The start-stop procedure: Estimation of temporal intervals by human subjects

CHRIS N. KLADOPOULOS, BRUCE L. BROWN, and NANCY S. HEMMES

*Queens College, City University of New York, Flushing, New York
and The Graduate School and University Center, City University of New York, New York*

and

SOLEDAD CABEZA DE VACA

Queens College, City University of New York, Flushing, New York

In four experiments investigating human timing, subjects produced estimates of sample durations by bracketing their endpoints. On each trial, subjects reproduced a sample duration by pressing a button before the estimated sample duration elapsed (start time) and releasing it after the estimated duration elapsed (stop time). From these responses, middle time ($\text{start} + \text{stop}/2$) and spread time ($\text{stop} - \text{start}$) were calculated, representing the point of subjective equality and the difference limen, respectively. In all experiments, subjects produced middle times that varied directly with sample duration. In Experiment 2, middle times lengthened when feedback was withheld. Consistent with Weber timing, spread times, as well as the standard deviation of middle times, varied directly with middle time (Experiments 1, 3, and 4). On the basis of an internal clock model of timing (Gibbon & Church, 1990), the data permitted inferences regarding memory processes and response threshold. Correlations between start and stop times and between start and spread times agreed with earlier findings in animals suggesting that the variance of temporal estimates across trials is based in part upon the selection of a single temporal memory sample from a reference memory store and upon one or two threshold samples for initiating and terminating each estimate within a trial.

A new method for the investigation of human timing is described. The method, termed the start-stop procedure, is unique in providing estimates of the point of subjective equality (PSE) and difference limen (DL) on individual trials. It also yields two dependent measures on each trial that permit inferences about a subject's criterion and threshold for responding.

The start-stop procedure is a modification of the method of reproduction. Under both procedures, a subject is presented with a sample stimulus of a specified duration, followed by an estimation stimulus. Under the traditional method of reproduction, the subject's estimate consists of a single response that terminates the second stimulus when its duration is perceived to match the sample duration. In contrast, the start-stop procedure requires the subject to produce a temporal bracket about the estimated sample time by pressing (start response) and releasing (stop response) a button. The subject is instructed to make the start response prior to the moment at which the second

stimulus duration is judged to be equal to the sample duration, and to make the stop response when the duration of the second stimulus is judged to exceed that of the sample duration. Accordingly, each trial under the start-stop procedure yields two dependent measures: Start and stop times are the latencies of the start and stop responses, respectively, measured from the onset of the second stimulus. The duration of the produced temporal bracket, spread time ($\text{stop} - \text{start}$), is assumed to represent the DL, analogous to the standard deviation of temporal estimates across trials. The midpoint of the bracket (middle time) is assumed to represent the PSE, analogous to the mean of temporal estimates across trials.

According to the foregoing analogy, the Weber fraction can be defined as the spread-time/middle-time ratio, calculated from individual trial performance. The validity of this ratio was determined by comparing its performance with that of the traditional Weber fraction for time estimation. The traditional Weber fraction has been defined as a ratio of standard deviation of time estimates to mean time estimate; the corresponding ratio in the present case is standard deviation of middle time/mean middle time (coefficient of variation, CV, of middle time). Specifically, we wanted to determine if spread time/middle time would yield a constant ratio across durations in accordance with Weber's law.

A second advantage of obtaining start and stop times is their implications for specific timing mechanisms. Guided by a structural model of human timing, the internal clock

A portion of the results was reported at the annual meeting of the Eastern Psychological Association, April 1995, in Boston. We thank Lester E. Krueger for his invaluable contribution to the revision of the manuscript. We thank Philip H. Ramsey for assistance with statistical analyses. We gratefully acknowledge the contributions of John Zhu, for his development of the timing programs, and Carol Wong, for her contributions to various aspects of the research. Correspondence regarding this article can be sent to C. N. Kladopoulos, Department of Psychology, Queens College-CUNY, Flushing, NY 11367 (e-mail: kladcuny@aol.com).

Figure 1. An information-processing model (the internal clock) of timing (after Gibbon & Church, 1990).

model (Treisman, 1963), which has undergone substantial development in the animal timing domain (e.g., Church, 1984; Gibbon, 1991; Roberts, 1981), Gibbon and Church (1990) drew inferences based on the functional relations between start, stop, and spread times. Figure 1, based on Gibbon and Church (1990), shows an internal clock model that includes the following processes: pacemaker, switch, accumulator, comparator, and reference memory store. During a temporal event, the pacemaker emits pulses that are gated by the switch to an accumulator. The number of pulses in the accumulator (also represented in working memory) and a value represented in the reference memory store are continually monitored by a decision process (comparator) that controls the output of the system (response). Values in the memory store are assumed to consist of representations of previously presented critical durations.

According to the clock model, start and stop times are determined by the reference memory and comparator processes. Gibbon and Church (1984, 1990) argued that direction of the functional relations between start and stop times and start and spread times reflects the operation of memory and threshold mechanisms, respectively. In a given time-estimation task, it is assumed that the subject continuously compares current subjective elapsed time with the representation (a sample from reference memory) of the criterion time. As time elapses, the discrepancy between current time and remembered time diminishes; judgment of a match will occur when a discrepancy threshold is met. A match will be detected at some value of subjective elapsed time that falls short of remembered time. If the procedure arranges for timing to continue beyond

the point at which current time is equal to remembered time, judgment of a match persists with further increases in elapsed time until the magnitude of the discrepancy between current and remembered times again meets threshold. Thus, memory and comparator mechanisms, represented by discrepancy and threshold values, respectively, jointly define a temporal match region, outside of which times are judged as subjectively "too short" or "too long." The boundaries of the match region are taken to correspond to start and stop times.

Gibbon and Church's (1990) reasoning is represented graphically in Figure 2, in which elapsed time is arrayed along the abscissa and the discrepancy between current subjective time and remembered time is represented on the ordinate. In animal research, the model requires that the discrepancy be expressed relative to remembered time, that is $|t - s|/s$, where t is current subjective time and s is remembered time (but see Wearden, 1992). Therefore, discrepancy begins with a value of 1 when $t = 0$ and declines linearly to zero when $t = s$, after which it increases linearly. The horizontal line, B , represents threshold, which intersects with the relative discrepancy function to define start and stop times, as indicated by the vertical lines. The model readily accommodates Weber timing: as the criterion time increases, the vertex (middle time), S , of the function in Figure 2 shifts to the right, the slope of the discrepancy function flattens, and spread time (stop - start distance) increases correspondingly (for a fixed value for threshold). This feature is an essential component of scalar timing theory (Gibbon, 1977) that is incorporated in the model. Spread time is also influenced by threshold. For a fixed-criterion time, spread time varies directly with fluctuations in threshold level (B).

The model has proven to be diagnostic for the interpretation of trial-to-trial fluctuations in performance in timing tasks. Figure 2 is shown with a single value for memory and for threshold. Gibbon and Church (1990) examined the patterns of correlations among start, stop, spread, and middle times that emerge from the model when memory (s) and threshold (B) are permitted to vary randomly among trials, and then interpreted obtained correlations as reflecting specific properties of memory and threshold mechanisms. In the clock model, a positive start-stop correlation across trials is implied by the sampling of a single value on each trial from a distribution of values in reference memory (i.e., variation across trials in location of the vertex, S , in Figure 2), while a negative start-spread correlation is implied by variability across trials in threshold value (i.e., variation across trials in height of the threshold, B , in Figure 2).

Correlations among these values have been obtained from animal timing research employing the peak procedure—a signaled (discrete-trials) fixed-interval (FI) procedure (Catania, 1970; Roberts, 1981). On unreinforced test trials that extend well beyond the FI duration in this procedure, mean response rate over many trials typically varies as a smooth unimodal function of elapsed trial time, with a peak located near the FI value. In contrast, performance on individual trials is characterized by a break-run-break pattern—that is, a shift from a stable low rate to a higher

Figure 2. Subjective discrepancy ratio as a function of elapsed trial time for a relative comparison rule, based on Gibbon and Church (1990). t = current subjective time, s = remembered time, B = discrepancy threshold, S = middle time. Start and stop times are indicated by the vertical lines.

rate and a return to a lower rate. The transition points define the start and stop times respectively, and typically bracket the FI value. In their study, Gibbon and Church (1990) obtained a positive correlation between start and stop times and a negative correlation between start and spread times. This pattern of correlations has been replicated in other animal timing research (Cheng & Westwood, 1993; Cheng, Westwood, & Crystal, 1993; Church, Meck, & Gibbon, 1994).

The present study extends this analytical tool to timing in human subjects. In Experiments 1 and 2, we examined temporal estimation and the effect of feedback, respectively, under the start-stop procedure. In Experiment 3, temporal estimates over a range of durations were obtained, and the relationship of middle time to each of the two DL measures (the standard deviation of middle times and spread time) was examined. Of particular interest was stability of the corresponding Weber fractions. Experiment 4 was a systematic replication of Experiment 3 undertaken to determine the same relations under conditions of restricted feedback. Correlations among start, stop, and spread times were examined in each study.

EXPERIMENT 1

Method

Subjects

Twelve undergraduate psychology students (4 male and 8 female) participated to satisfy an academic requirement. No subject in Experiments 1–4 participated in more than one experiment.

Apparatus and Setting

A microcomputer (Gateway 286, 16 MHz) with a VGA color monitor was used to run experimental sessions and to collect data. A two-panel Microsoft mouse, with the left response button covered by black electrical tape, served as the manipulandum. No sources of temporal information were available in the room, and the subjects were asked to remove their watches prior to each session.

Procedure

Subjects were seated facing the computer monitor with the mouse located directly in front of the computer monitor. All trials consisted of the sequential presentation of the following events at the center of the monitor screen: an initial message, a green sample stimulus, a purple estimation stimulus, and feedback about performance.

A trial began with the words "Click the button when ready." One second following this response, the sample stimulus was presented for a given duration (sample duration). The sample stimulus was a green circle (1.5-cm in diameter) with a white border (.7 mm thick). After a 1-sec delay following termination of the sample stimulus, the estimation stimulus appeared and remained visible until the subject made a response. The estimation stimulus was a purple circle that was otherwise identical to the sample stimulus. The subjects were told that their task was to terminate the estimation stimulus when they estimated that it had been visible for an interval equal to the duration of the sample stimulus. They were instructed to terminate the estimation stimulus by pressing the response button when they estimated that the sample duration was about to elapse (start time) and releasing it when they estimated that the sample duration had elapsed (stop time). The subjects were told that the 2-point response should always bracket the termination time of the sample stimulus, and that they should create the smallest possible bracket without missing the sample duration. No statements were made to the subjects to prohibit them from engaging in chronometric counting or any other timing strategy.

The actual duration is: (sample duration)sec.

The estimated interval is: (start-stop)sec.

Click a button when ready.

Figure 3. Graphic and numeric feedback provided following a response under the start-stop procedure. For graphic feedback, the dotted line perpendicular to the time axis represents the sample duration in a given trial, and the hatched area parallel to the time axis represents the bracket produced by the subject, beginning at the start time and ending at the stop time. Sample durations and brackets were presented to subjects as solid red and green lines, respectively.

Release of the response button was followed by termination of the estimation stimulus and presentation of the message "Click the button when ready." During trials in which feedback was provided, graphic and numeric feedback were presented above the message, as shown schematically in Figure 3. The sample duration was indicated by a red vertical line (a dotted line in the figure), perpendicular to the horizontal time axis. A subject's estimated bracket was indicated by a green bracket (hatched area in Figure 3), beginning at the subject's start time and ending at the subject's stop time. Regardless of the sample duration, the axis was scaled so that zero time was represented at the left edge of the monitor screen and the sample duration was placed at the center of the screen. The numeric value of the sample duration and start and stop times were presented in the upper half of the screen. Numeric feedback was always presented in seconds, carried to two decimal places. Feedback and/or the message remained on the screen until the subject started the next trial.

Practice. The subjects were exposed to one practice session, during which the experimenter described the task and asked the subject to restate the description. The experimenter remained in the room until the subject had performed the task correctly (attempted to produce start and stop responses that bracketed the target duration). There were 80 practice trials arranged in four 20-trial blocks. One of the two durations used during the subsequent test sessions was presented on all trials of a given block, starting with a 2-sec block and alternating with a 4-sec block. Feedback was presented after each trial.

Testing. Each test session consisted of 10 warm-up trials, 5 with the 2-sec sample duration followed by 5 with the 4-sec sample duration, followed by eight 20-trial blocks of test trials. The probability of feedback in a trial was always 1.0. There were three types of test sessions. In the 2-sec session and the 4-sec session, trials consisted exclusively of one sample duration (2 or 4 sec, respectively). In the mixed-duration session, 2- and 4-sec-sample durations were presented in random order equally often in 20-trial blocks. The order in which test sessions were presented was counterbalanced across subjects in each of two complete Latin squares. The 12 subjects were randomly assigned to one of the two squares, and within each square, randomly assigned to one of the six possible sequences of test sessions.

Data Analysis

Two time estimates, start time and stop time, were directly obtained on each trial. From these two estimates, two additional measures were calculated for each trial: Middle time was defined as the mean of start and stop times, $(\text{start} + \text{stop})/2$, and spread time was defined as the difference between start and stop times ($\text{stop} - \text{start}$).

For each of the four dependent measures, means and standard deviations were calculated for each subject on each session and duration. For each duration, trials in which a subject produced start or stop times that differed from the session mean by more than 3 SDs were rejected and the session values were recalculated. The percentage of data lost for Experiments 1-4 was 1.4, 1.7, 2.5, and 2.0, respectively.

Preliminary analysis of each of the four dependent measures showed no main or interaction effects involving the session factor. Therefore, means were calculated across sessions for subsequent analyses. At each duration, CVs, spread-time/middle-time ratios, and correlations between dependent measures were calculated for each subject and session, then the mean CV, spread-time/middle-time ratio, and correlations were calculated across subjects and sessions.

Results

Table 1 contains means and standard deviations of start, stop, middle, and spread times for the 2-sec sessions, the 4-sec sessions, and the 2- and 4-sec trials of the mixed sessions. On each type of session, start and stop times bracketed the corresponding sample durations and middle times closely approximated the sample durations.

To determine the degree of temporal control under within- and between-session manipulations of sample duration, a two-way analysis of variance (ANOVA) of middle times was conducted. Factors were sample duration (2 and 4 sec) and session type (within- vs. between-session manipulation of sample duration). The main effect of duration was significant [$F(1,11) = 10,839.00, p < .001$]. For all subjects, means on 4-sec trials exceeded those on 2-sec trials. There was no main effect of session type [$F(1,11) = 1.30, p > .20$]; nor was there an interaction between session type and sample duration [$F(1,11) = 3.19, p > .10$].

Variation in DL across conditions was evaluated in terms of standard deviation of middle times and spread times. Differences in variability among trial types were evaluated on the entire data set with Pitman's (1939) test for homogeneity of dependent variances ($\alpha = .05$). For 11 of 12 subjects, variances during the 4-sec sessions were significantly greater than those on the 2-sec sessions in middle times. During mixed sessions, 7 of the 12 subjects produced significantly more variability in the 4-sec trials than in the 2-sec trials in middle times. In no case was variance greater on 2-sec trials than on 4-sec trials.

Spread times were examined with a two-way ANOVA, with sample duration and session type as factors. The main effect of duration was significant [$F(1,11) = 15.93, p < .01$].

Table 1
Means and Standard Deviations of Start, Stop, Middle, and Spread Times (in Seconds) for the Single-Sample Duration Sessions (2 or 4 sec) and for the Mixed-Sample Durations Session (2- and 4-sec Trials) of Experiment 1

Sample Duration	Start		Stop		Middle		Spread	
	M	SD	M	SD	M	SD	M	SD
Single Duration								
2 sec	1.775	.209	2.184	.219	1.980	.186	.408	.212
4 sec	3.728	.299	4.208	.304	3.968	.286	.479	.187
Mixed Durations								
2 sec	1.773	.256	2.217	.240	1.995	.229	.444	.188
4 sec	3.674	.301	4.179	.319	3.927	.291	.504	.217

Table 2
Coefficients of Variation of Middle Time (CV) and Spread-Time/Middle-Time Ratios for the Single-Sample Duration Sessions (2 or 4 sec) and for the Mixed-Sample Durations Session (2- and 4-sec Trials) of Experiment 1

Measures	Single Duration		Mixed Durations	
	2 sec	4 sec	2 sec	4 sec
CV	.094	.072	.115	.074
Spread/middle	.207	.121	.223	.130

as spread times increased with sample duration (Table 1). Neither the main effect of session type [$F(1,11) < 1$] nor the session type $\times$ duration interaction [$F(1,11) < 1$] were significant.

Group mean coefficients of variation and the spread-time/middle-time ratios for each session type are presented in Table 2. Both ratio measures were consistently larger on 2-sec trials than on 4-sec trials. In two-way ANOVAs of sample duration and session type, both ratio measures were significantly larger on 2-sec trials than on 4-sec trials [CV, $F(1,11) = 14.36$, $p < .01$; spread/middle ratios, $F(1,11) = 9.36$, $p < .05$]. There were no significant main effects of session type for either CV [$F(1,11) < 1$] or spread-time to middle-time ratios [$F(1,11) < 1$]. There was no significant session type $\times$ sample duration interactions for CV [$F(1,11) = 1.22$, $p > .20$] or spread-time/middle-time ratios [$F(1,11) < 1$].

Inferences regarding memory and threshold processes were tested by examining correlations among start, stop, and spread times. For each subject, correlations across trials were obtained for each duration in each session. Group mean correlations, under each session type and sample duration are presented in Table 3. For each duration, there was a positive correlation between start and stop times and a negative correlation between start and spread times. All 12 subjects showed a positive start–stop correlation for each duration of the two session types. The start–spread correlation was negative in the 2- and 4-sec sessions for 11 subjects and in the 2- and 4-sec trials of the mixed sessions for 9 and 10 subjects, respectively. Although Table 3 suggests an effect of sample duration on magnitude of correlations, statistical analyses failed to indicate significant effects. Sample duration $\times$ session type ANOVAs of start–stop and start–spread correlation coefficients were conducted. The analyses yielded no significant main effect of sample duration for either start–stop [$F(1,11) = 3.62$, $p > .05$] or start–spread correlations [$F(1,11) < 1$] and no significant main effect of session type on start–stop [$F(1,11) = 3.87$, $p > .05$] or start–spread correlations [$F(1,11) < 1$]. The session type $\times$ duration interactions were not significant for either start–stop or start–spread correlations [$F(1,11) = 3.36$, $p > .05$, and $F(1,11) = 2.12$, $p > .10$, respectively].

Discussion

Performance on the start–stop procedure demonstrated temporal control by sample duration as middle times varied directly with sample duration on both within- and between-

session manipulations. Consistent with Weber's law, both estimates of the DL (standard deviation and spread time) varied directly with mean middle time. However, both the CV and spread/middle-time ratio varied with sample duration. This finding was further investigated in Experiments 3 and 4 by presenting subjects with a larger range of sample durations.

The positive correlations between start and stop times and the negative correlations between start and spread times obtained in the present study with human subjects are similar to those reported for animal subjects under the peak procedure (Cheng & Westwood, 1993; Cheng et al., 1993; Gibbon & Church, 1990), suggesting functional similarity between the start and stop measures across the two procedures.

EXPERIMENT 2

In Experiment 1, all trials were accompanied by feedback indicating accuracy of the subjects' estimates. Experiment 2 was conducted to assess the role of feedback on temporal estimates under the start–stop procedure, and to determine the generality of the Weber-like data obtained in Experiment 1. Following initial practice with feedback, subjects were exposed to sessions in which the probability of feedback following a response was 0, .5, or 1, while sample duration was held constant.

Method

The apparatus, setting, and training procedures were identical to those used in Experiment 1. Twelve undergraduate students (4 male and 8 female) were exposed to one practice and three test sessions in which the sample duration was always 2 sec. During the 80-trial practice session and the 10 warm-up trials of the test sessions, the probability of feedback was 1.0. The probability of feedback on the 160 test trials of each test session varied among values of 0.0, 0.5, or 1.0 across sessions. The order in which these session types were presented was counterbalanced across subjects using the Latin square design used in Experiment 1. Subjects were randomly assigned (2 subjects each) to one of six possible sequences of test sessions.

Results

Means and standard deviations of start, stop, middle, and spread times for each session type are presented in Table 4. While start and stop times bracketed the 2-sec sample duration under all conditions, the absence of feedback resulted in longer start, stop, and middle times than either of the two feedback conditions. An ANOVA yielded a significant effect of session type on middle times [$F(2,22) =$

Table 3
Start–Stop and Start–Spread Correlations for the Single-Sample Duration Sessions (2 or 4 sec) and for the Mixed-Sample Durations Session (2- and 4-sec Trials) of Experiment 1

Measures	Single Duration		Mixed Durations	
	2 sec	4 sec	2 sec	4 sec
Start–stop	.512	.808	.714	.756
Start–spread	–.455	–.287	–.448	–.276

Table 4
Means and Standard Deviations of Start, Stop,
Middle, and Spread Times (in Seconds) for the
Three Probabilities of Feedback of Experiment 2

Probability of Feedback	Start		Stop		Middle		Spread	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
.0	1.960	.421	2.504	.459	2.232	.418	.544	.273
.5	1.805	.283	2.258	.273	2.032	.252	.454	.235
1.0	1.806	.257	2.220	.234	2.013	.224	.414	.200

6.08, $p < .01$], and comparisons among individual means (Bonferroni procedure) showed significant differences between the no-feedback condition and each of the two feedback conditions.

In a test of the stability in performance during the no-feedback sessions, an ANOVA showed that middle times produced during the first half of the no-feedback sessions were not significantly different from middle times during the second half of the sessions [$F(1,11) < 1$]. Additional analyses indicated an absence of significant between-session effects of feedback. A 3×6 ANOVA with condition and sequence (session order) as factors showed no significant main effect of sequence [$F(5,6) = .32$] or significant condition $\times$ sequence interaction [$F(10,12) = .86$].

To determine whether changes in feedback probability affected the DL measures, time judgments were examined in terms of spread times and the variance of middle times. Table 4 shows an inverse relation between both measures and the probability of feedback. An ANOVA indicated a significant effect of feedback on spread times [$F(2,22) = 3.81$, $p < .05$], and post hoc analysis (Bonferroni procedure) yielded a significant difference between conditions in which the probability of feedback was 0.0 and 1.0. Differences in variance of middle times among trial types were evaluated with Pitman's test. Variances were significantly greater during sessions in which the probability of

feedback was 0 than when it was .5 or 1.0 for a majority of subjects (8 and 11 of 12, respectively).

Mean coefficients of variation for conditions 0.0, 0.5, and 1.0 were .187, .124, and .111, respectively, and spread-time/middle-time ratios were .245, .225, and .208, respectively. Both measures varied inversely with the probability of feedback; however, one-way ANOVAs showed no significant effects of feedback probability on CVs [$F(2,22) = 3.25$, $p = .058$] or spread-time/middle-time ratios [$F(2,22) = 1.55$, $p > .20$].

Table 5 (top two rows) shows start-stop and start-spread correlations for the three session types. In each case, there was a positive correlation between start and stop times and a negative correlation between start and spread times. One-way ANOVAs indicated no significant differences among conditions (0.0, 0.5, and 1.0) in either start-stop [$F(2,22) < 1$] or start-spread [$F(2,22) < 1$] correlations. In all three conditions, the start-stop correlations were positive for all 12 subjects and the start-spread correlations were negative for 11, 10, and 12 subjects in conditions 0.0, 0.5, and 1.0, respectively.

Discussion

Intermittent and continuous feedback generated a close correspondence between estimated time and sample duration. Temporal estimates were longer when the probability of feedback was 0 than they were under the two feedback conditions. The increase in middle time under the no-feedback condition was accompanied by an increase in spread time and in variance of middle times. Although the data appear to be consistent with the mediation of feedback effects by a Weber-like process, it is implausible that feedback affected the subjective representation of sample duration. Rather, it is more likely that differences in middle time and variance reflect the shaping of start and stop response times on which feedback was con-

Table 5
Start-Stop and Start-Spread Correlations for the Three Probabilities of
Feedback of Experiment 2, the Five Sample Durations of Experiment 3,
and the Five Sample Durations for the Two Groups of Experiment 4

		<i>p</i> (Feedback) for Each Session				
Experiment	Measures	.0	.5	1.0		
2	Start-stop	.810	.645	.670		
	Start-spread	-.180	-.457	-.500		
Sample Durations						
		2 sec	6 sec	10 sec	14 sec	18 sec
Subjects Trained With All Durations						
3	Start-stop	.653	.731	.771	.849	.823
	Start-spread	-.303	-.252	-.181	-.161	-.222
Subjects Trained With the 2-sec Duration						
4	Start-stop	.551	.757	.836	.901	.890
	Start-spread	-.086	-.084	-.320	-.109	-.193
Subjects Trained With the 18-sec Duration						
4	Start-stop	.250	.671	.812	.861	.853
	Start-spread	-.354	-.200	-.316	-.228	-.130

Figure 4. Mean middle time as a function of sample duration for Experiment 3 (panel A) and the 2- and 18-sec groups of Experiment 4 (panel B).

tingent. In addition, both CV and spread-time/middle-time ratios appeared to vary with feedback, although the effects were not statistically significant.

Correlations between start and stop times and between start and spread times were similar across conditions. If these correlations reflect memory and threshold processes, respectively (Cheng & Westwood, 1993; Cheng et al., 1993; Gibbon & Church, 1990), the effect of feedback on these processes was minimal. Whereas middle and spread times varied inversely with the probability of feedback, start-stop correlations were positive under all conditions and start-spread correlations were predominantly negative. Therefore, under the present conditions, feedback does not appear to have affected the process by which representations of temporal values were selected from memory or the process by which the comparator functioned to determine the duration of the estimated bracket.

EXPERIMENT 3

The generality of the results of Experiments 1 and 2 was studied in Experiment 3 by exposing individual subjects to five sample durations, ranging from 2 to 18 sec. The availability of a larger range of temporal estimates permitted a more sensitive assessment of Weber timing properties and examination of the psychophysical relation between sample duration and estimated time.

Method

The apparatus, setting, and trial structure were identical to those used in the first two experiments. Fifteen undergraduate psychology students (5 male and 10 female) were exposed to 2-, 6-, 10-, 14-, and 18-sec sample durations in one training and three test sessions. During the training session, the subjects were required to meet a performance criterion starting with the 2-sec sample duration and proceeding to each of the other sample durations in order of increasing duration. For each duration, training continued until a subject's start and stop times bracketed the target duration on four of five consecutive trials. A bracket was defined as the start time less than or equal

to the sample duration and the stop time greater than or equal to the sample duration in a given trial. All subjects met the criterion for all five durations within 30 min of instruction.

Each test session consisted of 5 warm-up trials with feedback in which each of the five sample durations was presented once in random order, followed by 60 test trials arranged in five 12-trial blocks. One of the five sample durations was presented on every trial in a given block. Feedback followed responding on the first trial only of each block. That trial and the following trial in each block were excluded from analysis. The order of presentation of sample-duration blocks in a given session was counterbalanced across subjects with incomplete Latin squares. The 15 subjects were randomly assigned to one of three Latin squares ($n = 5$ per square), and within each square, randomly assigned to one of five orders. A different set of Latin squares was generated for each session.

Results

Table 6 (top five rows) contains means and standard deviations of each dependent measure for the five sample durations. Start and stop times bracketed the five corresponding sample durations such that middle times closely approximated the durations, as shown in Figure 4A. A linear regression analysis of middle times against sample duration yielded a proportion of explained variance (r^2) greater than .999, a slope of 1.002 sec/sec, and an intercept ($-.048$ sec) that did not differ significantly from zero.

Weber properties were examined by obtaining the functional relations between DL and PSE measures and determining whether CV and spread-time/middle-time ratios were constant across middle time at each sample duration. The relations of the standard deviation of middle times and mean spread time to mean middle time are shown in Figures 5A and 5B, respectively. Linear regression lines are presented for each function. Both measures increased with middle time. Stepwise regression analyses showed a significant linear component and no significant quadratic component for standard deviation (slope = .051 sec/sec, intercept = .102 sec, and $r^2 = .987$) and spread time (slope = .054 sec/sec, intercept = .64 sec, and $r^2 = .960$) functions. The slopes were significantly greater than zero for both

Table 6
Means and Standard Deviations of Start, Stop, Middle,
and Spread Times (in Seconds) for Experiments 3 and 4

Sample Duration	Start		Stop		Middle		Spread	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experiment 3 (Feedback for All Durations)								
2 sec	1.644	.248	2.337	.255	1.990	.223	.693	.233
6 sec	5.502	.446	6.463	.440	5.982	.397	.960	.396
10 sec	9.277	.711	10.561	.607	9.919	.560	1.285	.697
14 sec	13.212	.975	14.635	.904	13.924	.863	1.423	.720
18 sec	17.291	1.090	18.844	1.096	18.068	1.012	1.553	.828
Experiment 4 (Feedback for the 2-sec Duration)								
2 sec	1.490	.338	2.567	.435	2.028	.332	1.077	.410
6 sec	5.121	.634	6.634	.677	5.877	.578	1.513	.617
10 sec	9.170	1.143	11.075	1.223	10.122	1.080	1.906	.968
14 sec	13.049	1.876	15.028	1.656	14.038	1.649	1.979	1.282
18 sec	16.485	1.875	18.754	1.707	17.620	1.657	2.268	1.370
Experiment 4 (Feedback for the 18-sec Duration)								
2 sec	1.046	.450	2.222	.406	1.634	.336	1.176	.553
6 sec	4.863	.699	6.576	.723	5.720	.609	1.714	.736
10 sec	8.976	1.048	10.987	.893	9.982	.900	2.012	.740
14 sec	12.822	1.106	15.128	1.067	13.975	1.039	2.306	.638
18 sec	16.654	1.572	18.991	1.542	17.822	1.477	2.336	.988

functions. The intercept was significantly greater than zero for the spread but not for the standard deviation function.

The similarity in the slopes of the difference limen functions suggests a close correspondence between spread time and the standard deviation of middle time. To determine the degree of covariation across sample duration between the two DL measures, a correlation between group mean spread time and the group mean standard deviation of middle times was calculated. There was a significant positive correlation between the two measures [$r(3) = .965, p < .01$]. The range of correlations across individual subjects was .693 to .995, and the relation was significant for 11 of the 15 subjects.

Mean CVs and spread-time/middle-time ratios are presented for each duration in Figures 6A and 6B (open circles), respectively. Both CV and spread-time/middle-time ratios showed a gradual decrease with larger sample durations. One-way ANOVAs showed a significant effect of sample duration for both CV [$F(4,56) = 40.63, p < .001$] and spread-time/middle-time ratios [$F(4,56) = 192.81, p < .001$]. Trend analyses yielded significant linear and nonlinear components for the CV measure [$F(1,14) = 57.44, p < .001$, and $F(1,14) = 72.84, p < .01$, respectively] and for the spread-time/middle-time ratio [$F(1,14) = 262.24, p < .001$, and $F(1,14) = 256.50, p < .001$, respectively].

Mean start-stop and start-spread correlations for each of the five sample durations are shown in Table 5, rows 3 and 4. There was a significant positive correlation between start and stop times at each duration [minimum $t(14) = 15.27, p < .001$] and a significant negative correlation between start and spread times [minimum $t(14) = 3.53, p < .01$]. An ANOVA indicated that start-stop correlations increased with sample-duration length [$F(4,56) = 5.28, p < .001$]. There was no significant effect of sample duration on start-spread correlations [$F(4,56) < 1$].

Discussion

Mean middle time varied directly with sample duration, indicating that subjects displayed temporal control when estimating intervals between 2 and 18 sec under the start-stop procedure. Furthermore, the linear relation between the spread and middle time and between the standard deviation of middle time and middle time is consonant with Weber timing.

However, a strict Weber rule does not anticipate an inverse relation between the DL/PSE ratios and sample duration, as was observed for both the CV and spread-time/middle-time ratios in the present experiment (Figure 6) and in Experiment 1. The lack of constancy in these ratios may be accommodated by the generalized form of Weber's law,

$$\Delta T = cT + k, \quad (1)$$

which provides a representation of the relations between the two measures of ΔT (the standard deviation of middle times and spread time) and middle time, T (Allan, 1979; Ekman, 1959; Treisman, 1963). According to the modified equation, the inverse relations between the two DL/PSE ratios and sample duration may arise from the nonzero intercepts in the linear relations between the standard deviation of middle times and middle time and between spread time and middle time. This can be observed by dividing both sides of Equation 1 by T , thereby converting it to a relation of the Weber fraction to middle time:

$$\Delta T/T = c + k/T. \quad (2)$$

The constants c and k represent the obtained slopes and intercepts in the DL functions. Equation 2 indicates a nonlinear decrease in the Weber fraction to c as a function of T .

The dotted lines in Figure 6 are based on Equations 1 and 2, substituting slope and intercept values (c and k , respec-

Figure 5. Standard deviations (panels A and C) and mean spread times (panels B and D) as a function of mean middle time for Experiment 3 (panels A and B) and the 2- and 18-sec groups of Experiment 4 (panels C and D). Lines of best fit are linear regressions for all data sets.

tively) taken from corresponding linear regression analyses of the difference limen functions in Figures 5A and 5B. The predicted ratios at each duration closely approximate the obtained ratios (open circles). Although the intercept of the standard deviation function was not significantly different from zero, when used as a measure of the intercept in the generalized equation above, it yielded an inverse relation between predicted CVs and sample duration (Figure 6A).

The increase in measures of the DL with sample duration is consistent with Weber timing. In terms of the clock model, the relation can also be understood as reflecting an increase in the variance of stored values in reference memory with an increase in the duration of the estimated interval.

As in Experiments 1 and 2, positive start–stop correlations and negative start–spread correlations were obtained at each duration. This pattern of correlations rules out several alternative modes of memory and comparator functions. Gibbon and Church (1990) considered three possible sampling modes for the manner in which two sources of clock variability (memory and threshold) could determine start and stop times. At one extreme, memory or threshold values remain constant from trial to trial (no sampling). A more plausible situation is variation across trials in each value (sampling), but sampling of a single value of mem-

ory or threshold within a trial for determination of both start time and stop time. The third alternative envisions two samples of memory or threshold values in a trial, providing for the independent control of start time and stop time within a trial. The application of these three alternatives (none, one, or two samples) to each source of variability yields nine possible combinations. The pattern of correlations obtained by Gibbon and Church (1990) from pigeon subjects (positive start–stop correlations, negative start–spread correlations) was consistent with only two of these, both of which require sampling of a single reference memory value from a variable set of stored values. The correlations implied variability in threshold across trials, but did not distinguish between sampling of one or two values within a trial. The similarity of the present data with data from animal studies (e.g., Cheng & Westwood, 1993; Gibbon & Church, 1990) supports the operation of similar clock mechanisms.

The clock model does not necessarily predict the obtained increase in magnitude of start–stop correlations with sample–stimulus duration, but it is not inconsistent with it. According to the model, as sample duration increases, variability in both start and stop times increases concomitantly. It may be assumed that variability in start and stop times also derives in part from sources that are in-

Figure 6. Coefficients of variation (panel A) and mean spread-time/middle-time ratios (panel B) as a function of sample duration for Experiments 3 and 4. Obtained ratios are indicated by symbols, and predicted ratios by dotted lines.

dependent of clock mechanisms (e.g., motor or performance factors) and are constant across sample durations. In that case, the increase in the magnitude of the start-stop correlation across sample durations reflects the diminishing relative contribution of nontemporal (additive) sources of variability to start and stop times.

The relation between start-stop correlation and sample duration is not consistent with the results of Church et al. (1994), who manipulated FI value (15, 30, and 60 sec) in the peak procedure across different groups of pigeons. While the pattern of correlations among start, stop, and spread times within groups was similar to those obtained in the present study, no relationship was found between start-stop correlation size and FI duration. In addition to experimental design, species, and task, differences between studies in the range of stimulus durations employed may contribute to the discrepant findings. Thus, additive sources of variability may have been more substantial at the shorter durations in the start-stop procedure than in the peak procedure, exerting a greater influence on start-stop correlation size across sample duration. However, the same process would also entail an increase in the absolute value of the start-spread correlations, which did not vary significantly with sample duration.

EXPERIMENT 4

Experiment 4 addressed the possibility that the results of Experiment 3 were influenced by the presentation of feedback. Evidence of feedback effects was found in Experiment 2, in which time estimates lengthened when feedback was omitted. Sample durations in Experiment 4 were the same as those in Experiment 3, but feedback was provided for estimates of only one sample duration. For different groups of subjects, feedback was provided on 33% of 2- or 18-sec trials.

Method

The apparatus and setting were identical to those used in the first three experiments. The subjects (4 male and 6 female undergraduate psychology students) were exposed to five sample durations, but were trained and received feedback with only one sample duration. According to their order of participation, the subjects were paired and, within a pair, were randomly assigned to a 2- and an 18-sec training condition (Groups 2 and 18, respectively, $n = 5$ per group).

Each subject received one training and two test sessions; the first test session immediately followed the completion of training. Training proceeded until the subjects' start and stop times bracketed the sample duration (2 or 18 sec) in four of five consecutive trials. All subjects met this criterion within 15 min of instruction. Before the first test session, the subjects were told that, in addition to the trained duration, they would be presented with other sample durations and that feedback would follow only some of their responses.

Test sessions began with a 5-trial warm-up consisting of one presentation of each of the five sample durations in random order with feedback provided only when a subject's training duration was presented. The remainder of the session consisted of 55 trials arranged in five 11-trial blocks. On the first trial of each block, the training duration was presented and responding was followed by feedback. On the remaining 10 trials, one of the five sample durations was presented without feedback. Feedback trials were excluded from analysis. The same incomplete Latin-square arrangement used in Experiment 3 was used to counterbalance the order of presentation of sample durations across subjects. The Latin squares for the two groups were identical within each session, and pairs of subjects were assigned randomly to the same row of a Latin square.

Results

Means and standard deviations of each dependent measure for the five sample durations are presented for each group (feedback for 2- or 18-sec sample durations) in Table 6 (Experiment 4, bottom 10 rows). For each group, start and stop times bracketed the corresponding sample durations and middle times closely approximated the sample durations. A group (2 vs. 18 sec) $\times$ duration ANOVA of middle times showed a significant main effect of dura-

tion [$F(4,32) = 3,263.32, p < .001$], no significant main effect of group [$F(1,8) < 1$], and no significant duration $\times$ group interaction [$F(4,32) < 1$].

Figure 4B illustrates mean middle time as a function of sample duration for Groups 2 and 18. Regression analyses indicated linear relations between middle time and sample duration in both cases ($r^2s > .999$). The slopes were .984 and 1.016 sec/sec for Groups 2 and 18, respectively. In both analyses, the intercepts (.101 and -.331 sec, respectively) did not differ significantly from zero.

Standard deviation of middle times and mean spread time are plotted as a function of mean middle time in Figures 5C and 5D, respectively. According to stepwise-regression analyses, there were significant linear components for each function and a significant quadratic component for the spread-time function of Group 18. In the figure, all functions are fit with the linear component.

Separate regression analyses on the standard deviation functions (Figure 5C) yielded slope, intercept, and r^2 values of .095 sec/sec, .115 sec, and .954 for Group 2 and .067 sec/sec, .217 sec, and .976 for Group 18. In both cases, slope and intercept values differed significantly from zero. Regression analyses of the spread-time functions (Figure 5D) yielded slope, intercept, and r^2 values of .073 sec/sec, 1.028 sec, and .953 for Group 2, and .072 sec/sec, 1.201 sec, and .924 for the linear component for Group 18. The slope and intercept values were significantly different from zero. For Group 18, the added quadratic component accounted for an increase in the proportion of variance explained from .924 to .995.

Correlations across durations were obtained between mean standard deviations of middle times and mean spread times for each group. As in Experiment 3, a significant positive correlation between the two DL measures was obtained for both Group 2 [$r(3) = .94, p < .05$] and Group 18 [$r(3) = .92, p < .05$].

CVs and spread-time/middle-time ratios are presented as a function of sample duration in Figure 6 (filled circles). Group $\times$ sample duration ANOVAs of each obtained ratio indicated a significant effect of duration for both CV [$F(4,32) = 8.04, p < .001$] and spread-time/middle-time ratios [$F(4,32) = 39.23, p < .001$]. There was no significant effect of group for CV [$F(1,8) < 1$] or spread-time/middle-time ratios [$F(1,8) < 1$]. The group $\times$ duration interaction was not significant for either CV [$F(4,32) < 1$] or spread-time/middle-time ratios [$F(4,32) = 1.72, p > .10$]; therefore, data presented in Figure 6 were pooled across groups. For both functions, ratios decreased with larger sample durations.

Trend tests yielded significant linear and nonlinear components for both CV [$F(1,9) = 16.89, p < .01$, and $F(1,9) = 10.58, p < .05$, respectively] and spread-time/middle-time ratios [$F(1,9) = 40.80, p < .001$, and $F(1,9) = 63.35, p < .001$, respectively]. The dotted line fit to the data in Figure 6 was based on the slope and intercept constants obtained from the solution of the generalized Weber relation (Equations 1 and 2). As in Experiment 3, the predicted ra-

tios closely approximated the obtained ratios at most sample durations.

Group mean correlations between start and stop times and between start and spread times are presented in Table 5 for each duration (bottom four rows). For each group at each sample duration, there was a positive correlation between start and stop times and a negative correlation between start and spread times, replicating the results of Experiment 3. Separate t tests showed significant mean positive start-stop correlations for Group 2 [$t(4) = 7.52, p < .01$] and Group 18 [$t(4) = 4.82, p < .01$]. The mean negative start-spread correlations for each group were not significantly different from zero [$t(4) = 1.04, p > .05$, and $t(4) = 1.75, p > .05$, respectively], but all subjects showed negative start-spread correlations (sign test, $p < .001$). Group $\times$ duration ANOVAs were conducted for start-stop and start-spread correlations. Start-stop correlations increased with sample duration [$F(4,32) = 10.28, p < .001$]. The main effect of group was not significant [$F(1,8) = 1.20, p > .05$], nor was the group $\times$ duration interaction [$F(4,32) < 1$]. For start-spread correlations, there were no significant main effects of duration [$F(4,32) < 1$] or group [$F(1,8) < 1$] and no significant group $\times$ duration interaction [$F(4,32) < 1$]. For both correlations, results regarding effects of duration replicated those of Experiment 3.

Discussion

Unlike the procedure of Experiment 3, subjects in Experiment 4 received feedback for only one of the five sample durations. Therefore, their performance on the four other sample durations was not subject to direct shaping by feedback stimuli. Nevertheless, both groups showed linear relations between middle time and sample duration, linear relations between DL measures and mean middle time (5 of 6 analyses), and decreasing CVs and spread-time/middle-time ratios with longer sample durations. Each of these relations replicated the findings of Experiment 3.

While the results of Experiment 2 demonstrated a role of feedback in time judgments, the locus of the feedback effect has not been determined. Feedback may influence clock processes, such as pacemaker rate, or more peripheral mechanisms involving the shaping of response labels attached to subjective time. In either case, the results of Experiments 3 and 4 indicate that duration-specific feedback is not responsible for the effect of duration upon behavior in the present procedure.

GENERAL DISCUSSION

The start-stop procedure yielded four major findings: (1) Human subjects produced time judgments that varied as a linear function of sample duration over a range of 2–18 sec (Experiments 3 and 4); (2) two measures of the DL increased with the PSE in a manner consistent with the generalized form of Weber's law (Experiments 3 and 4); (3) correlations among start, stop, and spread times were

in accord with those obtained from research with animals, and were interpretable in terms of an internal clock model of timing (Experiments 1–4); and (4) feedback influenced the magnitude and DL of estimates (Experiment 2). The present procedure differs from the methods of reproduction in human timing and the peak procedure in animal timing research. Unlike the former, it yields new measures, start and stop times, on individual trials and, unlike the latter, it provides direct measures of those variables.

In each experiment, positive correlations between start and stop times and negative correlations between start and spread times agreed with earlier findings in animals, suggesting that the variance of temporal estimates across trials is based in part upon the selection of a single temporal memory sample from a reference memory store (Cheng & Westwood, 1993; Cheng et al., 1993; Gibbon & Church, 1990) and one or two threshold samples for initiating and terminating each estimate within a trial (Church et al., 1994; Gibbon & Church, 1990). Moreover, the dissociation of start–stop and start–spread correlations under manipulation of sample duration (Experiments 3 and 4) lends indirect support to a two-threshold sampling model according to which start and stop times are governed by separate response criteria.

Further support for a two-threshold model was obtained from an analytic procedure developed by Church et al. (1994) that yields a test for one versus two thresholds. The test is based on the relation between spread to middle correlations and the ratios of CV for middle time to CV for spread time. Church et al. showed that these two measures would be approximately equal under a one-threshold model; however, if two thresholds were used, the ratio of

coefficients of variation would exceed the spread to middle correlation. Figure 7 shows this relation for the individual subjects of each experiment. Each dependent measure is based on a mean for each subject across durations (or conditions). The diagonal line indicates equality between the two measures (predicted by a one-threshold model), while points below the diagonal are consistent with a two-threshold model. All but 2 of the 49 values lie below the diagonal. The results of this analysis are similar to those reported by Church et al. for rats exposed to FI 15, 30, or 60 sec under the peak procedure, and are in line with previous data favoring a two-threshold model for pigeons' performance under the peak procedure (Gibbon & Church, 1990).

Spread times provide an estimate of the DL in a single trial that is analogous to the standard deviation of time estimates across trials in conventional procedures. This interpretation was supported in Experiments 3 and 4 by the obtained positive correlations between the two measures of the DL (spread time and standard deviation of middle time) and by the observed linear relations between those measures and middle times required by Weber timing.

The obtained decrease in CVs and spread-time/middle-time ratios with sample duration is inconsistent with Weber timing, but is not uncommon for human subjects (Killeen & Weiss, 1987; Treisman, 1963; Wearden, 1991). Current theoretical accounts acknowledge that variance in temporal estimates may arise from more than one component of the timing processes, but that at least one component may yield variance that is proportional to elapsed duration (Gibbon, 1991; Killeen & Weiss, 1987). Wearden (1991) recently demonstrated that the inverse relation be-

Figure 7. Scatter plot of spread-time–middle-time correlation versus the ratio of coefficients of variation based on middle and spread times. Data points represent individual subjects in Experiments 1–4.

tween CV and sample duration is consistent with timing based on chronometric counting in which the intercount intervals are scaled according to Weber's law.

The present data can be accommodated by a model based on the generalized form of Weber's law (Equation 1) that includes a multiplicative and an additive source of variance. As previously described, all of the difference-limen functions in Experiments 3 and 4 of the present study had positive intercepts, and the Weber fraction corresponding to each was well accommodated by Equation 2 (Figure 6). The structural clock model provides a useful way to partition these sources of variance. The model operates according to a formal scalar timing rule (Gibbon, 1977). According to the scalar principle, subjective representations of time are multiplicative transformations of physical time, such that, for different physical durations, equal proportions of elapsed time are judged as subjectively equivalent. When applied to variable time estimates, the scalar transform generates the direct relation between mean and standard deviation of estimates, that is, Weber's law.

The scalar property can derive from variability in pacemaker, memory, and comparator mechanisms. Other sources of variability in the model, including the switch latency and possible Poisson properties of the pacemaker, do not contribute to the scalar transformation (Gibbon, 1991; Gibbon & Church, 1984). For example, the switch gates pulses from the pacemaker to the accumulator (Figure 1). It is assumed that there is variance across trials in the latency between sample stimulus onset and the closing of the switch and between sample offset and the opening of the switch. This source of variability is orthogonal to scalar sources and may contribute to the additive component of the present data.

Another contributor to the intercepts may inhere in response production features of the task that are independent of the cognitive timer. In the present case, this could include components of start and stop motor responses that are subsequent to the decision process. The requirement of multiple responses during the estimation stimulus would be expected to create more response variance across trials than the single-response estimates characteristic of conventional psychophysical timing procedures. The consequence of multiple responses may be evident in the intercepts of the spread-time to middle-time relations.

In summary, the start-stop procedure produced results in human subjects that parallel those in animal research, suggesting the operation of common mechanisms of timing across species. Models derived from animal research (e.g., Church & Broadbent, 1990; Gibbon, 1991; Killeen & Fetterman, 1988) may therefore be applicable to human

timing. While the present data do not rule out the mediation of timing by behavior that may be unique to human observers, such as counting (e.g., Wearden, 1991), they are accommodated by an account that does not require it under the conditions prevailing in the present study.

REFERENCES

- ALLAN, L. G. (1979). The perception of time. *Perception & Psychophysics*, **26**, 340-354.
- CATANIA, A. C. (1970). Reinforcement schedules and psychophysical judgements: A study of some temporal properties of behavior. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 1-42). New York: Appleton-Century-Crofts.
- CHENG, K., & WESTWOOD, R. (1993). Analysis of single trials in pigeons' timing performance. *Journal of Experimental Psychology: Animal Behavior Processes*, **19**, 56-67.
- CHENG, K., WESTWOOD, R., & CRYSTAL, J. D. (1993). Memory variance in the peak procedure of timing in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **19**, 68-76.
- CHURCH, R. M. (1984). Properties of the internal clock. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (Annals of the New York Academy of Sciences: Vol. 423, pp. 566-582). New York: New York Academy of Sciences.
- CHURCH, R. M., & BROADBENT, H. A. (1990). Alternative representations of time, number, and rate. *Cognition*, **37**, 55-81.
- CHURCH, R. M., MECK, W. H., & GIBBON, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 135-155.
- EKMAN, G. (1959). Weber's law and related functions. *Journal of Psychology*, **47**, 343-352.
- GIBBON, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, **84**, 279-325.
- GIBBON, J. (1991). Origins of scalar timing. *Learning & Motivation*, **22**, 3-38.
- GIBBON, J., & CHURCH, R. M. (1984). Sources of variance in an information processing theory of timing. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 465-488). Hillsdale, NJ: Erlbaum.
- GIBBON, J., & CHURCH, R. M. (1990). Representation of time. *Cognition*, **37**, 23-54.
- KILLEEN, P. R., & FETTERMAN, J. G. (1988). A behavioral theory of timing. *Psychological Review*, **95**, 274-295.
- KILLEEN, P. R., & WEISS, N. A. (1987). Optimal timing and the Weber function. *Psychological Review*, **94**, 455-468.
- PITMAN, E. J. G. (1939). A note on normal correlation. *Biometrika*, **39**, 9-12.
- ROBERTS, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes*, **7**, 242-268.
- TREISMAN, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychological Monographs*, **77** (13, Whole No. 576).
- WEARDEN, J. H. (1991). Do humans possess an internal clock with scalar timing properties? *Learning & Motivation*, **22**, 59-83.
- WEARDEN, J. H. (1992). Temporal generalization in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 134-144.

(Manuscript received April 17, 1995;
revision accepted for publication February 28, 1997.)