Psychonomic Bulletin & Review
2004, 11 (1), 24–30

THEORETICAL AND REVIEW ARTICLES

"One-thousand *one* . . . one-thousand *two* . . .": Chronometric counting violates the scalar property in interval timing

SEAN C. HINTON and STEPHEN M. RAO
Medical College of Wisconsin, Milwaukee, Wisconsin

Weber's law applied to interval timing is called the *scalar property*. A hallmark of timing in the seconds-to-minutes range, the scalar property is characterized by proportionality between the standard deviation of a response distribution and the duration being timed. In this temporal reproduction study, we assessed whether the scalar property was upheld when participants chronometrically counted three visually presented durations (8, 16, and 24 sec) as compared with explicitly timing durations without counting. Accuracy for timing and accuracy for counting were similar. However, whereas timing variability showed the scalar property, counting variability did not. Counting variability across intervals was accurately modeled by summing a random variable representing an individual count. A second experiment replicated the first and demonstrated that task differences were not due to presentation order or practice effects. The distinct psychophysical properties of counting and timing behaviors argue for greater attention to participant strategies in timing studies.

Interval timing allows organisms to predict environmental events that occur in the seconds-to-minutes range. A nearly universal feature of interval timing is the scalar property that is observed across durations. The scalar property refers to a proportional relationship between the standard deviation of the response distribution and the duration being timed. Any perceptual process that obeys Weber's law implies such a relationship. Scalar expectancy theory successfully accounts for this property by defining the sources of timing variability that contribute to performance (Gibbon, 1977). The primary sources of variability derive from clock, memory, and decision processes (Gibbon & Church, 1984; Gibbon, Church, & Meck, 1984)

The scalar property has been observed across a variety of species, including fish, birds, rodents, and humans (Church, Getty, & Lerner, 1976; Gibbon, 1977; Rakitin et al., 1998; Schneider, 1969; Talton, Higa, & Staddon, 1999), and it spans durations ranging from milliseconds to hours (Crystal, 2001; Ivry & Hazeltine, 1995). The scalar property is also evident in a wide range of nontemporal tasks that involve numerical representations, including number discrimination in animals (Mechner, 1958; Meck & Church, 1983; Platt & Johnson, 1971), nonverbal numerical representation in both animals and humans (Gallistel & Gelman, 1992), and nonverbal counting in humans (Cordes, Gelman, Gallistel, & Whalen, 2001; Whalen, Gallistel, & Gelman, 1999).

Humans readily use language to develop counting strategies, referred to as *chronometric counting*, to mark the passage of time (Gilliland & Martin, 1940). One might reasonably conclude that this form of verbal counting behavior also obeys the scalar property. Chronometric counting enables the individual to divide a temporal duration into rhythmic subintervals with verbal labels (e.g., "one-thousand *one* . . . one-thousand *two* . . ."). Interval timing studies in humans do not always control for chronometric counting, yet timing with or without counting may have different psychophysical properties. It is unknown whether chronometric counting obeys the scalar property.

Published reports regarding how counting variability changes with duration are conflictual. A mathematical treatment that emphasized the contribution of counter error concluded that "for long intervals, Weber's constant for time is shown to be the Weber constant for counting" (Killeen, 1991, p. 213). In addition, a number of authors have presented empirical counting data that generally conform to the scalar property, although their studies were not specifically designed to assess variability changes (Gilliland & Martin, 1940; Kruup, 1961; Spivack & Levine, 1964). A recent

The authors thank Rebecca Thompson and Jill Dorflinger for technical assistance, Raymond Hoffmann for statistical advice, and Deborah Harrington and Warren Meck for helpful comments on earlier versions of this manuscript. We also gratefully acknowledge the thoughtful suggestions of reviewers Russell Church, Randy Gallistel, and Peter Killeen. This study was supported by grants from the National Institutes of Health (P01 MH51358, R01 MH57836) to S.M.R. and MCW's General Clinical Research Center (M01 RR00058). Correspondence should be addressed to S. C. Hinton, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226 (e-mail: shinton@mcw.edu).

study examined variability explicitly and found that counting a 12-sec duration was more precise than timing the same duration (Rakitin et al., 1998). Although this finding indicates that response variability for counting and timing can be distinguished, variability was not assessed across multiple intervals, a necessity for examining the scalar property.

During a counting task, participants divide a duration into a series of smaller subintervals. The variability of these subintervals exhibits the scalar property over a 0.5- to 1.3-sec range, suggesting that participants are in fact timing the subintervals (Wearden & McShane, 1988). Several theorists have predicted reduced variability during counting based on the mathematical property that total variance consists of the sum of variances of individual counts (Fetterman & Killeen, 1990; Getty, 1976; Killeen & Weiss, 1987). In this circumstance, *variance* (rather than the standard deviation) should scale proportionally with duration. Such a relationship is equivalent to the variability scaling with the *square root* of the duration being timed (which we refer to as the *square-root model*, SRM). An experiment examining duration discrimination in the range of 0.7–1.9 sec generally bore out these predictions when participants counted fast or did not count (Grondin, Meilleur-Wells, & Lachance, 1999). However, a subsequent experiment within the same study produced contrary results: when participants were given more practice and tested with twice as many trials, the variability in the counting condition remained constant from 1.3 to 1.9 sec. This observation is inconsistent with counting following either scalar or square-root variability, and it led to the suggestion that "when human observers are allowed to keep track of time as efficiently as they can, Weber's law encounters violations" (Grondin, 2001, p. 32).

Given these inconclusive findings, we sought to determine empirically whether counting variability exhibits the scalar, the square-root, or another property. In the present study, we explicitly compared counting and timing of three durations within the same group of subjects using the peak-interval timing procedure, also known as the "peak procedure" (Catania, 1970; Roberts, 1981). In Experiment 1, all participants first timed and then counted each duration. In Experiment 2, the tasks were presented in the opposite order to a different group of participants. The purpose of Experiment 2 was to assess the reliability of the findings of Experiment 1 and to determine whether those results were significantly influenced by task order or practice effects.

METHOD

Experiment 1

Participants

Thirteen healthy adults (ages 24–46; 6 males) participated. They were recruited from students and staff at the Medical College of Wisconsin (MCW) and were compensated $12/h. All provided written informed consent consistent with MCW Institutional Review Board guidelines.

Apparatus

Data were collected on an IBM-compatible personal computer with a 33-MHz AMD 80386SX microprocessor. Stimuli were presented and responses recorded with millisecond precision using a custom DOS-based application written in Turbo Pascal. Timing of events was controlled by assembly language residual-count processing of the 8254 timing chip (Bovens & Brysbaert, 1990; Graves & Bradley, 1987). Accuracy was within approximately 15 msec as a result of both keyboard imprecision and software keyboard polling rate.

Design and Procedure

All participants were tested at the same time of day over two sessions separated by at least 1 day. Each testing session lasted approximately 1 h. Because of practical considerations (e.g., subject fatigue), the three target durations (8, 16, and 24 sec) were presented in two successive blocks of trials, counterbalanced across participants. On one testing day, either two blocks of the 8-sec duration preceded two blocks of the 16-sec duration or vice versa, and on the other testing day, two blocks of the 24-sec duration were presented. For each duration, timing always preceded counting, in order to discourage participants from associating a particular verbal label with the duration prior to timing. Each block consisted of 5 fixed-time training trials and 40 testing trials.

The participant sat approximately 80 cm from the computer monitor and pressed the Enter key to begin each training trial. After a random variable delay ranging from 0.5 to 1.5 sec in 100-msec increments, a blue rectangle 10.6 cm wide × 12.5 cm long appeared on the computer monitor, indicating the beginning of the interval to be timed. The participants were instructed to attend to the interval during which the rectangle appeared blue. When the target duration was reached, the blue rectangle changed color to magenta and remained on the monitor for 1.5 sec. At the end of the color change, the monitor went black and a continuation prompt appeared: "press *Enter* to begin the next trial."

After 5 training trials, the testing phase of 40 trials began. These trials used the same stimulus as that described for training, except that the color change marking the end of the interval did not occur. The participants were instructed to produce a continual range of responses such that the target duration was centered within that range. They were told to tap the space bar at a comfortable rate while responding and to make at least five responses per trial. The participants could terminate the trial by pressing the Enter key when they had finished responding.

During the intertrial interval, feedback was provided in the form of a histogram that displayed the distribution of the participant's responses in the previous trial on relative time and response scales (see Rakitin et al., 1998, Figure 2). The target duration was indicated on the *x*-axis by a green letter "T," to exclude information about its absolute duration. A total period of twice the target duration was displayed in the histogram. The histogram was presented in bins of 10% of the target duration, and the height of the bars represented the proportion of the maximum number of responses per bin. The region of the feedback graph within ±30% of the target duration had a green bar above it, indicating a range of good performance, whereas earlier and later regions had a red bar above them. The feedback was designed to encourage evenly distributed responses ranging symmetrically around the target duration. Pressing the Enter key a second time allowed the participant to continue to the next trial after a short, variable delay.

In a timing block, the participants were instructed not to count or use any other process of subdivision (such as foot tapping, humming, singing, etc.) nor to use any external timing mechanism. They were instructed to use their internalized perception of the temporal duration of the blue rectangle presented during the training trials to guide their performance during the testing trials. In the subsequent counting block, the participants were instructed to count silently during the training trials to determine a count corresponding to the duration that had elapsed when the color change occurred and to use that information to determine when to respond.

Data Analysis

We refer to a combination of a particular duration (8, 16, 24 sec) and task (counting, timing) as a *condition*. For each condition, all responses from all trials were pooled into a single frequency distribution with 0.5-sec bins to generate a peak function, which was normalized by maximum response rate to allow comparison across individuals. These normalized peak functions were fit with two-parameter (mean and standard deviation) Gaussian curves using the SPSS PeakFit program (v. 4.11, SYSTAT Software, 2001). A least-squares method that minimized the residuals determined the best-fitting Gaussian function. Two measures of interest were extracted from the fitted curves: the peak time or center of the distribution and the width of the distribution at half the maximum height (full width at half-maximum, FWHM). Center is a measure of timing accuracy, and FWHM is a measure of timing precision. A measure of error was derived by subtracting the expected value for each duration from the center. The FWHM of a Gaussian distribution is numerically equivalent to 2.355 standard deviations. A coefficient of variation (CV) was calculated by dividing the standard deviation (FWHM/2.355) by the center. CV is a proportionality measure of relative precision that remains constant across durations when the scalar property is upheld.

Error, FWHM, and CV were subjected to a 3 × 2 (duration × task) repeated measures analysis of variance (SPSS for Windows, release 8.0.0, 1997). The Wilks' lambda test within the multivariate repeated-measures analysis was used whenever there were significant departures from sphericity; otherwise univariate analyses were used. Planned comparisons using paired *t* tests (two-tailed significance levels) were performed between counting and timing at each duration for error, FWHM, and CV. Regression analyses were performed on FWHM to assess the ability of different models to characterize its change with duration. The CV was plotted by condition for each individual in log–log coordinates. Linear regression equations and 95% confidence intervals for the two tasks were calculated using SPSS.

Some theoretical accounts suggest that the counting data collected in the present study might be modeled as a sum of random variables (Fetterman & Killeen, 1990; Getty, 1976; Killeen & Weiss, 1987).

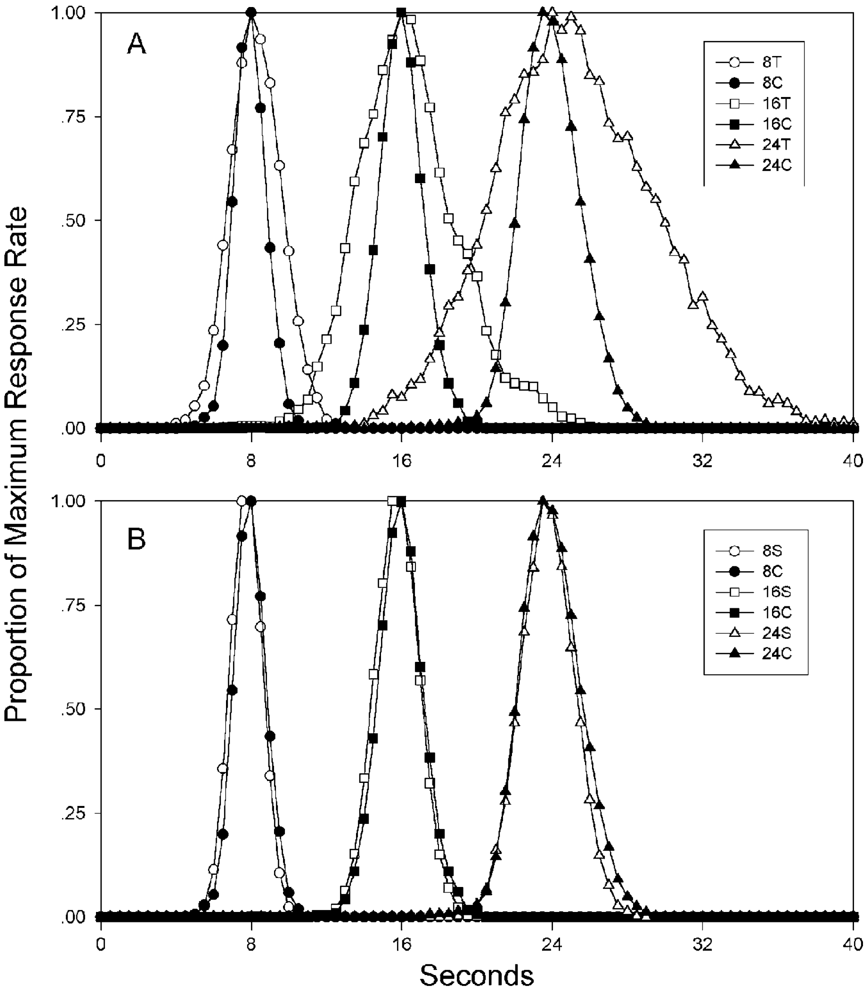

Figure 1. (A) Response distributions (peak functions) for Experiment 1 averaged across 13 subjects sorted by task (T, timing; C, counting) and target duration (8, 16, 24 sec) plotted in half-second bins on a relative response scale. Timing preceded counting for all conditions. (B) Peak functions for the Experiment 1 counting data (C) and a simulation (S) displayed by target duration (8, 16, 24 sec). See text for details of simulation.

The *rand* function in Excel was used to generate random numbers having a uniform distribution from 0 to 1 (Microsoft Excel 2000, v. 9.0.3821 SR-1). The observed data were fit well by a simulation in which four of these random numbers represented each second of the total duration (e.g., 32 draws for the 8-sec duration).[1] The sum of all the samples was divided by 2 to produce an individual "reproduction" trial, so that the sum corresponded approximately to the target duration. The simulation included 10,000 trials per condition.

Experiment 2

Experiment 2 was identical to Experiment 1, except that 8 healthy adults (ages 21–41; 4 males) participated, and counting preceded timing for each duration. Planned comparisons between Experiment 1 and Experiment 2 were performed using unpaired *t* tests (two-tailed significance levels) on error, FWHM, and CV for each condition.

RESULTS

Experiment 1

The normalized response distributions averaged across participants for each condition are displayed in Figure 1A. Each condition for each participant was modeled by a Gaussian distribution (goodness of fit was greater than .93 for all conditions), and the results are reported in terms of measures derived from these fitted functions.

Error (the deviation between the observed *center* and the target value) did not differ between timing and counting for 8 or 16 sec (Figure 2A). However, counting was significantly more accurate at the longest duration [$t(12) = 2.50, p < 0.03$; $F(2,11) = 4.64, p < .04$, for the interaction between task and duration]. Although accuracy was similar between the two tasks, precision was quite different, and counting produced significantly narrower distributions (FWHM) at all durations (Figure 2B) [$F(2,11) = 102, p < .001$].

The FWHM data for timing were fit with a linear model with a negative intercept, although fixing the intercept at 0 fit equally well (FWHM $= 0.41 * T - 0.42$ and FWHM $= 0.39 * T$; $r^2 = .73$ for both). The FWHM data for counting were well characterized both by the SRM (FWHM $= 0.67/T^2$, $r^2 = .55$) and by a linear model with a constant intercept (FWHM $= 0.099 * T + 1.03$, $r^2 = .57$).

The CVs were constant for the timing task (the scalar property), but they decreased with increasing duration for the counting task (Figure 3). The best-fitting linear regression to the log–log plot for timing was CV $= 0.00 * T - 0.81$ ($r^2 = .00$, slope parameter 95% confidence intervals -0.15 to 0.14). For counting, it was CV $= -0.40 * T - 0.69$ ($r^2 = .40$, slope parameter 95% confidence intervals -0.23 to -0.56). With the slope held constant at -0.5, the best-fitting equation was CV $= -0.5 * T - 0.58$ ($r^2 = .37$).

The simulation of counting closely modeled the observed counting data (Figure 1B).

Experiment 2

Experiment 1 and Experiment 2 did not differ significantly in any condition for error, FWHM, or CV (all *p*s > .22; see Figure 4 and Table 1).

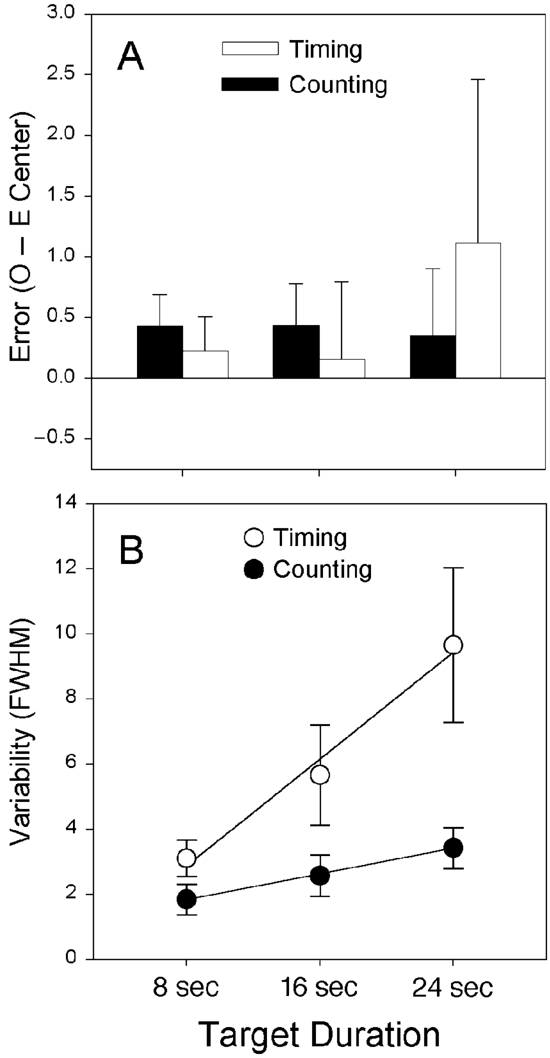

Figure 2. (A) Error measures by task and target duration for Experiment 1. Error is defined as the difference between the observed (O) and expected (E) centers of the Gaussian functions fit to individual participants' response distributions. Error bars represent ±1 *SD*. (B) Variability for Experiment 1 by task and target duration, defined as the full width at half-maximum (FWHM) in seconds of the Gaussian functions fit to individual participants' response distributions. Error bars represent ±1 *SD*.

DISCUSSION

These experiments have shown that counting and timing produce similar accuracy in temporal reproduction (Table 1, Figure 1A, Figure 2A, Figure 4). However, the FWHM of response distributions increases substantially less with counting than it does with timing (Table 1, Figure 1A, Figure 2B, Figure 4). In fact, the CV is constant for timing but decreases with increasing duration for counting (Table 1, Figure 3). These observations were replicated in Experiment 2, which generated comparable results and demonstrated that task order or practice effects cannot plausibly ac-

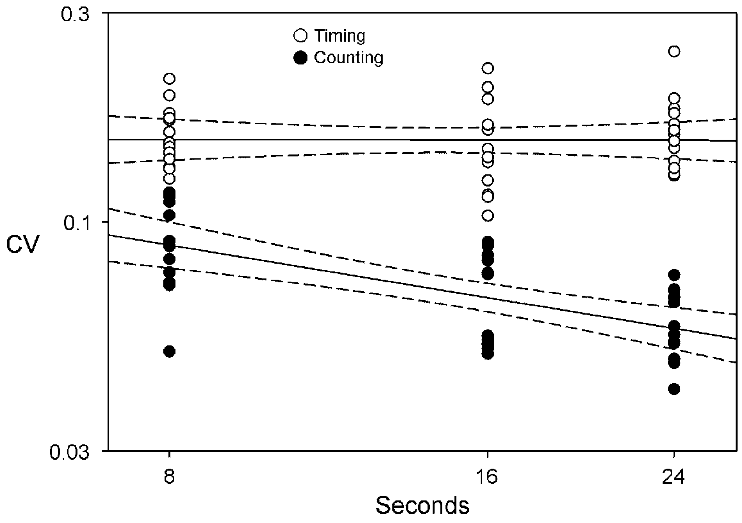

Figure 3. Coefficients of variation (CV) by task and target duration plotted in log–log coordinates for Experiment 1. The CV is derived by dividing the standard deviation (FWHM/2.355) by the center. Dashed lines indicate 95% confidence intervals. The slope for timing is not different from zero, indicating the scalar property, while the slope for counting is not different from −0.5, indicating the square-root property.

count for the data of Experiment 1 (Table 1, Figure 4). These experiments convincingly demonstrate that counting violates the scalar property, which requires a constant CV across durations. Nonetheless, how best to characterize changes in counting variability as a function of duration remains an open question. The simulation of the counting data (Figure 1B) and the plots of CV in log–log coordinates (Figure 3) suggest that the FWHM of the counting distributions are proportional to the square root of duration (FWHM $= a * \sqrt{T}$), unlike the timing distributions, in which FWHM scales linearly with duration (FWHM $= b * T$). However, a generalized version of Weber's law (GWL) also fits the counting data well (FWHM $= c * T + d$).

The fits to the counting data are similar for the SRM and GWL ($r^2 = .55$ and $.57$, respectively), and so the data themselves do not give a compelling reason to prefer one

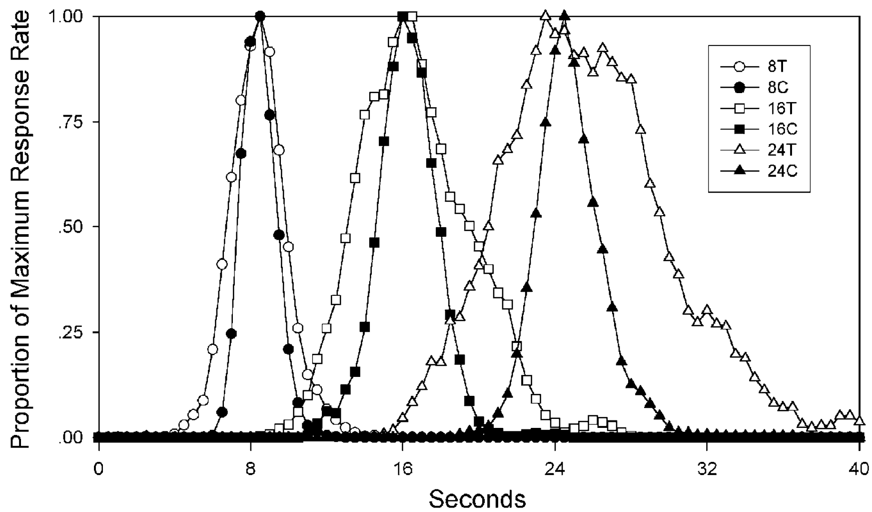

Figure 4. Peak functions for Experiment 2 averaged across 8 subjects sorted by task (T, timing; C, counting) and target duration (8, 16, 24 sec) plotted in half-second bins on a relative response scale. Counting preceded timing for all conditions.

Table 1
Means (*M*) and Standard Deviations (*SD*) for Error, FWHM, and CV by Task and Duration for Experiments 1 and 2

		Duration					
		8 sec		16 sec		24 sec	
Measure	Experiment	*M*	*SD*	*M*	*SD*	*M*	*SD*
		Counting					
Error	1	0.43	0.26	0.44	0.34	0.35	0.55
	2	0.41	0.27	0.24	0.48	0.62	0.75
FWHM	1	1.84	0.46	2.57	0.64	3.42	0.64
	2	1.92	0.49	2.96	0.91	3.13	1.14
CV	1	0.09	0.02	0.07	0.02	0.06	0.01
	2	0.10	0.02	0.08	0.02	0.05	0.02
		Timing					
Error	1	0.23	0.27	0.16	0.63	1.12	1.34
	2	0.35	0.38	0.37	1.37	1.18	1.70
FWHM	1	3.10	0.56	5.67	1.55	9.66	2.38
	2	3.23	0.41	6.24	0.96	9.14	1.38
CV	1	0.16	0.03	0.15	0.04	0.16	0.03
	2	0.16	0.02	0.16	0.02	0.15	0.02

Note—FWHM, full width at half maximum; CV, coefficient of variation.

over the other. However, the SRM is more parsimonious because it requires only a single free parameter to the GWL's two, and it makes intuitive sense because it reflects a mathematical truth derived from what we believe people are doing when they count: summing random variables. If a participant counts at a regular rate, the distribution of those counts has a reliable mean with relatively little variability, and the sum of individual counts represents the estimated duration. The standard deviation of a sum of random variables increases with the square root of the number of random variables, which is the mathematical basis of the SRM. If participants count at similar rates for all three durations, their count will be proportional to duration. Consequently, the standard deviations of their counting distributions will be proportional to the square root of the duration. Behavioral data are rarely so well modeled by such a simple process, and the simulation's close fit affirms common-sense notions about what participants are doing when they count (Figure 1B).

Some other objections to the GWL are also worth noting. In the log–log plot of the counting CV data in Figure 3, the GWL predicts a hyperbolic function of some form ($CV = a/x + b$) but makes no specific prediction about the value of the slope constant a. In contrast, the SRM specifically predicts that the CV for counting plotted in log–log coordinates will conform to a linear function with a slope of $-1/2$, which it does. In addition, the positive intercept modeled by the GWL asserts that all estimates based on counting must have a constant duration added to them that is independent of the duration being estimated. The value of this constant in terms of FWHM is 1.03 sec. Note that for the 8-sec duration, this constant accounts for 56% of the FWHM; for counting a 1-sec duration, this constant would account for 91% of the FWHM. When the timing data are so well fit by the classic version of Weber's law ($FWHM = c * T$), it seems problematic that counting would require the addition of such a substantial constant independent of time. The GWL provides no theoretically motivated account for this parameter, and the counting data are already well described by the simpler model, the SRM.

These experiments suggest that one can determine purely from behavioral data whether or not people are counting in order to estimate time, in that timing exhibits the scalar property whereas counting does not. The very different psychophysical characteristics exhibited by timing and counting further emphasize the critical importance of instructing participants not to count when one intends to study timing behavior.

REFERENCES

Bovens, N., & Brysbaert, M. (1990). IBM PC/XT/AT and PS/2 Turbo Pascal timing with extended resolution. *Behavior Research Methods, Instruments, & Computers*, **22**, 332-334.

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.

Church, R. M., Getty, D. J., & Lerner, N. D. (1976). Duration discrimination by rats. *Journal of Experimental Psychology: Animal Behavior Processes*, **2**, 303-312.

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review*, **8**, 698-707.

Crystal, J. D. (2001). Circadian time perception. *Journal of Experimental Psychology: Animal Behavior Processes*, **27**, 68-78.

Fetterman, J. G., & Killeen, P. R. (1990). A componential analysis of pacemaker-counter timing systems. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 766-780.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, **44**, 43-74.

Getty, D. J. (1976). Counting processes in human timing. *Perception & Psychophysics*, **20**, 191-197.

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

Gibbon, J., & Church, R. M. (1984). Sources of variance in information-processing theories of timing. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 465-488). Hillsdale, NJ: Erlbaum.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (Annals of the New York Academy of Sciences, Vol. 423, pp. 52-77). New York: New York Academy of Sciences.

Gilliland, A. R., & Martin, R. (1940). Some factors in estimating short time intervals. *Journal of Experimental Psychology*, **27**, 243-255.

Graves, R., & Bradley, R. (1987). Millisecond interval timer and auditory reaction time programs for the IBM PC. *Behavior Research Methods, Instruments, & Computers*, **19**, 30-35.

Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychological Bulletin*, **127**, 22-44.

Grondin, S., Meilleur-Wells, G., & Lachance, R. (1999). When to start explicit counting in a time-intervals discrimination task: A critical point in the timing process of humans. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 993-1004.

Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 3-18.

Killeen, P. R. (1991). Counting the minutes. In F. Macar, V. Pouthas, & W. J. Friedman (Eds.), *Time, action, & cognition: Towards bridging the gap* (pp. 203-214). Dordrecht: Kluwer.

Killeen, P. R., & Weiss, N. A. (1987). Optimal timing and the Weber function. *Psychological Review*, **94**, 455-468.

KRUUP, K. (1961). The influence of method on time judgments. *Australian Journal of Psychology*, **13**, 44-54.

MECHNER, F. (1958). Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior*, **1**, 109-122.

MECK, W. H., & CHURCH, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, **9**, 320-334.

PLATT, J. R., & JOHNSON, D. M. (1971). Localization of position within a homogeneous behavior chain: Effects of error contingencies. *Learning & Motivation*, **2**, 386-414.

RAKITIN, B. C., GIBBON, J., PENNEY, T. B., MALAPANI, C., HINTON, S. C., & MECK, W. H. (1998). Scalar expectancy theory and peak-interval timing in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, **24**, 15-33.

ROBERTS, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes*, **7**, 242-268.

SCHNEIDER, B. A. (1969). A two-state analysis of fixed-interval responding in the pigeon. *Journal of the Experimental Analysis of Behavior*, **12**, 677-687.

SPIVACK, G., & LEVINE, M. (1964). Consistency of individual differences in time judgments. *Perceptual & Motor Skills*, **19**, 83-92.

TALTON, L. E., HIGA, J. J., & STADDON, J. E. R. (1999). Interval schedule performance in the goldfish *Carassius auratus*. *Behavioural Processes*, **45**, 193-206.

WEARDEN, J. H., & MCSHANE, B. (1988). Interval production as an analogue of the peak procedure: Evidence for similarity of human and animal timing processes. *Quarterly Journal of Experimental Psychology*, **40B**, 363-375.

WHALEN, J., GALLISTEL, C. R., & GELMAN, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science*, **10**, 130-137.

NOTE

1. Let us assume that n represents the number of independent samples (i.e., count) used to estimate a target duration T. Let us further assume that those samples are drawn from a distribution D that has a mean of t, such that $T = n * t$. The variance, var(D), applies to any particular distribution of the samples, including Gaussian. A uniform distribution was chosen in this analysis merely for convenience of calculation. The same results could be obtained with many other possible combinations of n and var(D), as long as a proportional relationship was maintained between them. Mathematically, n trades off inversely with var(D) to keep the total variance of T constant. In other words, one would find the same result for T with twice as many samples taken from a distribution with half the variance, and so forth.

(Manuscript received May 22, 2002;
revision accepted for publication April 7, 2003.)