, not two, overlapping distributions of sensory effects. According to this analysis, the different configurations of stimuli $t_1 > t_2$, $t_1 = t_2$, and $t_1 < t_2$ were associated with three hypothetical distributions located at different points along the decision axis. The scale of this axis might reflect a difference between the durations or, alternatively, their ratio. (See Noreen, 1980, and Vickers, 1979, for discussions of these points and related applications of a "sensory difference model" to human S-D judgments.) This situation is more complex because it requires two decision criteria located on either side of the distribution centered over $t_1 = t_2$, and because subjects must classify two stimulus configurations ($t_1 > t_2$ and $t_1 < t_2$) under a single response category, "different." Table 1 and Figure 1 indicate that the animals had considerable difficulty in treating the two configurations similarly.

This analysis suggests procedural modifications that might simplify the task for the animals. For example, pairs of stimuli that contained only one configuration of *different* pairs ($t_1 > t_2$ or $t_1 < t_2$) could be used. This modification would simplify the task in several respects: It would eliminate the potential for response competition resulting from the grouping of positive and negative differences (or large and small ratios) in the same category; and it would eliminate the need for multiple decision criteria. Alternatively, a three-category task could be used in which different responses were associated with $t_1 > t_2$, $t_1 = t_2$, and $t_1 < t_2$. This three-response classification task might be easier for subjects to master than one that lumped the $t_1 > t_2$ and $t_1 < t_2$ categories together.¹ Experiment 2 employed the first of these procedural modifications to determine whether a level of performance superior to that observed in Experiment 1 might be attained.

EXPERIMENT 2

The pair comparison procedure was modified for Experiment 2 so that different-duration pairs included only those with a first duration longer than the second. This modification was motivated by the analysis of differences between the S-D and L-G tasks described above, and by the observation that all 3 birds were more accurate on *different* pairs with $t_1 > t_2$. Thus, the task was simplified in several ways in an attempt to encourage levels of performance superior to those observed in Experiment 1.

Experiment 2 also provided a within-subjects comparison of performance with the L-G and S-D duration comparison procedures. Although the levels of sensitivity ob-

served in Experiment 1 were below those found in earlier work with L-G judgments (Fetterman & Dreyfus, 1986, 1987), the comparisons were less than ideal because they were made across different subjects and experiments. In the second experiment, 2 pigeons with extensive histories on the L-G task were transferred to the S-D task, providing a direct comparison of performance on the two discrimination problems.

Method

Subjects. Five adult male pigeons served. Three of these birds (Pigeons 76, 84, and 80) had served in Experiment 1. The other 2 birds (Pigeons 53 and 68) had had prior experience on the L-G duration comparison task (see below). All were maintained at 80% of their free-feeding weights.

Apparatus. The apparatus was the same as that described in Experiment 1.

Procedure. The procedure was identical in most respects to that described in Experiment 1. For the subjects that had served in the first experiment, the appropriate response following same- and different-duration pairs was unchanged. For the 2 new birds, Pigeons 53 and 68, left- and right-key responses, respectively, were correct when the duration of red was different from that of green (i.e., $t_1 > t_2$); the alternative response was correct when the durations of red and green were the same. Table 2 lists the duration pairs used in Experiment 2—7 *same* and 24 *different* pairs. Note that all *different* pairs have a first duration longer than the second, that values of the second stimulus are limited to those used to construct *same* stimulus pairs, and that the ratios of these pairs cover a substantial range, from 6:1 to 1.2:1.

The two categories of stimuli, *same* and *different*, occurred equally often, and pairs within the *same* category were presented equally often. The probabilities of occurrence of individual *different* pairs were specified by their being grouped according to the value of the second stimulus. On *different* trials, one of the seven groups of stimuli was selected with each having an equal probability (approximately .14) of selection, and finally one pair from the group was presented with individual pairs having an equal probability of presentation.

The different values of the second stimulus (see Table 2) thus occurred with equal relative frequencies on *same* and *different* trials, and in this sense the value of the second stimulus was not predictive of the correct response. Other absolute features of the stimulus pairs (e.g., the value of the first stimulus or the total of the two durations) were sometimes (though infrequently) predictive of choice, but Fetterman and Dreyfus (1986, 1987) found no evidence of control by these features and only occasionally by the value of the second stimulus alone. It proved impossible to construct pairs of stimuli that eliminated all these potential absolute sources of control but, on the basis of past research, it seemed appropriate to eliminate the predictiveness of the second duration.

A noncorrection procedure with controlled reinforcement for the two choices was used throughout the experiment. Feedback (a feeder flash) was provided for correct responses that did not result in food,

Table 2
Duration Pairs Used in Experiment 2

Different Pairs (sec)	Same Pairs (sec)
6-1, 4-1, 2-1,	1-1, 2-2, 4-4, 6-6,
10-2, 8-2, 6-2, 4-2,	8-8, 10-10, 12-12
12-4, 10-4, 8-4, 6-4,	
16-6, 14-6, 12-6, 10-6, 8-6,	
16-8, 14-8, 12-8, 10-8,	
16-10, 14-10, 12-10,	
16-12	

Table 3
Summary Measures of Performance for Experiment 2

Pigeon	A' (1)	Range (2)	B'' (3)	Percent Correct	
				Same (4)	Different (5)
76	.78	.67-.84	-.15	77%	60%
84	.58	.55-.62	-.04	65%	43%
80	.55	.45-.61	.01	43%	56%
53	.71	.66-.76	.03	60%	66%
68	.67	.62-.74	.02	58%	63%

and sessions ended after 50 reinforcers were obtained. All birds were trained for 20 sessions; most of the results that follow are based on summary measures calculated over the last five sessions of training. All statistical tests were two-tailed.

Transfer birds, Pigeons 53 and 68, were transferred to the S-D task after approximately 100 sessions of training on the L-G duration comparison task. Details of the latter procedure were as follows: The stimulus pairs were generated by a probability system in which a probability gate was pulsed every 0.5 sec. The probability of an output was 0.10 for each pulse. An output from the probability gate ended a duration; each duration lasted until there was an output or until 32 pulses had occurred. Individual durations ranged between 0.5 and 16 sec with a mean value of 5 sec. This method of arranging stimuli resulted in over 700 combinations of the two durations. The pigeons' task was to make one response if $t_1 > t_2$ and the alternate response if $t_1 < t_2$. Other procedural details—key colors, intertrial interval, hopper duration, and so forth—were identical to those employed with the S-D comparison task.

The assignment of correct responses for birds transferred from the L-G to the S-D task was chosen to foster positive transfer across the two procedures. The correct response following duration pairs with a longer first stimulus remained unchanged after the change from the L-G to the S-D task. The response formerly designated correct when the second stimulus was longer was correct when the two stimuli were of the same duration.

Results and Discussion

Table 3 presents summary measures in a format comparable to that of Table 1. The data in the first three rows are from the subjects of Experiment 1; the bottom two rows show the results for the transfer birds. All A' measures are significantly greater than the chance level of .50. The A' scores over the last five sessions of Experiment 2 were compared against those of Experiment 1 for Pigeons 76, 84, and 80. Although the average value of A' was higher in the second experiment for 2 of the 3 birds, the average for the group was not significantly above that for Experiment 1. Pigeon 76 showed a bias in favor of the *same* alternative; the other two birds showed no evidence of a response bias. Pigeons 76 and 84 were more accurate on *same* than on *different* pairs, whereas Pigeon 80 showed the opposite result, a pattern identical to that found in Experiment 1. A' scores for the transfer birds were slightly, but not significantly, above those of the other 3 birds. Neither showed a strong bias, and both were more accurate on *different* than on *same* duration pairs.

Figure 3 shows accuracy on individual pairs of durations. The top panel shows the data averaged for Pigeons 76, 84, and 80; the bottom panel shows them for Pigeons 53 and 68. For the top panel, variability across individuals

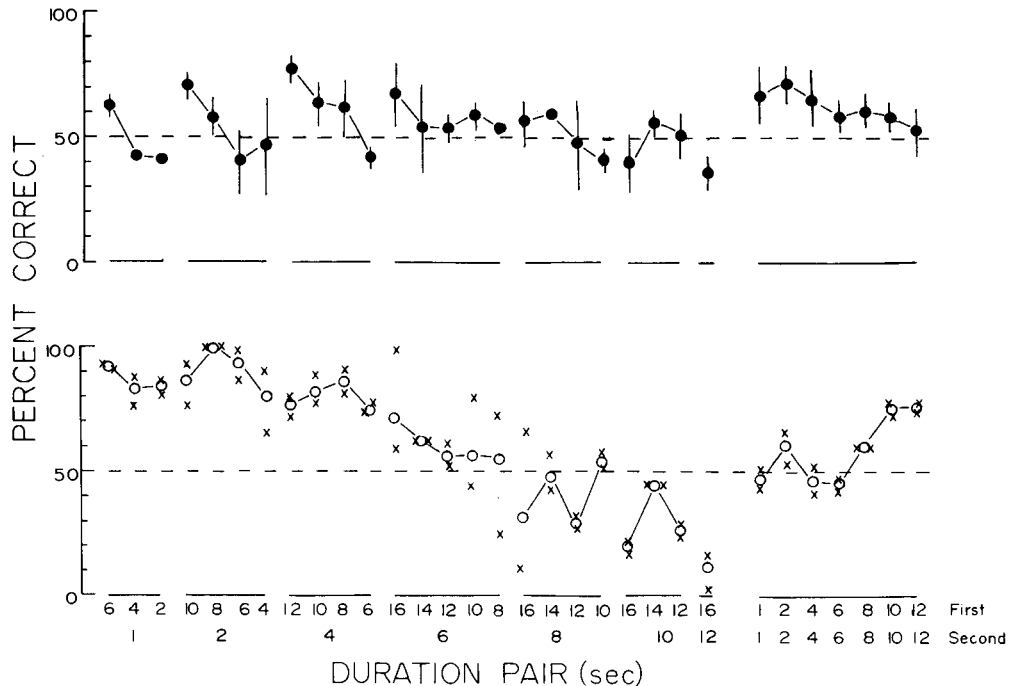


Figure 3. Percent correct responses for individual pairs of durations. The top panel shows the average for Pigeons 76, 84, and 80. The bottom panel shows the average for Pigeons 53 and 68. The lines through the symbols in the top panel indicate plus and minus one standard error of the mean for the 3 birds. The \times s in the bottom panel represent performance for individual birds.

is shown by the lines through the symbols, indicating plus and minus 1 standard error of the mean for the 3 birds; where no line is visible, the standard error was less than the diameter of the symbol. The \times s in the bottom panel represent performance for individual birds. Circles show the average of each pair of data points; the average was calculated from pooled data so that in some instances the symbols do not lie midway between the \times s. Individual *different* pairs are arranged in groups according to increasing length of the second stimulus. Performance with *same* pairs is shown in the far right of each panel.

Generally, accuracy decreased as the ratio of the stimuli approached 1:1 (equality), a pattern more obvious in the top than in the bottom panel, and with shorter rather than longer values of the second stimulus. Accuracy also decreased with increasing length of the second stimulus of *different* pairs, but, of course, the influence of this factor is comingled with a decreasing ratio of the stimuli. This change is most evident for the birds in the bottom panel. For these birds, accuracy on *different* pairs of stimuli averaged 77% when the second stimulus was less than or equal to 6 sec and 33% when the second stimulus was 8 sec or longer. These pigeons exhibited the opposite pattern for *same* pairs, a combination of results that indicates a strategy of responding "same" whenever the second stimulus exceeded a criterion value.

The data in Figure 3 are less revealing of absolute choice strategies than the data presented in Figure 1, and the trends within groups of stimuli could be taken as evidence for relational control by the ratio of a duration pair.

This interpretation is consistent with prior research involving L-G comparisons of duration (Fetterman & Dreyfus, 1986, 1987; Stubbs et al., 1987). Typically, however, a performance is considered relational to the degree that performance remains invariant over changes in the absolute values of the stimuli, and invariances of this sort appear rather elusive in studies of relational learning in animals. Thus, the issue in the present study is whether discrimination remained roughly constant for a given duration ratio (or duration difference) over the range of stimulus values. The data in Figure 3 do not resolve the issue; Performance was, in some cases, approximately invariant over changes in the absolute values of the stimuli (e.g., compare 2 vs. 1 sec, 4 vs. 2 sec, and 8 vs. 4 sec in the bottom panel), but not in others (e.g., compare 12 vs. 6 sec and 16 vs. 8 sec in the bottom panel with other pairs that stand in the ratio 2:1).

Figure 4 shows data across individual sessions before and after transfer from the L-G to the S-D task. Each panel shows performance for a single subject. From left to right, the panels show data from the last five sessions of training with the L-G task and the first and last (separated by dashed lines) five sessions of training with the S-D task. The top of each panel shows A' ; the bottom panels show percent correct for each category of stimuli. The symbols identify the stimulus categories.

A' scores decreased for both birds following the change from the L-G to the S-D task. However, both birds performed at levels significantly greater than chance on the first transfer session, and there was little change in per-

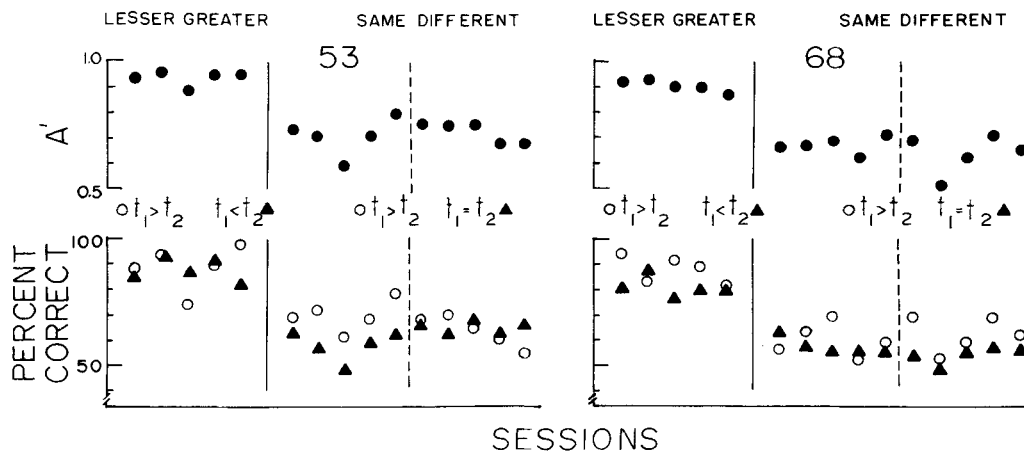


Figure 4. Performance before and after transfer from a lesser-greater (L-G) to a same-different (S-D) duration comparison task. From left to right, each panel shows performance during the last five sessions of the L-G task and for the first and last five sessions (separated by the dashed line) of training with the S-D task. The top row shows A' ; the bottom row shows percent correct for each category of stimuli. The symbols identify the stimulus categories (see text for details).

formance over the remaining 19 sessions of training. The bottom panels of Figure 4 show that accuracy was comparable for both categories of stimuli when choices were based upon a L-G comparison of durations. The change to the S-D task produced a decrease in accuracy for both choices, even though the appropriate choice following duration pairs with $t_1 > t_2$ was constant for both tasks; and there were changes in accuracy for the two categories over the course of training on the S-D task. Immediately after the change, Pigeon 53 showed a bias in favor of the $t_1 > t_2$ alternative, but accuracy was comparable for both categories over the last five sessions of training. This result indicates a shift in this animal's decision criterion produced by the replacement of $t_1 < t_2$ by $t_1 = t_2$ stimulus pairs. The change is not surprising, since substitution of the $t_1 = t_2$ for $t_1 < t_2$ pairs would decrease the distance along the decision axis between the hypothetical distributions of duration ratios. The changes for Pigeon 68 were not as clear in this regard, however.

In sum, procedural modifications in Experiment 2 produced marginal improvements in sensitivity for 2 of 3 birds, but there were fewer obvious "confusion" errors that suggested specific choice strategies. This was not the case in Experiment 1, in which the pattern of choices implicated specific choice strategies as the basis of the S-D discrimination (see Figure 1).

GENERAL DISCUSSION

The main result of these experiments was that pigeons had greater difficulty in learning a temporal pair comparison task when reinforcement was based on a S-D as opposed to a L-G rule. The differences in performance were substantial, and evident in comparisons made both within and between subjects and across two experiments employing different methods for constructing pairs of temporal stimuli. In Experiment 1, detailed analyses of perfor-

mance indicated that the animals responded primarily on the basis of specific instances, and not according to a temporal S-D concept. Analyses of the data of Experiment 2 suggested that the discrimination was, in that case, based upon the ratio of a duration pair. Whatever the basis of discrimination, however, the S-D task was clearly more difficult for pigeons than L-G judgments of duration.

What is the source of the difficulty of the S-D as compared with the L-G task? There are many plausible sources of difficulty. In Experiment 1, it seems plausible to view the relative difficulty of the discrimination in terms of the multiple criteria analysis discussed above (cf. Noreen, 1980). The pigeons were, obviously, unable to treat different configurations of a given pair of stimuli (i.e., $t_1 > t_2$ and $t_1 < t_2$) as comparable instances of the class *different*. That systematic confusions were frequent and similar for the 3 pigeons suggests a common method of engaging the problem; this finding merits further study.

Procedural modifications in Experiment 2 did not produce a significant improvement in discrimination as measured by the average for the 3 birds, but there were fewer "confusions" that would indicate control by absolute properties of the stimuli. Analyses of the data from Experiment 2 suggested control both by absolute and relational properties of the stimuli, and that the invariances were best characterized in terms of stimulus ratios (t_1/t_2), not stimulus differences ($t_1 - t_2$). This result is not surprising given the ubiquity of Weber's law in studies of animal timing (e.g., Gibbon, 1977).

The temporal S-D task is more difficult for animals than similar tasks involving nontemporal stimuli (e.g., colors; White, 1974). The differences are probably related both to the dimension of comparison and to the required comparison. Duration forms an intensive, prothetic (Stevens & Galanter, 1957) continuum and comparisons of stimuli along prothetic continua naturally involve judgments of relative magnitude (e.g., shorter-longer, brighter-dimmer,

etc.). S-D comparisons along prothetic continua may be difficult because they require a categorical coding of the stimuli, and this process is opposed by the effects of stimulus generalization (but see Kraemer, Mazmanian, & Roberts, 1985, for evidence suggesting a categorical coding of duration). The results of Creelman and Macmillan (1979) are consistent with this view. They studied human tone-frequency discrimination and found that same-different judgments were far more difficult than relative (higher-lower) judgments of pitch (60% vs. 88%). This result seems comparable to those obtained with pigeons making S-D versus L-G comparisons of duration.

We might also expect differences in performance on S-D tasks employing prothetic (e.g., duration) and metathetic (e.g., pictures) dimensions (noting that differences in the discriminability of the stimuli must be taken into account). Farrell (1985) reviewed the literature on human same-different judgments, noting that results differ depending upon whether "codable" or "uncodable" stimuli are used, and these differences were not artifacts of stimulus discriminability. Similarly, pigeons are more accurate on pair-comparison tasks with colors or pictures (e.g., Santiago & Wright, 1984) than was observed in the present experiment.

How do these and related results with duration comparison tasks fit in the context of research on relational learning in animals? As noted above, this area has its share of controversy, and there are differences of opinion regarding the status of relational learning in different species. D'Amato et al. (1985) reviewed several studies of relational learning in pigeons and offered alternative, non-relational interpretations of the data. But they also concluded that "possession of the identity concept is not an all-or-none affair" (p. 49) (cf. Premack, 1983). D'Amato and his associates concluded that relational concepts were best represented on a continuum by their degree of specificity-abstractness, as assessed by cross-dimensional and cross-modal transfer tests (and, it might be added, by intradimensional transfer tests).

We have just begun to explore the limits of pigeons' abilities to compare durations. For example, Fetterman and Dreyfus (1986) introduced novel pairs of durations on transfer tests and found good transfer to novel instances; but there were "confusions" indicating that the pigeons responded both to a temporal relation and to the absolute values of the stimuli. Subsequent work (Dreyfus, Fetterman, & Stubbs, 1986; Fetterman & Dreyfus, 1987; Stubbs et al., 1987) employed hundreds of stimulus combinations, and analyses indicated control by both relational and absolute properties of the stimuli, suggesting that the issue should not be treated in an "either-or" way.

For the present experiments, a case cannot be made that the animals learned a temporal same-different concept. However, the more important result concerns the substantial differences in discrimination sensitivity engendered

by what might be construed as a minor change in procedure. In one case (Fetterman & Dreyfus, 1986, 1987), pigeons readily learned the task and appeared to learn the discrimination on the basis of a relational comparison, whereas in another (the present experiments), pigeons learned only with great difficulty and did not appear to acquire the concept. These results call for a closer analysis of different relational tasks and ways that the nature of the required judgment may interact with the dimension of comparison.

REFERENCES

- CARTER, D. E., & WERNER, T. J. (1978). Complex learning and information processing by pigeons: A critical analysis. *Journal of the Experimental Analysis of Behavior*, *29*, 565-601.
- CREELMAN, C. D., & MACMILLAN, N. A. (1979). Auditory phase and frequency discrimination: A comparison of nine procedures. *Journal of Experimental Psychology: Human Perception & Performance*, *5*, 146-156.
- D'AMATO, M. R., SALMON, D. P., & COLOMBO, M. (1985). Extent and limits of the matching concept in monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 35-51.
- DREYFUS, L. R., FETTERMAN, J. G., & STUBBS, D. A. (1986, April). *Duration comparison: Variation in duration range*. Paper presented at the meeting of the Eastern Psychological Association, New York, NY.
- FARELL, B. (1985). "Same"-"different" judgments: A review of current controversies in perceptual comparisons. *Psychological Bulletin*, *98*, 419-456.
- FETTERMAN, J. G., & DREYFUS, L. R. (1986). Pair comparison of durations. *Behavioural Processes*, *12*, 111-123.
- FETTERMAN, J. G., & DREYFUS, L. R. (1987). Duration comparison and the perception of time. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value* (pp. 3-27). Hillsdale, NJ: Erlbaum.
- GIBBON, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*, 279-325.
- GREEN, D. M., & SWETS, J. A. (1974). *Signal detection theory and psychophysics*. Huntington, NY: Krieger.
- GRIER, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, *75*, 424-429.
- 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.
- NOREEN, D. L. (1980). Optimal decision rules for some common psychophysical paradigms. *Proceedings of the Symposium in Applied Mathematics*, *13*, 237-279.
- PREMACK, D. (1978). On the abstractness of human concepts: Why it would be difficult to talk to a pigeon. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 423-451). Hillsdale, NJ: Erlbaum.
- PREMACK, D. (1983). Animal cognition. *Annual Review of Psychology*, *34*, 351-362.
- SANTIAGO, H. C., & WRIGHT, A. A. (1984). Pigeon memory: Same/different concept learning, serial probe recognition acquisition, and probe delay effects on the serial-position function. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 498-512.
- STEVENS, S. S., & GALANTER, E. H. (1957). Ratio scales for a dozen perceptual continua. *Journal of Experimental Psychology*, *54*, 377-411.
- STUBBS, D. A. (1976). Response bias and the discrimination of stimulus duration. *Journal of the Experimental Analysis of Behavior*, *25*, 243-250.

- STUBBS, D. A., DREYFUS, L. R., & FETTERMAN, J. G. (1987). *Discrimination of duration ratios*. Unpublished manuscript.
- VICKERS, D. (1979). *Decision processes in visual perception*. New York: Academic Press.
- WHITE, K. G. (1974). Temporal integration in the pigeon. *British Journal of Psychology*, **65**, 437-444.
- are the same or different. When prothetic stimuli are used, the former can be converted to the latter by requiring subjects to make the response "different" whenever the first stimulus is greater *or* less than the second (Vickers, 1979).

NOTE

1. There is, of course, a close relationship between three-category judgment tasks and tasks that require a decision as to whether two stimuli

(Manuscript received February 18, 1987;
revision accepted for publication July 29, 1987.)

NOTICE

Nominations for the Editorship of *Memory & Cognition*

Nominations are solicited for the editorship of *Memory & Cognition*. The term of the present editor, Alice F. Healy, expires at the end of 1989. The new editor will begin to receive manuscripts in 1989, with an official 4-year term to begin January 1, 1990. Nominations must be submitted by March 1, 1988.

Nominations (including self-nominations) should be sent to:

Richard M. Shiffrin
Chair, *Memory & Cognition* Search Committee
Department of Psychology
Indiana University
Bloomington, IN 47405