Journal of Experimental Psychology:
Human Perception and Performance
2003, Vol. 29, No. 6, 1102–1125

0096-1523/03/$12.00 DOI: 10.1037/0096-1523.29.6.1102

Modeling Effects of Rhythmic Context on Perceived Duration: A Comparison of Interval and Entrainment Approaches to Short-Interval Timing

J. Devin McAuley
Bowling Green State University

Mari Riess Jones
Ohio State University

Relative merits of interval and entrainment conceptions of the *internal clock* were assessed within a common theoretical framework by 4 time-judgment experiments. The timing of tone onsets marking the beginning and ending of standard and comparison time intervals relative to a context rhythm were manipulated; onsets were on time, early, or late relative to the implied rhythm, and 2 distinct accuracy patterns emerged. A quadratic *ending profile* indicated best performance when the standard ended on time and worst performance when it was early or late, whereas a flat *beginning profile* (Experiments 1–3) indicated uniform performance for the 3 expectancy conditions. Only in Experiment 4, in which deviations from expected onset times were large, did significant effects of beginning times appear in time-discrimination thresholds and points of subjective equality. Findings are discussed in the context of theoretical assumptions about clock resetting, the representation of time, and independence of successive time intervals.

Research on timing and time perception has a long, well-established history in experimental psychology (Fraisse, 1963; James, 1890; Wallin, 1911; Woodrow, 1951; for excellent reviews, see Allan, 1979; Grondin, 2001b; Killeen & Weiss, 1987). This is perhaps not surprising, because arguably all human behaviors involve some form of implicit or explicit timing. Nonetheless, a number of basic questions about the functional and neural mechanisms underpinning the timing of behavior remain unanswered. One concerns the nature of the *internal clock* involved in the perception of short time intervals.[1] Does the brain's timekeeper for short intervals operate like an hourglass, which can be started and stopped at arbitrary points in time, or does it resemble a self-sustaining oscillator with rhythmic characteristics?

In the short-interval timing literature, hourglass and oscillator conceptions of the internal clock are associated with two classes of model, typically referred to as *interval* and *entrainment* models, respectively (Barnes & Jones, 2000; Ivry & Hazeltine, 1995; Jones, 1976; Keele, Nicoletti, Ivry, & Pokorny, 1989; McAuley & Kidd, 1998; Pashler, 2001; Schulze, 1978). The present article compares these approaches within a common theoretical framework. The merits of this framework lie in its potential for generating a family of models that crystallize key distinctions between interval and entrainment explanations of timing. The empirical studies used to evaluate these models rely on the task illustrated in Figure 1A. Participants experience a tone sequence that marks out a series of context time intervals followed by a standard–comparison pair, and they are asked to judge the duration of the comparison interval relative to the standard. Our primary interest concerns the effect of the relative timing of various tone onsets within a sequence on the perceived duration of the standard.

Theories of Short-Interval Timing: Interval Versus Entrainment

Interval models of short-interval timing involve three independent processing components: a clock used to estimate duration, a

J. Devin McAuley, J. P. Scott Center for Neuroscience, Mind, and Behavior, Department of Psychology, Bowling Green State University; Mari Riess Jones, Department of Psychology, Ohio State University.

Portions of this research were presented at the 39th Annual Meeting of the Psychonomic Society, Dallas, Texas, November 1998. A portion of the data reported in this article were collected while J. Devin McAuley was a Music Cognition Postdoctoral Fellow at Ohio State University. This research was supported, in part, by National Science Foundation Grant BCS 9809446 to Mari Riess Jones.

We thank Simon Grondin and Richard Ivry for many helpful comments on an earlier version of this article. We also thank Heather Moynihan, Noah MacKenzie, Amandine Penel, Jennifer Puente, and Bruno Repp for comments on drafts of this article. We are also grateful to Nathaniel Miller, Jonathan Miller, Kevin Pang, Jacqueline Toft, and the other members of the Timing Research Group at Bowling Green State University for their comments and suggestions and to Kelli Rider for her assistance in preparing the manuscript.

Correspondence concerning this article should be addressed to J. Devin McAuley, Department of Psychology, Bowling Green State University, Bowling Green, Ohio 43403. E-mail: mcauley@bgnet.bgsu.edu

[1] By short intervals, we mean time intervals between approximately 200 and 2,000 ms. These durations are shorter than what are typically referred to as "short intervals" in the animal timing literature, but they are comparable to durations so designated in other human studies. One interesting question that is beyond the scope of this article is whether the same timing mechanism applies to durations shorter than about 2.5 s and longer than about 2.5 s. Perceived rhythm tends to break down for durations longer than a couple of seconds, exceeding the temporal window that Fraisse (1963) referred to as the "psychological present" (see also Grondin, 2001c).

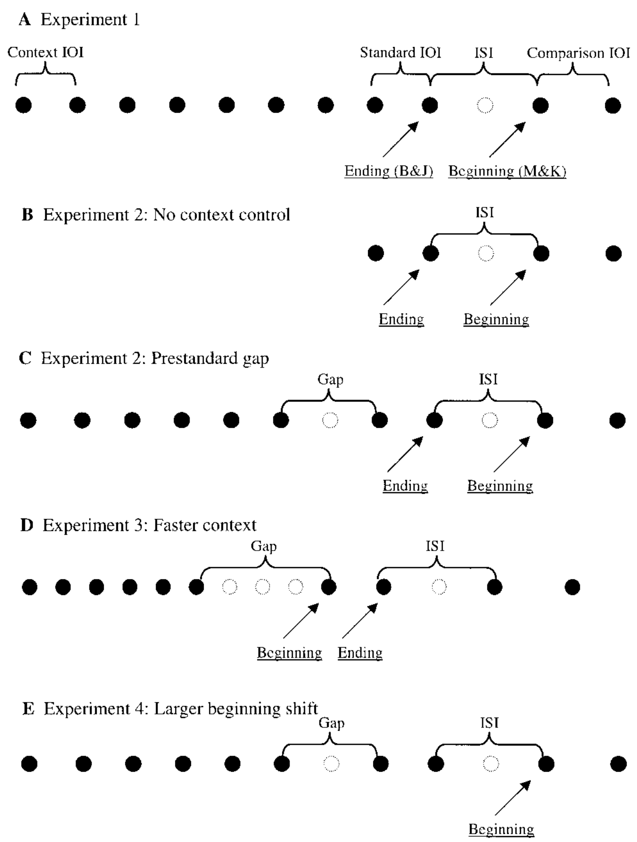

Figure 1. The time-judgment task in Experiments 1–4. Context sequences were followed by a single standard time interval and a comparison. The comparison differed from a given standard value, T, by $\pm\Delta T$ or 0 ms. Variants used by Barnes and Jones (2000) and McAuley and Kidd (1998) are denoted by B&J and M&K, respectively. Arrows indicate the onsets that were shifted in each experiment. Prestandard gaps (open circles indicate missing tones) are shown for Experiments 2, 3, and 4. Context interonset intervals (IOIs) were 600 ms in Experiments 1, 2, and 4; the context IOI was 300 ms in Experiment 3. The interstimulus interval (ISI) is the time interval between the stimulus (tone) onset ending a standard and the onset that initiates a comparison interval. See the *Method* section for details.

reference memory that stores duration information, and a comparison mechanism that measures a comparison's duration relative to that of a remembered standard (Church & Broadbent, 1990; Gibbon, 1977; Treisman, 1963). Perhaps the most influential interval model is scalar expectancy theory (SET; Church, Meck, & Gibbon, 1994; Gibbon, 1977; Gibbon, Church, & Meck, 1984; Rakitin et al., 1998). A schematic of SET is shown in Figure 2. In SET, the clock is a neural "pacemaker" that emits a continuous stream of pulses. Stimuli marking the beginning of a time interval (e.g., tone onsets) trigger the closing of a switch that allows pulses to enter an accumulator; stimuli marking the ending of a time interval trigger the opening of the switch. The number ("count") of pulses that enter the accumulator between the closing and opening of the switch serves as a code for the duration of the interval. Duration codes ("counts") for each time interval are maintained in working memory and then transferred to a more permanent reference memory. Judgments about a standard–comparison pair of intervals involve an online assessment of an updated comparison code in working memory relative to a remembered standard code in reference memory.

Entrainment models offer an alternative conception of short-interval timing. Their basic assumption is that the timekeeper is a self-sustaining (i.e., entrainable) oscillator (or collection of oscillators) that peak(s) in amplitude (a gross measure of neural activity) at regular temporal intervals (Large & Jones, 1999; McAuley, 1995; McAuley & Kidd, 1998). This view of time perception suggests that the internal clock for short intervals shares functional characteristics with the circadian clock (Jones, 1976; see Winfree, 2001, for an overview of various endogenous rhythms in biological systems). A schematic of an entrainment model is shown in Figure 3. Figure 3A shows a single oscillator, in which the time interval between successive amplitude peaks defines the oscillator's *period* (cycle duration, P). Relative phase, denoted as *phase* (ϕ), refers to a time point within each cycle that is measured relative to peak amplitude (the start of the oscillator's cycle). Peaks in amplitude occur at $\phi = 0$ and represent expected time points for stimulus onsets. Thus, stimulus onsets may occur at expected ($\phi = 0$) or unexpected ($\phi \neq 0$) phases (Figure 3C). Onsets that occur just prior to a given amplitude peak have a negative phase value,

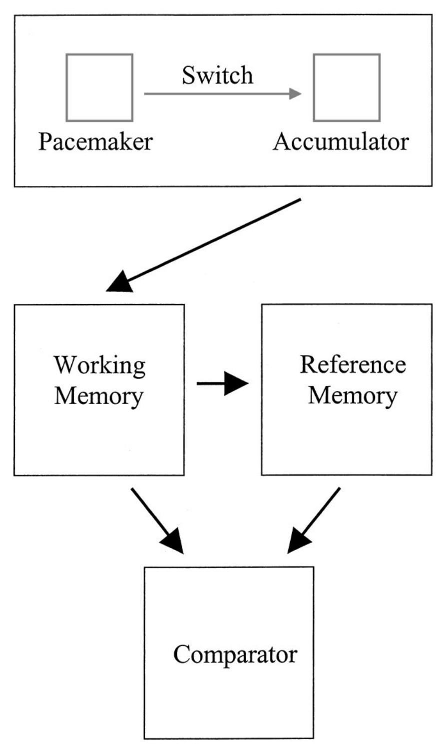

Figure 2. Schematic of a prototypical interval model with independent clock, memory, and comparison components. The clock consists of a pacemaker that emits a continuous stream of pulses; stimuli marking the beginning and ending of a time interval trigger the closing and opening of a switch that gates pulses into an accumulator. Duration is temporarily represented in working memory by the number ("count") of pulses that accumulate over the time interval. Successive counts are transferred to a reference memory. Judgments about duration involve a comparison between the current count (in working memory) and a sampled count (from reference memory).

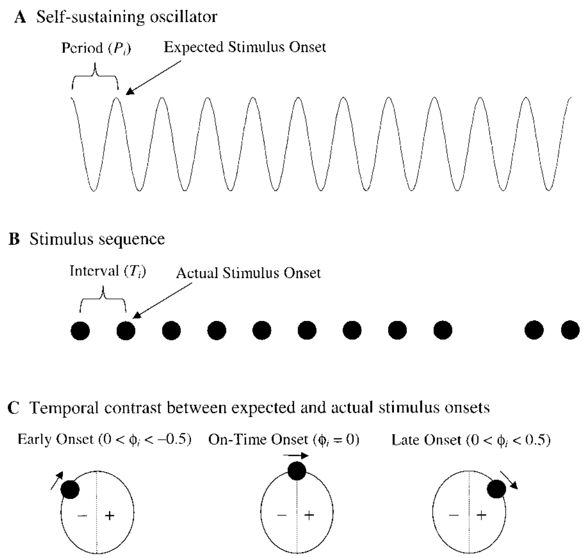

Figure 3. Schematic of a prototypical entrainment model. A: The clock and memory stages of the interval approach are integrated in a single oscillator that peaks in amplitude at roughly regular intervals, P; each amplitude peak represents an expectation for a stimulus onset. Phase, ϕ, measures time relative to the start of the oscillator's cycle. The start of the oscillator's cycle occurs at peak amplitude, which specifies $\phi = 0$. B: Example of a stimulus sequence marking a series of time intervals; the timing of stimulus onsets is measured only in terms of phase, ϕ, which serves to phase advance–delay the start of the oscillator's cycle and may also shorten–lengthen its period. C: Examples of stimulus onsets that occur early, on time, and late relative to an expected onset. Early onsets have a negative phase value and occur during the half of the oscillator's cycle that precedes the peak in amplitude (i.e., the expected time point). Late onsets, in contrast, have a positive phase value and occur during the half of the oscillator's cycle that follows the peak in amplitude.

meaning that the expected time point lags behind a given stimulus onset; that is, these onsets are unexpectedly "early." Conversely, stimulus onsets that occur shortly after a peak have a positive phase value and are unexpectedly "late." Duration judgments are based on the magnitude and sign of the relative time difference (contrast) between the actual onset that ends a comparison interval and the expected ending (cf. Jones & Boltz, 1989; McAuley, 1995). This difference is measured by phase and is termed *final temporal contrast* because it occurs at the end of the sequence; we denote final temporal contrast as C_{final}.

A key property of a self-sustaining oscillator is its responsiveness to unexpected onsets; specifically, such an oscillator can accommodate unexpected stimulus onsets by advancing or delaying its phase and/or shortening or lengthening its period. Thus, with an entrainable oscillator, expectancies (specific anticipations) about the temporal occurrence of forthcoming stimulus onsets are partially dependent on the timing of preceding stimulus onsets. If a sequence is rhythmic, then temporal expectations will come to align with stimulus onsets in the sequence, and time judgments will reflect greater accuracy than they do when stimulus onsets are misaligned with the oscillator.

Entrainment models differ from interval models in several ways. First, in entrainment models, stimulus markers merely serve to advance or delay an oscillator rather than to arbitrarily start or stop an internal clock. Second, time intervals are represented implicitly by the oscillator's period rather than explicitly as a stored code. Third, successive duration estimates are not independent but depend on the oscillator's response to the previous stimulus onsets. However, in a fourth respect, entrainment and SET approaches do not differ: They are conceptually similar in their common assumption that duration judgments reflect proportional (relative-time) differences between to-be-judged time intervals. In general, time judgments based on relative rather than absolute differences express Weber's law. In SET, Weber's law (also referred to as the *scalar property*) is explained by either explicitly mapping absolute time units into relative time units or by assuming scalar sources of variability at one or more stages of the model (Gibbon, 1977; Gibbon et al., 1984).

Effects of Rhythmic Context on Perceived Duration

One current issue in the time-perception literature concerns interval and entrainment explanations of effects of sequential context on duration perception (Barnes & Jones, 2000; Ivry & Hazeltine, 1995; Large & Jones, 1999; McAuley & Kidd, 1998; Pashler, 2001). Previous research has suggested that rhythmic characteristics are important for understanding the detection of deviations in sequence timing (Bharucha & Pryor, 1986; Halpern & Darwin, 1982; Jones & Yee, 1993; Monahan & Hirsh, 1990; Yee, Holleran, & Jones, 1994). A common finding is that increasing the number of repetitions of a target time interval within an isochronous sequence improves duration-discrimination performance (Drake & Botte, 1993; Grondin, 2001a; Ivry & Hazeltine, 1995; Jones & Yee, 1993; McAuley & Kidd, 1998; Michon, 1964; Schulze, 1989; Vos, van Assen, & Franek, 1997; Yee et al., 1994); conversely, variability in timing of sequence presentations typically worsens duration-discrimination performance (Drake & Botte, 1993; Yee et al., 1994).

In a recent study relevant to both interval and entrainment theories, Barnes and Jones (2000, Experiment 1) had participants listen to a sequence of seven interonset intervals (IOIs; all 600 ms) marked by brief tones, as shown in Figure 1A. This context sequence was followed by two intervals: a standard and a comparison. Participants judged the duration of the comparison interval relative to the standard, with the instruction to ignore the context sequence. The critical manipulation involved the onset of the tone marking the end of the standard; this tone was early, on time, or late relative to the rhythm implied by the context sequence, yielding three (or more, depending on the experiment) different standards (ranging between 524 and 676 ms). For example, for a context rhythm composed of 600-ms intervals, the onset of the tone ending the 524-ms standard was relatively early, whereas those of the 600- and 676-ms standards were on time and late, respectively. In this rhythmic context, Barnes and Jones found that comparative time estimates were more accurate, as measured by the overall proportion of correct shorter, same, and longer responses, for a standard that ended on time than for standards ending early or late; this resulted in a significant quadratic trend, shown schematically in Figure 4. This inverted-U pattern of accuracy was termed an *expectancy profile* by Barnes and Jones. Note that a null context effect is expressed by a flat accuracy profile, shown also in Figure 4.

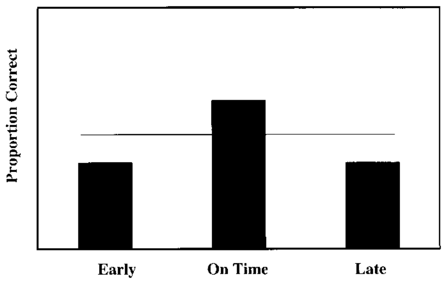

Figure 4. A schematic of two possible accuracy patterns associated with judgments about standard and comparison durations presented in a rhythmic context. Solid bars illustrate the inverted-U function reported as a quadratic expectancy profile by Barnes and Jones (2000); the dashed line illustrates a flat accuracy profile, consistent with null effects of rhythmic context on time judgments.

The Barnes and Jones (2000) task is similar to one introduced by McAuley and Kidd (1998, Experiment 2). However, as shown in Figure 1A, the critical manipulation in the research of McAuley and Kidd involved the relative timing of the tone delineating the beginning of the comparison time interval, not the tone ending the standard (cf. Figure 1A). To distinguish these variants, we refer to Barnes and Jones (2000) as B&J and to McAuley and Kidd (1998) as M&K. The findings of M&K generally converge with those of B&J in that deviations from an expected rhythm influenced perceived duration, with early and late beginnings of a comparison interval producing subjective over- and underestimates of the standard, respectively.

Both interval and entrainment theories can explain the patterns of outcomes associated with rhythmic context effects that we have reviewed. In the following paragraphs, we briefly summarize each explanation.

First, consider the interval approach to context effects. Several interval theories propose effects of rhythmic and irregularly timed contexts on the internal clock. Treisman, Faulkner, Naish, and Brogan (1990) hypothesized that a series of context clicks might speed up the internal clock by increasing arousal, and there is some empirical evidence to support this view (cf. Wearden, Philpot, & Win, 1999). A second way that an interval model has been extended to address context effects assumes that the separate duration codes for time intervals making up a rhythmic context sequence are combined in reference memory using some form of averaging. This second suggestion resembles the multiple-look hypothesis proposed by Drake and Botte (1993; Monahan & Hirsh, 1990; see also Keele et al., 1989). In the multiple-look hypothesis, repetitions of the standard produce lower discrimination thresholds because more precise statistical estimates of a standard duration obtain when multiple samples (“looks”) are averaged compared with when estimates are based on a single sample. This view is consistent with a long-standing tradition in psychology that appeals to the concept of assimilation through averaging to explain contextually induced distortions in perception (Hellstrom, 1985; Hollingworth, 1910; Michels & Helson, 1954; Pashler, 2001; Turchioe, 1948; Woodrow, 1951).[2]

Under the right conditions, an interval model that averages successive time intervals can explain the quadratic trend associated with the expectancy profiles reported by B&J (Experiment 1). To illustrate, let three standard time intervals (540, 600, and 660 ms) determine, respectively, early, on-time, and late conditions following a sequence of 600-ms context intervals. If participants do not ignore the context intervals but instead use them to produce an aggregate memory code (i.e., a simple average of context-plus-standard intervals), then the 540- and 660-ms standards (those ending early and late) will be over- and underestimated, respectively. As a consequence, judgments about the comparison intervals based on overestimated standards will be skewed toward *shorter* responses, whereas those based on underestimated standards will be skewed toward *longer* responses. In summary, an interval model addresses context effects by using an interval-averaging approach to assimilation; this leads to the prediction of an inverted-U pattern of accuracy similar to the quadratic expectancy profile observed by B&J.[3]

Next, consider the entrainment view of context effects. Entrainment models are also designed to address both rhythmic and irregularly timed contexts (e.g., Large & Jones, 1999). In rhythmic contexts, an entrainment explanation differs from interval averaging in its emphasis on the degree of synchrony between an internal oscillator and the timing of stimulus onsets. According to this approach, quadratic expectancy profiles are a consequence of C_{final} values that are different for standards ending early and late than for standards that end on time. For example, with a context sequence of 600-ms intervals, a 600-ms standard interval (i.e., one that ends on time) continues the preceding context rhythm; thus, in most cases, C_{final} will be an unbiased estimate of the relative difference between the comparison and standard intervals. However, if the standard ends unexpectedly (i.e., it ends early [540 ms] or late [660 ms]), then the oscillator must correct its phase and/or period in the direction implied by the relative phase of the unexpected onset to stay on track (i.e., remain synchronized). Early onsets serve to advance the oscillator and shorten its period, whereas late onsets serve to delay the oscillator and lengthen its period. In these cases, an entrainable oscillator corrects both its phase and its period. However, for such an oscillator, *complete* phase and/or period correction is not mandatory! This is important because partial corrections of phase and/or period tend to produce distortions in the internal representation of a presented time interval, which persist to affect C_{final}.

Systematic differences in C_{final} values generate the following general predictions. When an unexpected standard is shorter than the context IOI (e.g., 540 ms in the example above), C_{final} values tend to be exaggerated for shortened comparisons and attenuated for lengthened comparisons, increasing the tendency for participants to respond *shorter*. Conversely, when the standard is longer than the context IOI (e.g., 660 ms), C_{final} values tend to be exaggerated for longer comparisons and attenuated for shorter comparisons, increasing *longer* responses. For the latter, participants should make proportionally more correct longer responses and fewer correct shorter responses than they do when the standard ends on time. Thus, in rhythmical contexts, C_{final} values based on unexpected standards predict more judgment errors than do those

[2] *Assimilation* refers to the tendency to distort one stimulus in the direction of presented context stimuli.

[3] The durations used by B&J differed from those cited here; these values parallel ones used in the present research.

based on expected standards. This alternative approach to assimilation yields systematic errors in duration judgments that can account for the expectancy profiles reported by B&J.

In summary, research has indicated that rhythmic contexts systematically distort comparative time judgments. Two theoretical approaches (interval and entrainment) offer different explanations of the resulting patterns of errors. In the present research, we developed specific models that realize aspects of both theoretical approaches within a more general framework, and we tested these models in new experimental designs.

A General Framework

In this section, we present a general theoretical framework for modeling effects of rhythmic context on duration perception. This framework derives from a discrete formalization of a single oscillator that incorporates linear phase and period correction. The framework enables the generation of a family of models that address key differences between interval and entrainment explanations of expectancy profiles. Relevant differences revolve around three issues: (a) timekeeper response to stimulus onsets (closing a switch vs. oscillator phase correction), (b) internal representation of duration (stored code vs. oscillator period), and (c) successive duration estimates (independent vs. dependent). A fourth issue concerns the process of comparing two durations.

Timekeeper Response to Stimulus Onsets

The first issue concerns the effect of stimulus onsets on the internal clock. Do stimulus onsets automatically start and stop (i.e., reset) the internal clock (as proposed by interval theorists), or does the impact of an onset depend on its phase relative to an internal oscillator (as proposed by entrainment theorists)? To model both interval and entrainment assumptions about the timekeeper's response to stimulus onsets, we propose a linear phase-correction term:

$$\phi_{i+1} = (1 - W_{\phi})C(\phi_i, \mathrm{IOI}_i, P_i). \tag{1}$$

In this equation, C represents temporal contrast, which measures the phase of the ith stimulus onset prior to any phase advance or delay; see Appendix A for calculation of C. Formally, C varies between $-.5$ and .5 and is recursively determined by ϕ_i (the relative phase of the ith stimulus onset), IOI_i (the interval between the ith and $i + 1$ stimulus onset), and P_i (the current oscillator period). The parameter W_{ϕ} determines the amount of phase correction ($0 < W_{\phi} < 1.00$). When $W_{\phi} = 1.00$, the oscillator restarts its cycle after each stimulus onset. This is full phase correction; it mimics the arbitrary reset assumption of the interval approach. In contrast, $W_{\phi} = 0$ represents the opposite extreme; here, the oscillator entirely fails to correct its phase following a change in stimulus timing. $W_{\phi} = 0$ simulates a rigid oscillator that does not reentrain following a phase shift.

Internal Representation of Duration

The second issue concerns the internal representation of duration. Is duration represented in memory as an explicit duration code (as proposed by interval theorists), or is it only implicit in the period of a self-sustaining oscillator (as proposed by entrainment theorists)? We capture both views by assuming that time intervals are represented by an oscillator period, P_i. Indeed, recent formalizations of interval models have replaced the pacemaker–accumulator conception of the clock with a collection of oscillators, all of fixed periods (Church & Broadbent, 1990; Matell & Meck, 2000; Miall, 1989). Consistent with such models, each IOI in a sequence then determines an explicit duration code, P_i. Alternatively, from an entrainment perspective, the period (P_i) is that of single oscillator, which changes in real time; as such, it may be interpreted as an implicit momentary representation of duration. In either case, an oscillator model (interval or entrainment) that completely corrects its period to an unexpected standard interval will reliably represent the duration of that standard.

In this framework, the internal representation of duration can change over time through the mechanism of period correction. To model differences in the amount of period correction, we propose a linear period correction term:

$$P_{i+1} = [1 + W_p\, C(\phi_i, \mathrm{IOI}_i, P_i)]\, P_i. \tag{2}$$

Equation 2 describes the change in P_i from the ith to $i + 1$th stimulus onset. As with phase correction, period correction depends on temporal contrast $C(\phi_i, \mathrm{IOI}_i, P_i)$ and occurs only when stimulus onsets are out of phase with the oscillator. The parameter W_p ($0 < W_p < 1.0$) determines the amount of period correction in response to successive IOIs. Complete period correction is specified by $W_p = 1.00$; the period of the oscillator is corrected to match the current IOI, up to a limit ($\pm P_i/2$). The opposite extreme is specified by $W_p = 0$; in this case, the oscillator retains its initial (or preset) period, P_0, throughout any sequence of (arbitrary) time intervals.

Successive Time Estimates

The third issue concerns the independence of successive estimates of duration. Are these estimates independent (as proposed by interval theorists) or dependent (as proposed by entrainment theorists)? Although Equations 1 and 2 are difference equations, expressing i to $i + 1$ state changes, they can simulate either independence or dependence, depending on the values of the parameters W_{ϕ} and W_p. When $W_k = 1.00$ ($k = \phi$ or p, respectively), the system at time $i + 1$ is independent of its preceding state i. In contrast, complete dependence of phase or period is given by $W_k = 0$ ($k = \phi$ or p, respectively). When $W_{\phi} = W_p = 0$, the system at Time $i + 1$ is determined entirely by initial phase (ϕ_0) and period (P_0). Intermediate values of W_{ϕ} and W_p, respectively, specify the degree to which phase and period at Time i are transformed into different values at Time $i + 1$ and, hence, are dependent. Finally, entrainment explanations of assimilation rest on dependencies among representations of successive time intervals. Indeed, predictions about expectancy profiles (e.g., flat vs. quadratic) express the degree of assimilation that results from interactions of period and phase correction over time (i.e., between onsets i and $i + 1$).

Judgments About Relative Duration

A fourth issue concerns the process of comparing two durations. Both interval- and entrainment-model conceptions of the decision process involve proportional (relative-time) differences. In SET and related interval models, duration judgments based on proportional differences are appealing because they account for the scalar property of timing (Gibbon, 1977; Gibbon et al., 1984). In entrainment models, relative time is implied in the relational definition of phase (McAuley, 1995; McAuley & Kidd, 1998). In the present analyses, differences between an internal standard and a comparison are uniformly indexed using C_{final}. That is, for all of the quantitative models reported in this article, values of C_{final} are mapped into response probabilities using a Luce-choice rule (see Appendix A). Generally, values of C_{final} less than and greater than zero correspond, respectively, to judgments of *shorter* and *longer*, whereas values around zero correspond to *same* responses.

Four Cardinal Models

Four cardinal models occupy the corners of a parameter space that is specified by binary values of W_ϕ and W_p: (a) full-reset ($W_\phi = 1.00$, $W_p = 1.00$), (b) phase-reset ($W_\phi = 1.00$, $W_p = 0$), (c) period-reset ($W_\phi = 0$, $W_p = 1.00$), and (d) the no-reset ($W_\phi = W_p = 0$) models. These cardinal models define limiting cases of our framework within a two-parameter space, as shown in Figure 5. Thus, in this space, a model with high period correction is located closer to 1.0 on the *y*-axis, whereas one with high phase correction is located closer to 1.0 on the *x*-axis. The two cardinal models with complete phase correction (full-reset and phase-reset) characterize prototypical interval approaches to timing (Church & Broadbent, 1990; Gibbon, 1977; Gibbon et al., 1984; Matell & Meck, 2000), whereas the other two cardinal models, those without phase correction (period-reset and no-reset), reflect the two corresponding limiting cases that characterize typical "beat-based" models. We assess predictions of all models under the common assumptions that initial values for phase and period are $\phi_0 = 0$ and $P_0 =$ initial-context IOI.

In this research, a central theme involves the potential of the four cardinal models and the related models to predict different accuracy profiles in the task of B&J–M&K as a function of two variables: (a) *beginning manipulations,* which are defined by timing variations of stimulus onsets that delineate the beginning of a to-be-judged time interval (standard and/or comparison), and (b) *ending manipulations,* which are defined by timing variations of stimulus onsets that mark the ending of a standard (hence, the duration of a standard). A key to understanding model predictions with respect to manipulations of stimulus onsets involves distinguishing those onsets that mark the beginning from those that mark the ending of some to-be-judged time interval. In this respect, it is important to note that manipulations of beginning onsets never change the duration of any to-be-judged time interval (standard or comparison), whereas manipulations of ending onsets always do so. Theoretically, this fact is important because period correction is required only in the latter case, in order for the oscillator to accurately represent the standard's duration.

Full-Reset Model ($W_\phi = 1.00$, $W_p = 1.00$)

In this limiting case, the general model responds to each stimulus onset by resetting its phase (i.e., it is fully corrected to $\phi_i =$

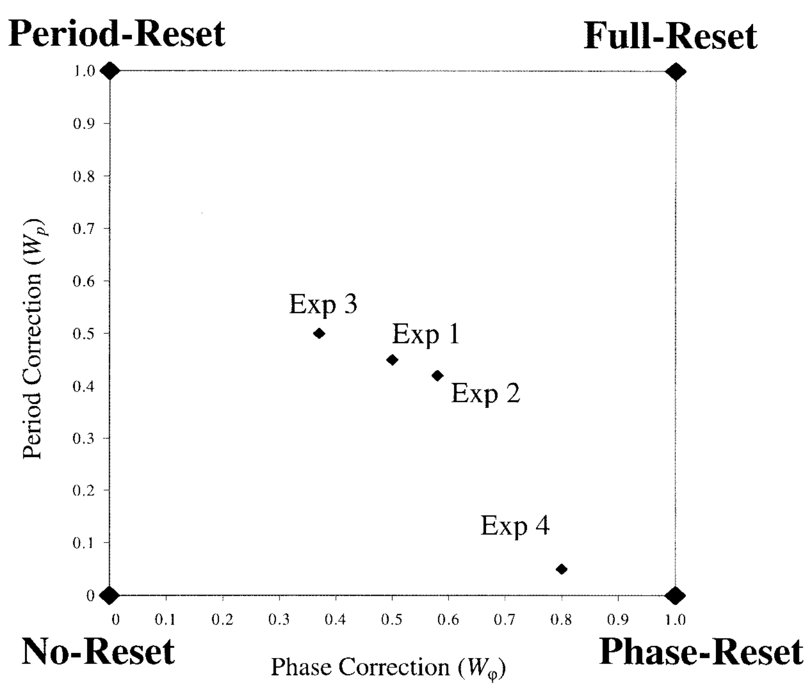

Figure 5. Representation of the parameter space of the proposed theoretical framework; the four corners of the space represent the four cardinal models, defined by binary values of phase correction and period correction. The full-reset and phase-reset models ($W_\phi = 1.00$) capture the arbitrary phase-reset property of interval models. The period-reset and no-reset models ($W_\phi = 0$) capture no-rest properties of "beat-based" models. Within this space, plotted points indicate estimates of phase and period correction for the best quantitative fits to the data from Experiments 1–4. Exp = experiment.

0) and by matching its period to each IOI in the sequence (P_i = IOI_i). This model is characteristic of interval models in several important ways. First, $W_\phi = 1.00$ implies that the internal clock is automatically reset in response to each stimulus marker, thereby simulating an interval timer's switch (e.g., as in SET). Second, $W_p = 1.00$ implies that the internal clock obtains independent estimates of successive time intervals, with each value P_i fixed by IOI_i. Thus, this model is context free in that each new time interval (IOI) erases memory of the preceding one because the oscillator's period completely adjusts to match the new IOI (given that a new IOI is within $\pm P_i/2$ of P_i).

The full-reset model predicts that flat accuracy profiles result from manipulations of the tone onsets of either the beginning or ending of a to-be-judged time interval that follow a rhythmic context (e.g., for sequences of Figure 1A). These null predictions derive from the model's full phase- and period-correction properties, respectively. A summary of the general predictions of the full-reset model for the B&J–M&K task appears in Table 1 (first row).

Phase-Reset Model ($W_\phi = 1.00$, $W_p = 0$)

In this case, the general model responds to each stimulus onset by resetting its phase, but it does not adjust period. Like the full-reset model, the phase-reset model shares characteristics of interval models. The phase-resetting assumption ($W_\phi = 1.00$) mimics an interval timer switch. However, the phase-reset model is more sensitive to context than the full-reset model; without complete period correction, it does not obtain independent estimates of successive IOIs. Rather, $W_p = 0$ represents the opposite extreme. Period never changes throughout the sequence; therefore, estimates of the standard duration are determined by the initial period, P_0 (i.e., P_0 = initial context IOI). Consequently, the phase-reset model predicts maximal assimilation of the standard with the context.

The phase-reset model predicts a flat accuracy profile for manipulations of onsets marking the beginning of a to-be-judged time interval; this is because of the model's strong phase-resetting properties. However, unlike the full-reset model, the phase-reset model predicts a quadratic ending profile. The latter prediction is the result of maximum assimilation of the standard interval with isochronous context intervals. A summary of the general predictions of the phase-reset model for the B&J–M&K task appears in Table 1 (second row).

Table 1
Predicted Accuracy Profiles for the Four Cardinal Models on the Basis of Binary Values of Phase and Period Correction for Early, On-Time, and Late Comparison Beginnings and Standard Endings in the B&J–M&K Task

	Type of correction		Expectancy profile	
Model	Phase	Period	Beginning	Ending
Full-reset	Yes	Yes	Flat	Flat
Phase-reset	Yes	No	Flat	Quadratic
Period-reset	No	Yes	Quadratic	Flat
No-reset	No	No	Quadratic	Quadratic

Note. B&J = Barnes & Jones (2000); M&K = McAuley & Kidd (1998).

Period-Reset Model ($W_\phi = 0$, $W_p = 1.00$)

In this case, the general model alters its period to match successive time intervals but does not advance or delay its phase. This model resembles a "reckless" oscillator because without phase correction to connect expectancies to onsets, peaks in amplitude drift relative to tone onsets. As a consequence, phase may not always accurately specify the direction and magnitude of the period change required to correctly match variations in sequence timing. However, as long as initial phase synchrony ($\phi_0 = 0$) is stipulated, and beginning times of any subsequent to-be-judged time intervals align with the implied rhythm, then the period-reset model will accurately correct its period.

Given the above assumptions, the predictions of the period-reset model are the opposite of those of the phase-reset model. The period-reset model predicts a quadratic accuracy profile for manipulations of beginning markers and a flat accuracy profile for manipulations of ending markers. The flat profile is predicted because complete period correction permits a veridical estimation of the relative difference between an unexpected standard and its comparison (via C_{final}). A summary of the predictions of the period-reset model for the B&J–M&K task appears in Table 1 (third row).

No-Reset Model ($W_\phi = W_p = 0$)

In this limiting case, the general model reduces to a "rigid" oscillator, which projects a predetermined path of expected points in time. That is, lacking both phase and period correction, C_{final} is rigidly determined by initial conditions (P_0, ϕ_0). The no-reset model captures important aspects of prominent beat-based models. It reflects Fraisse's (1978) hypothesis that people respond to temporal relationships via resonant internal oscillations (cf. also Schulze, 1978). The model also echoes the clock grid of Povel and Essens's (1985) beat-induction model, in which neither period nor phase of the internal clock adapts in real time. Finally, the no-reset model resembles the beat-based oscillator model typically evaluated by interval theorists (Keele et al., 1989; Pashler, 2001; Schulze, 1978, 1989).

The no-reset model predicts quadratic expectancy profiles associated with both unexpected beginnings and unexpected endings in the task of B&J–M&K. Here, C_{final} reflects temporal distortions due to failures in phase correction whenever a to-be-judged duration begins early or late; it also reflects distortions resulting from failures in period correction whenever a standard ends early or late. A summary of the general predictions of the no-reset model for this task appears in Table 1 (bottom row).

Overview of Experiments

Four experiments evaluated predictions of the four cardinal models (and variants) with regard to effects of rhythmic context on time judgments (Table 1). In this research, we operationalized beginnings and endings as stimulus onsets (i.e., physical markers) that followed and preceded, respectively, a gap in the sequence (e.g., Garner, 1974). Figure 1 summarizes the specific beginnings and endings manipulated in the four experiments.

Experiment 1 provided an overview of model predictions in the task of B&J–M&K. We combined the ending manipulation of B&J with the beginning manipulation of M&K. Experiment 2 extended Experiment 1 by introducing a gap between the context sequence and the standard, designed to weaken the influence of the rhythmic context sequence. Experiment 3 preserved the prestandard gap and introduced a faster context rate, which permitted examination of the effects of further timing constraints on performance. Finally, a different task was employed in Experiment 4, one which used a wider range of beginning manipulations to assess the limits, if any, of phase correction.

Experiment 1: Standard Endings and Comparison Beginnings

Experiment 1 had two goals. First, we evaluated the general predictions of all four cardinal entrainment models by manipulating the beginning of the comparison interval and the ending of the standard interval relative to a 600-ms context rhythm in a single experiment (as illustrated in Figure 1A). Table 1 indicates that these models cover the four combinations of predicted accuracy profiles resulting from (a) manipulation of the beginning of the comparison and (b) manipulation of the ending of the standard. By *beginning profile*, we mean an accuracy pattern that results from early, on-time, and late onsets of stimuli that mark the beginning of a to-be-judged time interval (e.g., a comparison). By *ending profile*, we mean an accuracy pattern that results from early, on-time, and late onsets of stimuli that mark the ending of a to-be-judged time interval, specifically a standard. The no-reset model predicts inverted-U (i.e., quadratic) beginning and ending profiles; the period-reset model predicts a quadratic beginning profile and a flat ending profile; the phase-reset model predicts a flat beginning profile and a quadratic ending profile; and the full-reset model predicts two flat profiles, indicating effects of neither the beginning nor the ending manipulation. We evaluated these models using a design in which the standard duration was either predictable (blocked) or unpredictable (randomized) over trials in a session.

Our second goal was in the service of the first; it entailed isolating determinants of expectancy profiles by use of a design that does not confound examination of the effects of unexpected beginnings and endings. M&K manipulated the beginning of the comparison, whereas B&J manipulated the ending of the standard. An issue that arises in the B&J experiments is that their variations in the ending of the standard also resulted in variations in the beginning of the comparison interval. Therefore, the expectancy profile reported by B&J could have been a function of unexpected standard endings (i.e. support for phase-reset), comparison beginnings (support for period-reset), or both (support for no-reset).

Method

Participants. Participants were 40 native English-speaking students from an introductory psychology class at Ohio State University who volunteered in return for course credit. All reported having normal hearing. They were randomly assigned to either a randomized standard ($n = 22$) or a blocked standard ($n = 18$) condition in one of four presentation orders.

Design. Experiment 1 had a 3 × 3 × 2 mixed-factorial design. Three standard endings were crossed with three comparison beginnings for each of three comparison IOIs (shorter, same, longer). Both standard ending and comparison beginning were within-subject manipulations. The single between-subjects variable was presentation condition: randomized or blocked ending of the standard. The three standard IOIs were 600 ms (on time), 540 ms (ending 60 ms early), and 660 ms (ending 60 ms late), with the duration of the comparison IOI yoked to the standard. On shorter trials, the comparison IOI was 10% shorter than the standard; on longer trials, it was 10% longer than the standard; and on same trials, it was equal to the standard. Manipulating the comparison beginning involved changing the time interval between the tone onset marking the ending of the standard and the tone onset marking the beginning of the comparison; to distinguish this interval, we use the term *interstimulus interval* (ISI; see Figure 1).

Comparisons could begin 60 ms early, on time, or 60 ms late relative to the expected onset time implied by a continuation of the 600-ms context rhythm. For the 600-ms standard (the on-time ending condition), the ISIs corresponding to early, on-time, and late comparison beginnings were 1,140, 1,200, and 1,260 ms, respectively. For the 540-ms standard, which ended early relative to the context rhythm, the ISIs for the three beginning conditions were lengthened by 60 ms in order to compensate for the shortened standard. Similarly, for the 660-ms standard, which ended late relative to the context rhythm, the ISIs for the three beginning conditions were shortened by 60 ms in order to compensate for the lengthened standard.

Stimuli and apparatus. All stimulus sequences comprised a series of 60-ms 440-Hz sine tones presented at a comfortable listening level (60 dB). Stimulus generation and response collection was controlled by the MIDILAB software package (Todd, Boltz, & Jones, 1989). Stimuli were generated on an IBM PC-compatible computer interfaced by a Roland MPU-401 MIDI processing unit that controlled a Yamaha TX81Z FM tone generator. The stimuli were then transmitted to a separate experimental room and amplified using a Rane HC 6 headphone console. Each participant listened to the stimuli over AKG K270 headphones.

Procedure. On each trial, participants judged the duration of a comparison IOI relative to the standard by pressing one of three buttons, labeled *shorter*, *same*, and *longer*, within a 2.5-s response interval. A trial began with a 600-ms (high-pitched) warning tone followed by 600 ms of silence prior to the onset of the first tone. Participants heard recorded instructions and studied a task diagram (e.g., Figure 1A). They received two sets of practice trials with corrective feedback (i.e., the correct alternative was supplied), followed by six blocks of test trials with noncorrective feedback (correct vs. incorrect).

Initial practice trials required participants to judge nine single pairs of standard–comparison time intervals involving the three standards (standard IOI = 540, 600, or 660 ms) crossed with three different comparisons. Next, participants received 27 practice trials involving an isochronous context sequence that preceded the standard IOI. Participants were told to ignore the context and concentrate on standard and comparison intervals; they were informed that attending to the context pattern would distract them. Practice trials contained an equal number of early, on-time, and late standard endings and comparison beginnings.

The test session comprised six blocks of 54 trials. In the random presentation condition, participants heard each combination of standard ending and comparison beginning six times in each 54-trial block. In the blocked presentation condition, the standard ending was fixed within a block of trials, whereas the comparison beginning varied; each comparison beginning was presented 18 times in each 54-trial block. Presentation order (across blocks) for the different standard-ending conditions was counterbalanced between participants. In both presentation conditions, equal numbers of shorter, same, and longer comparison IOIs occurred every 27 trials. Overall, there were 12 judgments for each condition.

Results

Figure 6 reports mean proportion of correct (PC) responses for the three standard endings and the three comparison beginnings, respectively (also shown are model predictions, discussed below). Mean PC collapses over the three comparison IOIs (i.e., shorter, same, and longer comparisons) for both presentation conditions (blocked vs. randomized). Blocking (vs. randomization) of standard durations did not significantly affect judgment accuracy; mean PCs were .66 and .64 for the blocked and the randomized presentation conditions, respectively. A $3 \times 3 \times 2$ mixed analysis of variance (ANOVA) on overall PC revealed a main effect of standard ending, $F(2, 76) = 27.18$, $MSE = 0.023$, $p < .01$, but no other main effects or interactions (all ps $> .10$).

These findings reveal two different accuracy profiles. One is a quadratic *ending profile* (Figure 6B) that replicates B&J. Higher overall PCs occurred whenever the standard time interval ended on time (mean PC $= .73$) compared with when it ended early (mean PC $= .60$) or late (mean PC $= .62$). Post hoc comparisons between the early, on-time, and late ending conditions (using Tukey's honestly significant difference [HSD]) revealed significantly lower PCs for both early and late ending standards relative to on-time standards ($p < .01$ for both comparisons). This profile did not differ as a function of blocking versus randomization of the standard duration. The other accuracy profile is a flat *beginning profile* (Figure 6A). Overall, accuracy was nearly identical for comparisons that arrived early, on time, and late; mean PCs for early, on-time, and late beginning conditions (collapsed over standard endings) were approximately .64. Taken together, these findings confirm that the quadratic expectancy profiles reported by B&J depended on standard endings, not on comparison beginnings. Moreover, the quadratic ending profiles in Experiment 1 did not depend on whether the standard was blocked (fixed) or randomized (roving) from trial to trial.

Discussion

The data of Experiment 1 reveal different accuracy profiles for beginning and ending manipulations. In this section, we discuss possible determinants of these profiles, examining both qualitative and quantitative predictions of interval and entrainment models.

Qualitative model predictions concern the general shape of the beginning and ending profiles (flat vs. quadratic). For Experiment 1, these predictions appear in Table 1 for the four cardinal models. Clearly, only the phase-reset model (which is consistent with an interval view of timing) correctly predicted both a quadratic ending profile and a flat beginning profile. Predictions of the phase-reset model also appear graphically in Figures 6A and 6B, along with the data. This support for the phase-reset model is not qualified by the predictability of standard durations, because the expectancy profile was similar in blocked and randomized ending conditions. As indicated in Table 1, neither the rigid oscillator (no-reset) nor the reckless oscillator (period-reset) performed well because both lack phase correction; thus, both incorrectly predicted a quadratic beginning profile. The full-reset model also performed poorly, because it incorporates full period correction and hence failed to predict the quadratic ending profile.

We next examined the pattern of errors responsible for the quadratic ending profile. In Experiment 1, the time-judgment

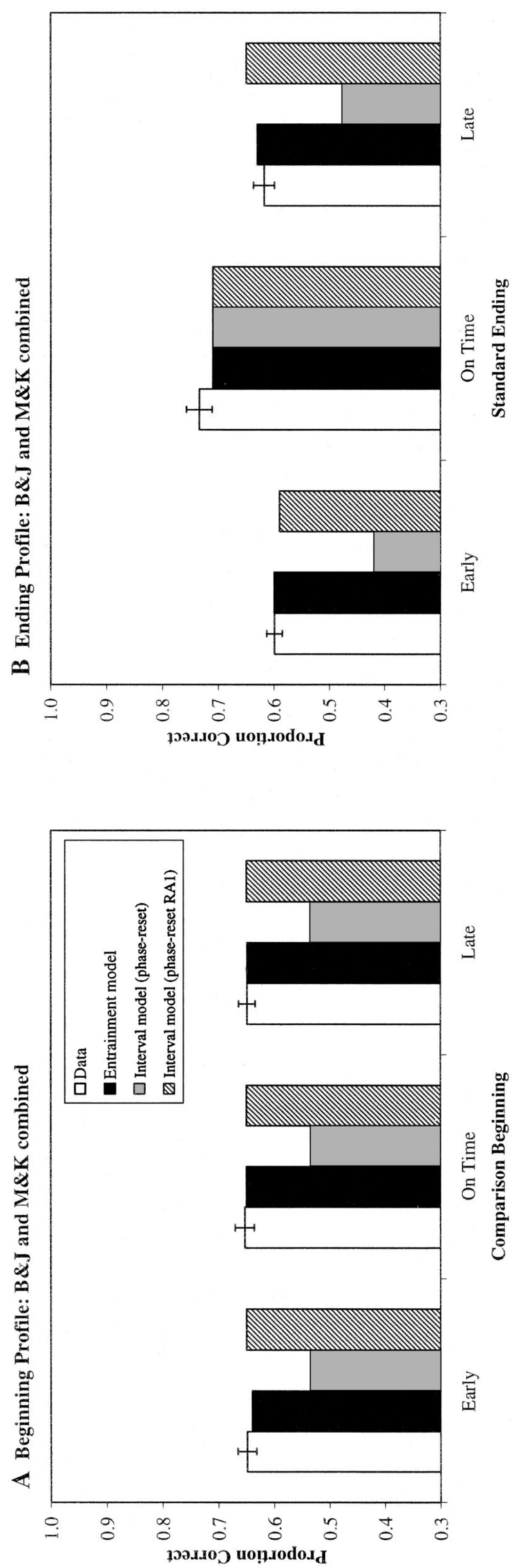

Figure 6. Mean proportion correct (with standard error bars) for participants in Experiment 1 as a function of comparison beginning (A) and standard ending (B). Quantitative predictions are shown for the best-fitting entrainment model ($W_\phi = 0.50$, $W_p = 0.45$), the cardinal phase-reset model, and the phase-reset running average (RA) 1 model ($W = 0.50$). B&J = Barnes and Jones (2000); M&K = McAuley and Kidd (1998).

errors we observed were typical of those reported by B&J. They are summarized by an 81-cell 3 (responses) × 3 (comparison durations) × 3 (standard durations) × 3 (comparison beginnings) confusion matrix (see Appendix B). This matrix reveals that lengthened (shortened) unexpected standards tended to be judged as shorter (longer) than they really were, whereas expected standards were less distorted by context. These errors are not consistent with the view that a quadratic ending profile (Figure 6B) results from uncertainties instilled in participants regarding the standard. According to an uncertainty hypothesis, errors occur because of increased effort to differentiate the standard from the preceding context IOIs and, hence, to determine “when” the standard appears in a sequence. This hypothesis implies that durations of all unexpected standards (early or late) should be underestimated because more effortful time judgments are typically underestimated (presumably due to an accumulator missing counts; e.g., Block & Zakay, 1997; Zakay, 1993). However, the confusion matrix data indicate that unexpected standards in Experiment 1 were both under- and overestimated. That is, the observed pattern of errors suggests a symmetrical assimilation of the standard into the preceding context: short (early-ending) standards were overestimated, whereas long (late-ending) standards are underestimated.

Quantitative model fits provide a more rigorous assessment of the Experiment 1 data. All model fits report root-mean-square errors of approximation (RMSEAs) between predicted and observed response proportions making up the average confusion matrices.[4] All RMSEA values are presented in Table 2 for Experiment 1; all fits used common initial conditions (P_0 = 600 ms, $\phi_0 = 0$). Note that the phase-reset model, which yielded the best qualitative fit of the four cardinal models, did not provide an exceptionally good quantitative fit to the full confusion matrix (see Table 2 and Figure 6). This model predicted too much assimilation because it allows no period correction. Consequently, all duration judgments are based on a remembered standard of 600 ms, and this yielded an overly pronounced expectancy profile.

In spite of these shortcomings, the phase-reset model remains a promising version of the interval model. Therefore, we corrected its limitations using an interval-averaging algorithm. We let P_i represent a running average of time interval codes for context and standard IOIs. In this variant of the cardinal phase-reset model, P_i is updated by taking a weighted average of the current IOI and P_i; the equation describing the running average is given by $P_{i+1} = W\,\text{IOI}_i + (1 - W)\,P_i$. This model, which has a single weight parameter, W, we refer to as the *phase-reset running average 1* (or phase-reset RA1) model. A formal assessment of the phase-reset RA1 model, allowing W to vary, indeed shows that this version of the phase-reset model improved the quantitative fit considerably (see Table 2). The estimate of W that minimized RMSEA was 0.5, corresponding to duration estimates of 570, 600, and 630 ms for standards of 540, 600, and 660 ms, respectively. Quantitative predictions of the phase-reset RA1 model appear in Figures 6A and 6B.

Finally, we compared the explanatory power of the phase-reset RA1 model with that of an entrainment model that permits partial correction of both phase and period. To determine whether an entrainment model with partial phase and period correction can improve the fit of the Experiment 1 data, we estimated values of W_p and W_ϕ that minimized RMSEA. These best-fit parameter values of phase and period correction corresponded to none of the four cardinal oscillator models, as shown in Figure 5 ($W_\phi = 0.50$, $W_p = 0.45$). The RMSEA value obtained for this fit was identical to that for the phase-reset RA1 model; not surprisingly, the entrainment model also predicted PC scores that closely matched observed values for both the beginning and ending profiles (see corresponding bars in Figures 6A and 6B).

An interesting outcome of the present modeling endeavor was our finding that the quadratic ending profile and the flat beginning profile can be explained in two different ways: either by a phase-reset model supplemented by a running average of intervals (phase-reset RA1) or by an entrainment model with only partial phase and period correction. Earlier, we discussed how these two approaches explain a quadratic accuracy profile. Here, we consider predictions of flat accuracy profiles. A flat profile is predicted by the phase-reset model because early and late beginnings of a comparison IOI simply serve to reset the oscillator to the beginning of its cycle. Consequently, C_{final} values do not differ for early, on-time, and late beginnings of the comparison IOI, and the net result is a flat beginning profile. By contrast, in an entrainment model equipped with partial phase and period correction, the effects of early and late stimulus onsets can linger. Moreover, in this view, it is necessary to carefully consider situations in which these lingering effects result in converging influences of phase and period correction on future alignments of the oscillator. A stimulus onset arriving early or late means that the oscillator either lags behind the stimulus sequence or is ahead of it, respectively. If phase is only partially corrected (rather than reset) in response to this onset, then the oscillator will still be lagging behind or be ahead, respectively, when the next stimulus onset occurs (with all

Table 2
Quantitative Fits of the Four Cardinal Models, the Entrainment Model, and the Interval-Averaging Model (Phase-Reset Running Average [RA] 1) to the Data From Experiment 1

Model	RMSEA
Full-reset	0.315
Phase-reset	0.132
Period-reset	0.217
No-reset	0.220
Entrainment	**0.028**
Phase-reset RA1	**0.028**

Note. Bolded values are those of the two best-fitting models. RMSEA = root-mean-square error of approximation.

[4] RMSEA was calculated by taking the square root of the mean summed squared difference between the observed and predicted response probabilities comprising the average confusion matrices, in which response probabilities varied between 0 and 1. RMSEA measures the average distance between the observed and the predicted values. In Experiments 1 and 2, there were 81 cells in the observed and predicted matrices that entered into calculation of RMSEA. In Experiment 3, in which the location of the time change was blocked, there were 54 cells making up the observed and predicted matrices. In Experiment 4, which involved six different comparisons in three different expectancy conditions (early, on-time, and late beginning of the comparison) but only two responses (*shorter* or *longer*), 18 points entered into the RMSEA calculation.

else constant). However, because the oscillator also period corrects, and because a beginning manipulation does not affect the duration of any to-be-judged time interval, the net effect of period correction in response to an unexpected beginning marker is to contribute to the amount of phase correction. That is, if an oscillator still lags behind following phase correction to an early onset, the remaining lag time may be essentially "picked up" by shortening the period (period correction) such that, together, partial phase and period correction bring the oscillator closer to synchrony. The reverse can happen for late onsets, but the net effect is the same: An oscillator's pulse peak precedes a late onset; therefore, the effect of lengthening the period again brings the oscillator into synchrony. Thus, if a change in stimulus timing does not affect the duration of a to-be-judged time interval, then period correction can contribute to phase correction. We refer to this as a *convergence effect.* In some cases, convergence is functionally equivalent to a phase reset, and it results in the prediction of a flat expectancy profile. This occurred for the beginning manipulation in Experiment 1 when both phase- and period-correction parameter values were approximately equal.

In summary, Experiment 1 revealed two expectancy profiles: a quadratic ending profile and a flat beginning profile. The ending profile, which replicated the B&J expectancy profile, depended primarily on variations in the standard duration, not on variations of a comparison's beginning time. The quadratic ending profile did not differ for blocked and randomized (roving) standard endings. An assessment of the four cardinal models showed that the phase-reset model (P_0 = 600 ms, $W_\phi = 1.00$, $W_p = 0$) provided the best qualitative explanation of these data. However, a phase-reset model augmented with a running average of intervals (phase-reset RA1), provided a much better quantitative fit to the Experiment 1 data. An equally good quantitative fit was found with an entrainment model having intermediate values of phase and period correction ($W_\phi = 0.50$, $W_p = 0.45$). Overall, we conclude that the phase-reset RA1 model is the most parsimonious interpretation of the Experiment 1 data because it has only one free parameter, compared with the two free parameters of the best-fitting entrainment model.

Experiment 2: The Role of a Silent Gap

One issue raised by Experiment 1 was whether participants can perceptually distinguish the standard from context intervals. Experiment 2 directly addressed this issue by inserting a longer interval (gap) prior to the standard, as shown in Figure 1C. Predictions about the role of such a prestandard gap vary. One of the best-fitting models of Experiment 1 (phase-reset RA1) implies that a gap may be included in a running average of context IOIs. But, given an averaging model, at least three possibilities for the role of a silent gap must be considered. First, because the prestandard gap is also a context IOI, one possibility is that the gap interval indeed is included in the running average. In this case, participants' estimates of the standard should vastly overestimate its true duration. This is predicted by the phase-reset RA1 model on the basis of a weighting parameter similar to that obtained for Experiment 1 (i.e., $W = 0.50$).

A second possibility is that the gap functions merely to differentiate the standard and does not enter into the running average. In this case, introducing a much longer interval (gap) prior to the standard should reduce confusion about the standard and enhance participants' ability to ignore the context IOIs. This second possibility is beyond the scope of the phase-reset RA1 model because it requires a second parameter permitting certain intervals to be excluded from the running average and others to be included. However, if an exclusion parameter is incorporated, then a two-parameter running average model, phase-reset RA2, can predict a quadratic ending profile for the design of Experiment 2. Furthermore, given the differentiation function of the gap, the predicted ending profile should be weaker (i.e., flatter) in Experiment 2 than in Experiment 1 (with correspondingly different best-fitting parameter estimates); this is because participants will use the gap to place greater weight on the standard than on the context IOIs preceding the gap.

The third possibility is that the gap permits participants to completely ignore the context IOIs. In this case, we would expect the data to conform to the predictions of the full-reset model, in which prior context plays no role. To thoroughly assess this last possibility, we incorporated a control condition that conformed to the conventional two-interval judgment task, in which no rhythmic context sequence is present (see Figure 1B). If participants in the experimental group indeed ignore context IOIs, then they should perform equivalently to control participants; moreover, both groups should perform more efficiently with unexpected standards than did participants in Experiment 1.

However, a totally different outcome is possible if participants do not engage in interval averaging of context IOIs. Instead, a participant's perception of a standard's duration may be based neither on averaging context IOIs nor ignoring them but, rather, may depend on momentary values of the persisting period of an oscillator entrained to the context rhythm. In this case, experimental participants should continue to be affected by context IOIs despite their clear separation from the standard, especially when the gap is a multiple of the recurrent context IOI (e.g., a 1,200-ms gap). Recall that the best-fitting entrainment model examined in Experiment 1 explained the data as well as the phase-reset RA1 model; application of this entrainment model to Experiment 2 required no changes. From an entrainment perspective, the gap simply introduces a "missing beat" through which a self-sustaining oscillator continues. Thus, the performance of experimental participants should be nearly identical to that of participants in Experiment 1 and should differ from that of the control group.

Method

Participants. Thirty-one students were recruited in the manner described for Experiment 1, and they were randomly assigned to one of two context conditions (experimental, $n = 14$; control, $n = 17$).

Design. A 2 × 3 × 3 × 3 mixed-factorial design crossed two context conditions (experimental, control) with three standard endings (early, on time, late), three comparison beginnings (early, on time, late) and three levels of the comparison IOI (shorter, same, longer). The single between-subjects variable was context condition. The remaining three (within-subject) variables varied randomly over trials. Auditory sequences in the experimental condition were identical to those of Experiment 1, with one exception: The sixth and seventh context IOIs were replaced by one of 1,200 ms; this is referred to as a *prestandard gap* (see Figure 1C). The control group received only a standard IOI and a comparison IOI on each trial; the standard followed the warning tone by 1,200 ms (see Figure 1B). The ISIs used in both the control and the experimental groups were identical to those used in Experiment 1.

Stimuli and apparatus. These were identical to those of Experiment 1.

Procedure. The procedure was identical to that of Experiment 1, with two exceptions. First, following practice trials involving only standard and comparison intervals, participants in the experimental group received practice sequences that included context sequences with prestandard gaps. These participants were explicitly told to ignore all context IOIs preceding the gap and to concentrate fully only on standard and comparison intervals; they were informed that attending to the context pattern would distract them. Second, control participants received only practice trials with standard and comparison IOIs.

As in Experiment 1, the test session comprised six blocks of 54 trials. Participants in both experimental and control conditions heard each combination of standard ending and comparison beginning six times in each 54-trial block. Equal numbers of shorter, same, and longer comparison IOIs occurred every 27 trials. Overall, there were 12 judgments for each condition.

Results and Discussion

Figure 7 reports mean PC scores (averaged over comparison IOI) for the experimental and control conditions for the three standard endings and comparison beginnings, respectively. A mixed-factorial ANOVA on overall PC indicated no significant difference between the overall levels of performance of the control and experimental groups, $F(1, 29) = 2.05$, $MSE = 0.079$, $p > .15$. Consistent with Experiment 1, we observed a main effect of standard ending, $F(2, 58) = 22.37$, $MSE = 0.01$, $p < .01$, but no main effect of comparison beginning ($p > .60$). However, a significant Context Condition (experimental vs. control) × Standard Ending interaction obtained, $F(2, 58) = 10.45$, $MSE = 0.01$, $p < .01$. The results for control participants, who experienced no contextual rhythm, revealed a flat ending profile, with mean PCs for early, on-time, and late standard endings of .57, .60, and .56, respectively. Experimental participants, in contrast, produced a quadratic ending profile, with mean PCs for early, on-time, and late standard endings of .56, .72, and .59, respectively. Post hoc analysis of the experimental participants' data only showed that early and late performance for standard endings was significantly lower than on-time performance ($p < .01$), whereas early and late conditions did not significantly differ. In addition, the mean PCs for experimental participants did not differ significantly from those in Experiment 1, in which no silent gap appeared. RMSEA values for tested models are reported in Table 3.

On the basis of an interval model of duration perception, it is evident from the Experiment 2 data that participants did not include the gap in the running average. If they did so, then they should have overestimated the standard duration in all conditions, which was not the case. Indeed, the quantitative fit of the phase-reset RA1 model to the Experiment 2 (experimental group) data (for $W = 0.50$) was poor, producing subjective estimates of the standard of 720, 750, and 780 ms for standards ending early, on time, and late, respectively (see corresponding bars in Figure 7). In addition, it is clear that introducing a gap did not enable participants to ignore the context sequence. If they had ignored the context sequence, then the performance of the experimental group would have been similar to that of the control group, which was not the case.

To test the possibility that participants excluded the gap from the running average, we assessed the phase-reset RA2 model, which includes an *exclusion parameter*, ε. This parameter is sensitive to differences between each IOI and the standard IOI (e.g., Hellstrom, 1985). Suppose that T = standard IOI, and ΔT is the absolute time difference between each IOI and the standard IOI. On the basis of this notation, values of $\Delta T/T$ that are equal to or exceed a critical relative difference (ε) are excluded from the averaging process; thus, ε varies such that best-fitting values depend on sequence structure and task. In this way, durations longer or shorter than the standard by some estimated amount will not affect the standard's perceived duration. The phase-reset RA2 model greatly improved the fit to Experiment 2 data (for $W = 0.50$, $\varepsilon = 0.5$); see Table 3 and corresponding bars in Figure 7.

Experimentally, we found no evidence that introducing a gap weakened participants' reliance on the context sequence. Indeed, the quadratic ending profiles in Experiments 1 and 2 were essentially equivalent, yielding similar estimates of the running average weight (the best estimate of W was 0.5 in both experiments).[5] The lack of difference between context conditions in the two experiments is somewhat surprising from the perspective of an interval account because it suggests that although the gap was excluded from an averaging process, it nonetheless failed to differentiate and weaken context effects. An alternative interpretation comes from an entrainment perspective: A self-sustaining oscillator, with P_0 = 600 ms, is assumed to persist through the prestandard gap, having a time span that is twice its period. Supporting this interpretation, the quantitative fit of the entrainment model to Experiment 2 data was excellent and revealed similar estimates of phase and period correction to those of Experiment 1 ($W_\phi = 0.58$, $W_p = 0.42$).

Overall, the consistency of outcomes across Experiment 2 (experimental participants) and Experiment 1 suggests that the prestandard gap in Experiment 2 neither enabled participants to ignore context IOIs nor weakened the effect of the context on the perceived duration of the standard, as might be expected on the basis of a one-parameter interval model of the internal clock. The addition of a second (exclusion) parameter to the interval model accommodated the data from Experiment 1 (as in the phase-reset RA2 model). The phase-reset RA2 model performed as well as the best-fitting two-parameter entrainment model.

Experiment 3: A Faster Context Rate

Together, Experiments 1 and 2 illustrate three important points: (a) Manipulations of standard endings produce quadratic accuracy profiles, whereas manipulations of comparison beginnings do not; (b) beginning and ending accuracy profiles are unaffected by the presence of a gap introduced prior to the standard; and (c) manipulations of comparison beginnings appear to produce flat expectancy profiles. Theoretically, variants of both interval and entrainment models can account for all three findings, but they do so in conceptually different ways.

An interval model explains the ending profiles in the presence of a gap using two estimated parameters: a running average weight

[5] Another explanation for the expectancy profile, suggested by one reviewer (especially for the data of Experiment 1), is that participants shifted from a prospective to retrospective mode of listening after an unexpected standard occurred. The design of Experiment 2, in which a prestandard gap signaled the standard, ensured a common prospective attentional set for all standards. The findings and model fits in the current paragraph argue against such an interpretation of either experiment.

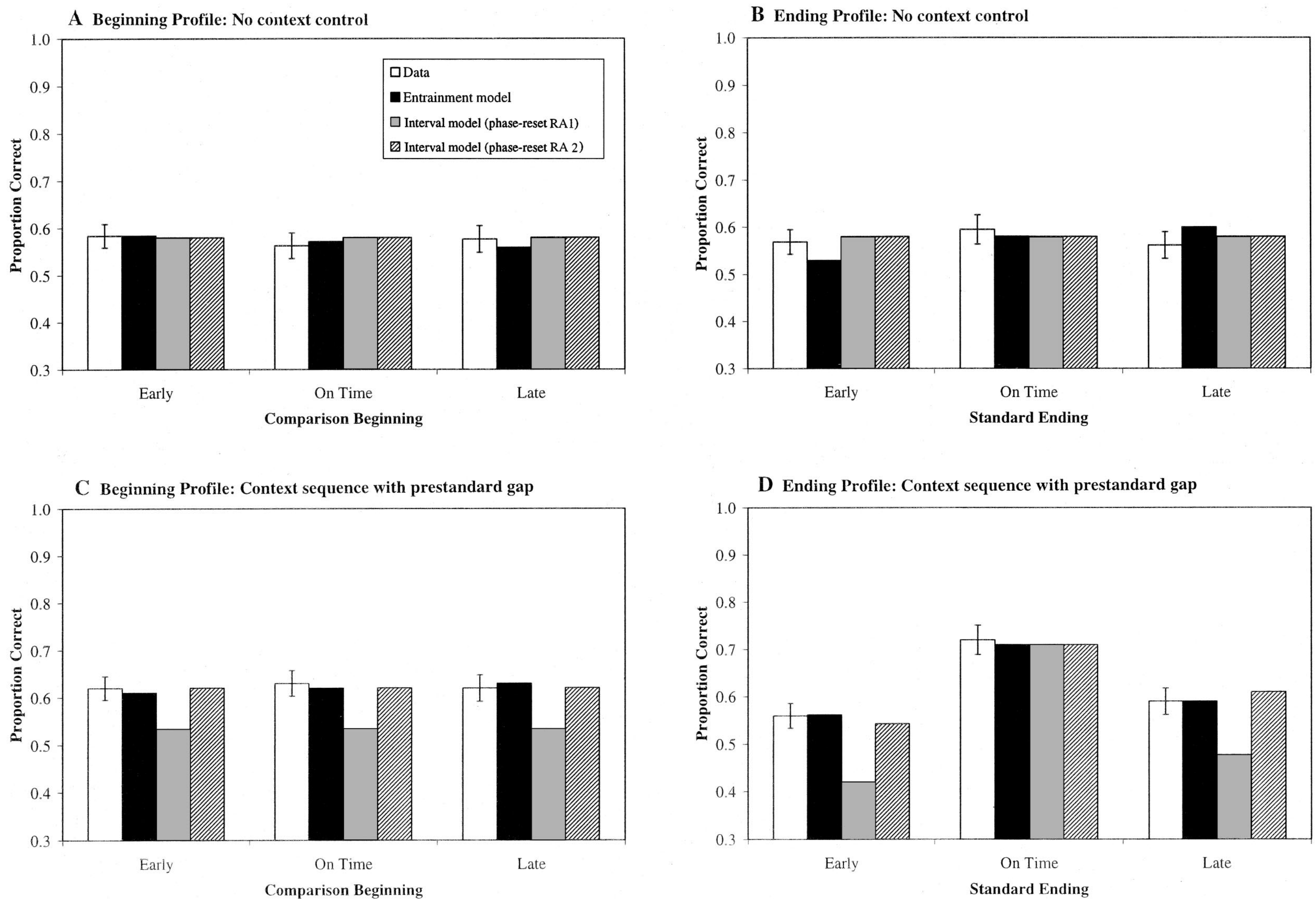

Figure 7. Mean proportion correct (with standard error bars) for participants in Experiment 2 as a function of comparison beginning (A and C) and standard ending (B and D). Control participants are shown in Panels A and B; experimental participants are shown in Panels C and D. Corresponding quantitative predictions are included for the best-fitting entrainment model (W_ϕ = 0.58, W_p = 0.42) and the interval model variants (phase-reset running average [RA] 1 and phase-reset RA2; W = 0.50).

Table 3
Quantitative Fits of the Four Cardinal Models, the Entrainment Model, and Interval-Averaging Model Variants (Phase-Reset Running Average [RA] 1 and 2) to the Data From Experiment 2

Model	RMSEA	
	Experimental	Control
Full-reset	0.338	0.021
Phase-reset	0.106	0.021
Period-reset	0.182	0.119
No-reset	0.183	0.286
Entrainment	**0.015**	**0.042**
Phase-reset RA1	0.270	0.021
Phase-reset RA2	**0.023**	**0.021**

Note. Bolded values are those of the two best-fitting models. RMSEA = root-mean-square error of approximation.

(W), which combines the context IOIs with the standard in memory, and an exclusion parameter (ε), which permits IOIs substantially longer or shorter than the standard to be excluded from the running average. This model explains the flat beginning profile associated with the manipulation of comparison beginnings by assuming that $W_{\phi} = 1.00$ for all stimulus onsets; this instantiates automatic phase resetting to any stimulus marker.

An entrainment model explains ending profiles in the presence of a gap, also using two parameters (a phase-correction parameter and a period-correction parameter). According to this model's explanation of the quadratic ending profile, the oscillator period (akin to a working memory) persists through the gap and partially corrects its period to unexpected standards, generating systematic errors in responding to the comparison. The flat beginning profile is predicted to arise from a convergence effect resulting from partial period correction (in combination with partial phase correction); in some cases, the convergence effect mimics phase resetting when a standard duration does not change. This occurred in Experiment 2 and Experiment 1 when estimates of W_{ϕ} and W_p were near 0.50 (as shown in Figure 5).

Experiment 3 continued to distinguish between the interval and entrainment interpretations of rhythmic context effects by changing the rate of the context sequence. Both preceding experiments supplied context IOIs that were close, if not equivalent, in duration to the presented standard. In Experiment 3, the context rate involved IOIs that were far (on an interval scale) from the standard. The design of Experiment 3 halved all pregap IOIs but retained the same three standard durations (540, 600, and 660 ms) and the gap interval (1,200 ms), as illustrated in Figure 1D. Note that although 300-ms IOIs are a significant interval distance from 600 ms, on a ratio scale they are harmonically related to 600 and 1,200 ms.

An interval model offers at least two possible outcomes for Experiment 3. One possibility is that the 300-ms context IOI will be included in the running average, as predicted by the phase-reset RA1 model. In this case, all standard durations should be underestimated. Perhaps a more plausible interpretation is offered by the phase-reset RA2 model (which provided the best interval explanation of both Experiments 1 and 2). This model predicts relatively little impact of the 300-ms context sequence or of the 1,200-ms gap because both would be excluded from the running average (or, alternatively, assigned low averaging weights) because of the absolute values of their interval distances from the three standard durations. In either case, we would expect to find both flat beginning and ending profiles, as with the control data of Experiment 2. The full-reset model makes the identical prediction.

By contrast, if participants' perception of the duration is not based on an interval clock but rather involves an entrainable oscillator, then they should continue to be affected by the context sequence in a manner similar to that found in the first two experiments. From an entrainment perspective, doubling the rate of the context sequence should still result in a quadratic expectancy profile because, in terms of relative phase, the three standard durations still end early, on time, and late relative to the context rhythm. On the basis of parameter estimates obtained from the best-fitting entrainment model from Experiments 1 and 2, in this experiment we expected to see a quadratic ending profile and a flat beginning profile, similar to those we observed in the first two experiments.

Finally, although our main interests in Experiment 3 concerned sequence rate, here we also address a prediction about beginning manipulations common to both interval and entrainment models. Both views predict that manipulations of beginning times of standards should provide flat accuracy profiles, similar to those found for manipulations of beginnings of comparisons. Accordingly, in one condition of Experiment 3, we manipulated the beginning times of the standard, holding standard duration constant. Our aim was to confirm that the flat accuracy profile found for comparison beginnings in Experiments 1 and 2 was not specific to comparison intervals but holds more generally for the beginning of any to-be-judged time interval.

Method

Participants. Forty-one students were recruited in the manner described for Experiment 1. They were randomly assigned to one of two expectancy conditions (beginning of standard varied, ending of standard varied).

Design. A 3 × 2 × 3 mixed-factorial design crossed three expectancy levels (early, on time, late) with two locations (beginning vs. ending of standard) for each of three comparison durations (shorter, same, longer). The location of the expectancy manipulation was a between-subjects variable; the levels of other variables varied randomly from trial to trial, as in the previous two experiments. All sequences comprised six context IOIs, followed by a standard IOI and a comparison IOI, as shown in Figure 1D. The first five IOIs were 300 ms; the value of the sixth IOI (the gap) depended on the condition. The two location conditions were distinguished by manipulations of beginning (Group B) and ending (Group E) times of the standard IOI.

Group B (n = 12; beginning varied condition) heard the fast induction sequences, with the sixth context IOI varied from trial to trial, so that the beginning of a (600-ms) standard was early (gap = 1,140 ms), on time (gap = 1,200 ms), or late (gap = 1,260 ms) relative to the implied induction rate of 300 ms. The standard IOI was constant at 600 ms. For Group E (n = 29; ending varied condition), the sixth context IOI was fixed at 1,200 ms, and the standard ended early (standard IOI = 540 ms), on time (standard IOI = 600 ms), or late (standard IOI = 660 ms), as in Experiment 2. For both location conditions (Group B, Group E), the ISI between the standard and the comparison was constant at 1,200 ms. All other aspects of the design were identical to those of the context conditions in Experiments 1 and 2.

Stimuli and apparatus. These were identical to those of the previous experiments.

Procedure. These were identical to those of the previous experiments.

Results and Discussion

Figure 8 reports mean PC (averaged over comparison IOI) for the beginning and ending groups as a function of timing condition (early, on time, late). Consistent with the results of Experiments 1 and 2, two different accuracy profiles appear: a flat accuracy profile for early, on-time, and late beginning of the standard (mean PC = .78) and a quadratic ending profile. Mean PCs for early, on-time, and late standard endings were .57, .70, and .62, respectively. An ANOVA confirmed the main effect of time-change location (beginning vs. ending), $F(1, 39) = 18.95$, $MSE = 0.028$, $p < .01$, with participants who received only the beginning time manipulation performing better overall than those who received the ending manipulation. This difference was probably due to the constancy of the standard's duration in the former group. Timing condition (early, on time, late) was also significant, $F(2, 78) = 7.45$, $MSE = 0.006$, $p < .01$, and there was a Timing Condition × Time-Change Location interaction, $F(2, 78) = 5.02$, $MSE = 0.006$, $p < .01$. Together, these findings confirm that flat accuracy profiles are associated with beginning time manipulations of both standards and comparison intervals, whereas quadratic profiles are associated with standard endings. We return to this point in the General Discussion.

Quantitative fits of the entrainment model and the phase-reset RA2 model, based on the full confusion matrix, appear with the data in Figure 8.[6] Relevant RMSEA values are given in Table 4. As in the previous experiments, the entrainment model accurately predicted a quadratic ending profile and a flat beginning profile. Estimates of phase and period correction were 0.37 and 0.50, respectively, using an initial oscillator period of 300 ms (i.e., for sequence rate). In contrast, for the same initial conditions, the phase-reset RA2 model erroneously predicted a flat ending profile in addition to a flat beginning profile. It failed to predict a quadratic ending profile because it excluded both the 300-ms context IOI and the 1,200-ms gap from the running average (for $\varepsilon = 0.5$). If the 300-ms IOI is not excluded (using a larger ε estimate) but, rather, contributes to the running average, then the phase-reset RA2 model would predict that all three standard durations will be underestimated, which did not occur.

Up to this point, the modeling results lead to the conclusion that the best quantitative explanation for the data from the three experiments is an oscillator with partial phase and period correction. For the first two experiments, it was not possible to distinguish this model from an interval model with two parameters (a running average weight and a parameter that selectively permits some of the context intervals to be ignored in the computation of the referent interval stored in memory). However, for Experiment 3, the phase-reset RA2 model failed when it either excluded or included the 300-ms context IOI from the running average.

One further modification, however, allows the interval model to explain the Experiment 3 data. If it is supposed that participants infer a 600-ms referent interval (standard) by concatenating independently stored codes for two successive 300-ms context IOIs, then from an interval perspective it is possible to accurately predict a quadratic ending profile and a flat beginning profile. This assumes that the resultant coded intervals (of 600 ms) contribute to the running average. We assessed this last possibility by adding a *concatenation* operation to the phase-reset RA2 model. This variant is termed *phase-reset RA3.*

With concatenation, two adjacent context IOIs, initially excluded from averaging because of their difference from a standard, may be included in the running average if they combine to form a single IOI that is sufficiently similar to the duration of that standard. This proposal is related to research on production of successive time intervals by Vorberg (1978; Vorberg & Hamburg, 1984), who concluded that serial intervals of equal length are generated by concatenation rather than being synchronized with a periodic internal beat.[7] Similarly, in perception, Keele et al. (1989) proposed that individuals can generate an apparent beat by concatenating separate codes of successive intervals where such intervals are not further synchronized with a beat source. In the present design, a correct concatenation (i.e., 600 ms) required not only an additional operation but also a limit to the span of the operator (i.e., two 300-ms IOIs). Finally, with incorporation of these operations, the phase-reset RA3 concatenation model provided a fit that was superior to that of the phase-reset RA2 model, as indicated by RMSEA values (Table 4), and which did not differ substantially from that of the best-fitting entrainment model. Fits of all three models (phase-reset RA2, phase-reset RA3, and entrainment) are shown with the data in Figure 8.

Experiment 4: Large Beginning Shifts

A final experiment further distinguished the interval and entrainment approaches by directly testing the phase-resetting assumption of the interval model. In the previous experiments, we consistently observed flat accuracy profiles associated with manipulations of beginning times of standard or comparison intervals. In Experiment 4, we probed the limits of those profiles to directly address the phase-resetting assumption of the interval approach. We included larger shifts in the beginning time of the comparison interval. A defining property of the interval approach is full phase correction (i.e., resetting) to unexpected onset markers of comparison intervals. The phase-resetting assumption permitted the correct prediction of a flat beginning profile in Experiments 1–3.

Entrainment models also predicted a flat beginning profile in Experiments 1–3 because of the convergence effect involving partial phase and period correction; this simulated a phase-reset response to unexpected beginnings. In particular, this response occurred in response to our manipulations of the timing of beginning markers in Experiments 1–3 when phase and period correction were approximately equal and differences in beginning times

[6] The confusion matrix data also permitted us to eliminate an explanation of these data (and those of the earlier experiments) on the basis of competition of response labels whenever an unexpected standard is labeled shorter (longer) than the context and the comparison is labeled longer (shorter) than the standard. In the present design, both the standard and the comparison intervals were always longer than the context intervals and always shorter than the prestandard gap. Thus, there can have been no competition associated differential response labeling in Experiment 3.

[7] Concatenation of successive but different time intervals is implied in several interval models (Vorberg & Wing, 1996, pp. 248–258). Such concatenation is also a common serialization operation for successive codes in the perception of serial patterns (e.g., Restle, 1970; see Jones, 1981, for a review). The phase-reset RA3 model requires this operation as an additional computational step to ensure that exactly two context IOIs contribute to the running average.

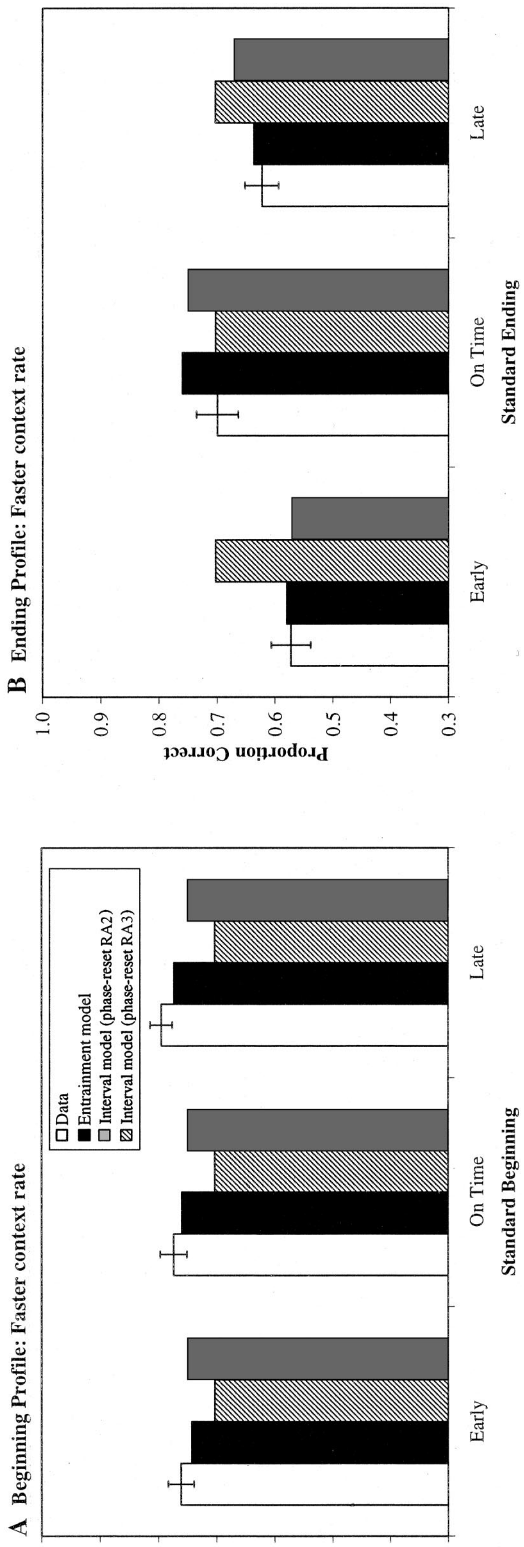

Figure 8. Mean proportion correct (with standard error bars) for participants in Experiment 3 as a function of standard beginning (A) and standard ending (B). The corresponding quantitative predictions are reported for the best-fitting entrainment model ($W_\phi = 0.37$, $W_p = 0.50$) and the interval model variants (phase-reset running average [RA] 2 and phase-reset RA3).

Table 4
Quantitative Fits of the Entrainment Model and Interval-Averaging Model Variants (Phase-Reset Running Average [RA] 1, 2, and 3) to the Data From Experiment 3

Model	RMSEA
Entrainment	**0.029**
Phase-reset RA1	0.082
Phase-reset RA2	0.082
Phase-reset RA3	**0.034**

Note. Bolded values are those of the two best-fitting models. RMSEA = root-mean-square error of approximation.

were relatively small (e.g., 60 ms). Empirically, small differences in beginning onset times may render the assessment of significant failures in phase correction of a hypothetical entrainment process difficult. This issue is less relevant in the case of ending manipulations because, in addition to producing a phase shift, the ending manipulation affects the duration of a to-be-remembered standard interval.

To better assess the effect of unexpected beginnings, we adapted the design of Experiment 2 by (a) holding constant the standard duration and varying comparison ending times in smaller steps; (b) varying the beginning time of the comparison interval in larger steps (±180 ms vs. ±60 ms); and (c) giving participants two rather than three identification responses, permitting a straightforward signal-detection analysis. If phase is only partially corrected, as predicted by the best-fitting entrainment model, then increasing the magnitude of the beginning manipulation may reveal effects of this variable that were not evident in the previous experiments. However, if the best explanation of the data in Experiments 1–3 involves the phase-resetting assumption of the interval approach, as incorporated into the cardinal phase-reset model and its variants (RA1, RA2, RA3), then C_{final} will necessarily register the same difference between the standard and a given comparison interval, regardless of when that comparison begins. In short, as in prior experiments, for Experiment 4, all of these models predict null effects of manipulations of comparison beginning times, regardless of their magnitude.

Method

Participants. Fourteen participants were recruited in the manner described for Experiment 1, with one exception: In addition to recruiting 6 Ohio State University students, we recruited 8 students from Bowling Green State University as well.

Design. A 3 × 6 repeated measures design crossed three levels of comparison beginning (−180, 0, and 180 ms) with six comparison durations (550, 570, 590, 610, 630, and 650 ms). Both variables were within-subject variables and varied randomly over trials. On each trial, participants judged the duration of the comparison IOI relative to the standard, responding *shorter* or *longer*. The sequences were identical to those of Experiment 2 (i.e., all contained a prestandard gap of 1,200 ms), with two exceptions: (a) The standard duration was always 600 ms, and (b) the three ISIs were 1,020, 1,200, and 1,380 ms, corresponding, respectively, to the 180-ms early, on-time, and 180-ms late comparison beginning conditions (see Figure 1E).

Stimuli and apparatus. These were identical to those in the previous experiments.

Table 5
Mean Observed and Predicted Points of Subjective Equality (PSEs) and Duration-Discrimination Thresholds (Just-Noticeable Differences; JNDs), in Milliseconds, for Experiment 4

	Average participant		Entrainment model		Interval model	
Beginning	PSE	JND	PSE	JND	PSE	JND
Early	647.5 (10.2)	56.7 (10.3)	627	30	600	30
On-time	601.6 (3.0)	24.3 (4.7)	600	30	600	30
Late	587.7 (4.1)	29.2 (5.3)	573	30	600	30

Note. Standard errors appear in parentheses. Estimates of phase correction and period correction for the best-fitting entrainment model were 0.8 and 0.05, respectively.

Procedure. Participants received a block of 12 practice trials (with context IOIs), followed by four test blocks, each with 54 trials. On each trial, participants judged the duration of the comparison relative to its standard by responding *shorter* or *longer*. They were explicitly told to ignore all context IOIs preceding the gap and to concentrate only on the standard and comparison intervals. Equal numbers of each level of the comparison IOI occurred in each block.

Results and Discussion

Points of subjective equality (PSEs) and duration-discrimination thresholds (just-noticeable differences; JNDs) were calculated for each participant in each condition, using the following algorithm. Individual estimates of PSE and JND were obtained by first transforming proportions of *short* responses for the six comparison intervals to *z*-coordinates and then fitting a straight line though the transformed points, as described by Macmillan and Creelman (1991, pp. 219–220). PSE is the duration judged *shorter* 50% of the time. JND is half of the stimulus distance between the 25th and 75th percentiles, corresponding to $d' = 0.675$.

Average PSE scores appear in Table 5. A repeated measures ANOVA on participant PSE scores revealed a significant main effect of comparison beginning, $F(2, 26) = 22.46$, $MSE = 611$, $p < .01$. Comparisons that began on time produced subjective estimates of the standard that were approximately equal to the standard duration (mean PSE = 601 ms), whereas comparisons that began unexpectedly early or late produced subjective over- and underestimates, respectively, of the standard (mean PSE of 647.6 ms vs. mean PSE of 587.7 ms). All pairwise contrasts were significant ($p < .05$).

The mean JND, averaged over participants, is shown for each of the three comparison beginning times in Table 5. A repeated measures ANOVA on participant JND scores revealed a significant main effect of comparison beginning, $F(2, 26) = 6.31$, $MSE = 677$, $p < .01$. However, only the pairwise contrasts between the early and on-time conditions and the early and late conditions were significant (ps $< .05$). Thresholds for early beginnings were almost double those for on-time and late beginnings, with the thresholds observed in the later two conditions being roughly equivalent. Overall, these results indicate that discrimination thresholds are higher when an unexpected comparison arrives very early but not when it arrives unexpectedly late.

Table 5 also presents predictions of the best-fitting interval and entrainment models. Although the two classes of model do not differ with regard to predicted JNDs, they differ markedly regarding PSE predictions, as well as the overall quality of the quantitative fits. The cardinal phase-reset model and its variants (RA1, RA2, RA3) predict null PSE differences, whereas the best-fitting entrainment model predicts that PSEs for unexpected comparisons that arrive very early or very late will manifest systematic distortions as a result of large phase shifts. Indeed, as predicted by the best-fitting entrainment model, early beginning comparisons produced PSEs greater than 600 ms (overestimates), and late beginning comparisons produced PSEs less than 600 ms (underestimates). However, two aspects of these data require discussion. First, overall, the quantitative fits to the Experiment 4 data (for both classes of model) were worse than those in the previous three experiments.[8] Second, poorer fits can be attributed to two clear asymmetries associated with early versus late conditions, which neither class of model accounts for. Although the entrainment model explained the general pattern of results, it does not predict the finding that early beginnings have a greater impact on both PSEs and JNDs than do late beginnings. We return to both points in the General Discussion.

The findings from Experiment 4 offer a qualification to conclusions from Experiments 1–3 regarding the null effects of manipulating beginnings of comparison time intervals in rhythmic contexts. Relatively large changes in onset times lower discriminability, especially for comparisons arriving unexpectedly early. In addition, both very early and very late beginning times significantly shift a participant's PSE. Together, these findings lead to the primary conclusion from Experiment 4, namely that models that incorporate phase resetting (full phase correction) do not accommodate the negative effects of relatively large expectancy violations on performance (e.g., $/\phi/ > 0.25$ of the standard). The limits of phase correction become evident only with large deviations from an expected time. This is notable because small time deviations (e.g., $/\phi/ < 0.15$) are in common use. Nevertheless, our data suggest that small time deviations do not challenge

[8] One hypothesis, suggested by Richard Ivry, proposes code confusion as a result of participants encoding the ISI preceding the comparison as *shorter* (1,140 ms) or *longer* (1,260 ms). Presumably, such codes would interfere with encoding of either the preceding standard or the following comparison. However, this does not explain the asymmetries observed in both the PSEs and JNDs in Experiment 4. In addition, it seems unlikely that participants would have encoded these ISIs in such a fashion, given that all ISIs were longer than the two time intervals surrounding them.

the limits of a participant's phase-correction abilities and, therefore, fail to reveal limits of phase-resetting models.

General Discussion

This article introduces a general theoretical framework for modeling effects of rhythmic context on the perceived timing of auditory events. The framework builds on the idea that a single oscillator can be viewed from different theoretical perspectives (e.g., interval, beat-based, or entrainment timers), depending on its parameters. Assuming linear phase and period correction, the values of parameters governing phase and period correction result in a family of oscillator models. Some members of this family are capable of entrainment, because they permit both phase and period correction, whereas others cannot entrain, because they do not permit both kinds of adaptivity. We have evaluated four cardinal models, along with variants, that highlight limiting cases of phase correction and period correction. A central theme in this experimental evaluation concerned whether an oscillator's phase or period (or both) responds adaptively to manipulations of beginnings versus endings of to-be-judged time intervals in different rhythmic contexts.

Across four experiments, we tested numerous predictions about accuracy patterns associated with specific failures of an oscillator to adapt. We discovered two different accuracy profiles: one suggesting significant adaptive constraints on responding to ending time manipulations and another suggesting relatively efficient adaptive responding to manipulations of beginning times. Those experiments that manipulated the ending time of the standard interval (Experiments 1–3) revealed a quadratic expectancy profile in accuracy, indicating poor performance when a standard ended early or late relative to an expected ending time. In contrast, in three of the four experiments, when the beginning time of an interval (standard or comparison) was manipulated, a flat beginning profile was observed, revealing uniformly good performance when either a standard (Experiment 3) or a comparison (Experiments 1 and 2) began early, on time, or late.[9] Both kinds of profiles were robust in the first three experiments. The ending profile was unchanged over modest variations in the beginning time of the comparison interval and across different prestandard gaps (Experiments 1 and 2). This inverted-U ending profile also remained strong across variations in the beginning time of the standard itself and under a harmonic rate change (Experiment 3). Related research has shown that the quadratic ending profile is eliminated when the context IOI is not harmonically related to the expected standard duration (Barnes & Jones, 2000).

The story of Experiments 1–3 appears straightforward: Manipulations of ending times of standards yield a quadratic expectancy profile consistent with incomplete period correction, whereas manipulations of beginning times of any to-be-judged interval yield a flat accuracy profile that is consistent with complete phase correction. But appearances are deceiving in the latter case. First, an entrainment model can explain both profiles using parameter values that suggest that neither phase nor period correction is complete. Second, Experiment 4 showed that the difference between ending and beginning profiles is a matter of degree. Increasing the magnitude of the beginning time manipulation relative to the ending manipulation significantly affected measures of both PSE and JND. Comparisons that arrived unexpectedly early and late produced subjective over- and underestimates, respectively, of the standard duration. Temporal acuity (JNDs) was especially poor for comparisons that arrived early relative to on-time and late comparisons. Two important differences between the distinctive (i.e., nonflat) profiles associated, respectively, with ending and beginning times were that (a) distinctive ending profiles emerged with smaller time changes than did distinctive beginning profiles, and (b) distinctive ending profiles were symmetrical (quadratic), whereas nonflat beginning profiles were asymmetric, at least with respect to JNDs. The latter findings extend those of M&K.

Within our general framework, we considered two different explanations of these findings using a single oscillator: one based on an interval view of timing and the other based on an entrainment view of timing. These views represent the poles of a current theoretical debate on the nature of the functional mechanism(s) underpinning the perception of time. Two of the four cardinal oscillator models (phase-reset and full-reset) capture defining properties of prototypical interval approaches to timing because they propose complete phase correction (Church & Broadbent, 1990; Gibbon, 1977; Gibbon et al., 1984; Treisman, 1963). The other two cardinal models (no-reset and period-reset) studied by previous researchers (Keele et al., 1989; Pashler, 2001; Schulze, 1978) capture defining aspects of the typical beat-based approach to timing because they propose no phase correction. Only one of the four cardinal models, the phase-reset model, correctly predicts the general shape of both expectancy profiles, namely the quadratic ending profile and the flat beginning profile observed in Experiments 1–3. This model instantiates the classic interval-timer view. However, its explanatory superiority is limited to a qualitative description of expectancy profiles.

It turns out that the phase-reset model overpredicted assimilation, resulting in an overly sharp expectancy profile. A rigorous assessment of this model, using quantitative fits to the confusion matrix, ruled out not only the phase-reset model but all four cardinal models, suggesting that constraining W_p and/or W_ϕ to binary values (0, 1) was inappropriate (Experiment 1). Consequently, in further modeling efforts, we (a) allowed period-correction (W_p) and phase-correction (W_ϕ) parameter values to vary and (b) introduced new parameters to modify the phase-reset model, leading to several phase-reset variants (RA1, RA2, RA3), all based on an averaging algorithm. Taken together, the most parsimonious fits of the data from Experiments 1–3 were given by entrainment models with partial phase and period correction. For these studies, estimates of phase and period correction were in the middle of a parameter space (cf. Figure 5) defined by binary values of these parameters ($W_\phi = 0.37–0.58$, $W_p = 0.45–0.50$); however, for Experiment 4, the estimate of phase correction was much larger ($W_\phi = 0.80$) than that found in previous experiments, whereas the estimate of period correction was much smaller ($W_p = 0.05$). Assessments of these models lead to several conclusions. One is that nonadaptive oscillators instantiating beat-based models do not adequately describe our findings. It is noteworthy that we can rule out the no-reset model, which involves a rigid oscillator, because this conclusion is consistent with previous research that has ques-

[9] An additional experiment, not reported here, confirmed that modest time changes in the beginning of the standard interval yield a flat expectancy profile.

tioned the generic beat-based approach (Keele et al., 1989; Pashler, 2001). However, beat-based models represent only a special case of the general class of entrainment models. In other words, by ruling out a rigid oscillator, we do not eliminate all oscillator timing models involving adaptive (i.e., self-sustaining) oscillators. The framework we have developed considers the range of possibilities for an oscillator timekeeper with linear phase and period correction.

Our analyses also lead to a second conclusion—namely, that the most successful quantitative explanations of our findings are given by variants of the cardinal phase-reset model and by the two-parameter adaptive oscillator model. In comparing these two classes of model, we showed that certain interval models based on running averages of context intervals performed better than ones based on simple averaging (e.g., Drake & Botte, 1993; Hirsh, Monahan, Grant, & Singh, 1990). A model based on the running average of context intervals can precisely predict quadratic ending and flat beginning accuracy profiles in different tasks. Moreover, additional parameters substantially improved the quantitative fits of this class of model to the data from Experiments 1–3. The phase-reset RA1 model, which incorporates a running average of context plus standard, improved the fit to the Experiment 1 data, rendering this model indistinguishable from the best-fitting entrainment model. This interval model failed, however, to provide an adequate quantitative account of Experiment 2, which added a silent prestandard gap. A further modification (phase-reset RA2), permitting substantially shorter or longer intervals to be excluded from the running average, provided excellent fits to the data from both Experiments 1 and 2. This model is comparable to the best-fitting entrainment model in that both operate in real time and have two free parameters. Nevertheless, the phase-reset RA2 model failed to accommodate the data from Experiments 3 and 4. In particular, it failed in Experiment 3 because it inappropriately averaged or ignored context IOIs on the basis of their interval properties. An additional modification to the phase-reset RA2 model, permitting concatenation of context intervals, succeeded in accommodating the Experiment 3 data; phase-reset RA3 realized earlier suggestions about perceptual and motor timing (Keele, Pokorny, Corcos, & Ivry, 1985; Vorberg, 1978; Vorberg & Hamburg, 1984). However, costs mount for the interval model with the addition of parameters, such as exclusion and concatenation. Finally, in spite of these additional parameters, all variants of the phase-reset model failed to account for the data in Experiment 4, in which we more thoroughly tested the fundamental phase-resetting assumption of the interval approach.

In this context, we note that our conclusion favoring an entrainment over an interval interpretation is at odds with the conclusion offered by Ivry and Hazeltine (1995), who reported findings similar to some of our findings in the current Experiments 1–3. Using an experimental design similar to ours (in their Experiment 4), Ivry and Hazeltine found minimal effects of variations in beginning times of comparison intervals following a gap of roughly twice the duration of a repeated standard (500 ms). Moreover, because the presence or absence of the precomparison gap also appeared to have little effect on performance, these authors concluded that the data failed to support an entrainment model. Our findings indeed agree with those of Ivry and Hazeltine, but they lead us to a different conclusion. Why? First, because gap durations in the Ivry and Hazeltine design were roughly double the context IOI, an entrained periodicity would have persisted through them, thereby explaining the null difference between gap (our Experiments 2 and 3) and no-gap conditions (our Experiment 1). Second, although Ivry and Hazeltine manipulated onsets of to-be-judged durations, they averaged performance over three different comparison beginning times (i.e., gaps). Thus, small effects of this manipulation may have been obscured, as we show in Experiment 4 of the present research. Third, our modeling results suggest that null effects of modest beginning-time manipulations are predicted if the oscillator is adaptive rather than rigid and when values of phase correction and period correction have converging effects on oscillator alignments. In other words, findings of both Ivry and Hazeltine and our Experiments 1–3 are consistent with either an interval model that is equipped with a running average weight and the ability to exclude and include intervals or a self-sustaining, two parameter oscillator having partial phase and period correction. The findings from Experiment 4, however, are only consistent with the latter interpretation.

The remainder of this article addresses general strengths and weaknesses of entrainment and interval models in the context of the present findings. We structure this discussion around the three central issues in time perception outlined in the introduction.

Timekeeper Response to Stimulus Onsets

The first issue concerns the response of an internal clock to events in the environment. All interval models subscribe to some form of complete phase correction (i.e., phase resetting). A defining feature of these models is that a pacemaker switch resets (restarts) the clock at arbitrary points in time in response to stimulus markers. Although recent studies have demonstrated that marker differences (e.g., auditory vs. visual) appear to modulate the switch (e.g., Grondin, Ivry, Franz, Perreault, & Metthé, 1996; Lustig & Meck, 2001), to our knowledge no research has suggested that the relative timing of identical markers (here, tone onsets) affects the efficiency of starting and stopping the clock. In light of this fact, the findings of Experiment 4 are especially difficult to explain for any model that assumes that the internal clock can be started or stopped at arbitrary points in time, much like a stopwatch, hourglass, or up counter (Hinton & Meck, 1997).

Although parameter estimates of phase correction varied somewhat across experiments, all quantitative fits of the best-fitting entrainment model suggested partial, rather than complete, phase correction. In Experiment 4, this model accurately predicted shifts in PSE values as a function of relatively large shifts in beginnings of comparisons. However, this entrainment model did not account for the observed asymmetry in PSEs, nor did it account for the reduced discriminability (higher JNDs) found with very early onsets of comparison intervals in Experiment 4. In other entrainment models, discriminability measures are linked to the width of an attentional pulse, which measures the standard deviation of a circular probability distribution (e.g., see Large & Jones, 1999). Large and Jones argued that a wider attentional pulse corresponds to greater uncertainty about the timing of future onsets, hence larger JNDs. However, in the simplified model presented here, we held pulse width constant in order to concentrate on its locus in time. The JND differences observed between early and late onsets are intriguing and have important implications for future modeling. They suggest that a wider attentional pulse obtains for early than for on-time and late onsets. An analysis of the unexpected conditions offers the following rationale. Unexpectedly early onsets are potentially more surprising than unexpectedly late onsets because they occur prior to an expected time point rather than after it. With unexpectedly late onsets, participants first experience the absence of a marker at the expected time point and then, after a delay, they experience the marker onset itself. One proposal is that in late-onset conditions, uncertainty is reduced about when a future marker occurs, producing a corresponding reduction in attentional pulse width (Jones, in press).

A related issue concerns differences in phase- (and period-) correction estimates for entrainment models in the four experiments, as shown in Figure 5 (e.g., for Experiment 4 vs. Experiments 1–3). Given the preceding discussion, it is possible that differences in parameter estimates are due to differences in task constraints. Thus, only in the task of Experiment 4 was the standard duration always equal to the main context IOI, and here the estimate of phase correction was high, suggesting the prominence of phase correction when no compensatory inputs from period correction are required. By contrast, in tasks in which period correction was clearly essential because standard durations varied (Experiments 1–3), phase parameter estimates were lower. Participants may consciously or unconsciously modulate the degree to which they phase and/or period correct depending on the temporal constraints of the task.

Internal Representation of Duration

The second issue concerns the way in which participants internally represent time. Is the duration of the standard conveyed explicitly by a nontemporal duration code or implicitly by the period of an oscillator? From an interval perspective, one suggestion is that the duration code is a count of the pulses that are accumulated over a to-be-timed interval (Gibbon, 1977; Hinton & Meck, 1997). To account for effects of context on the representation of such an interval, interval models assume a blending of multiple duration codes in memory. This blended memory might involve simple, after-the-fact averaging of independent duration codes for the context IOIs with the standard's code, which occurs at the time of a judgment; alternatively, blending may transpire in real time as a sequence unfolds. It is the latter possibility that we modeled in the three running average variants of the cardinal phase-reset model (RA1, RA2, RA3).

From an entrainment perspective, an oscillator's period functions implicitly as a dynamically changing working memory (i.e., over context IOIs). Quantitative fits to the Experiment 1–4 data favor this interpretation over ones involving discrete duration codes. Memory for the time span of an IOI is expressed, more or less accurately, by the period of an oscillator that does not have a fixed value but, rather, is influenced by the relative phasing of stimulus markers. This view offers a parsimonious account of effects of harmonically related context rates on time-judgment performance without positing a concatenation operation (Experiment 3; see also Barnes & Jones, 2000).

Successive Time Estimates

A third important difference between interval and entrainment approaches to short-interval timing concerns how these models treat the perception of successive time intervals. The interval approach assumes independent estimates of successive time intervals, whereas the entrainment approach assumes dependent estimates of successive time intervals.

Interval models enlist interval averaging to explain assimilation of independently stored time codes for successive intervals (i.e., context, standard IOIs, etc.). Reliance on assumptions of independence and simple averaging of these intervals is often justified in terms of predictive success and parsimony. Predictive successes for an interval-averaging model, using the B&J–M&K task, were clear in the present research. That is, when an aggregated internal code, based on averaging of internal representations of context with that of an unexpected standard, was used as a referent for judgments of a comparison interval, interval models nicely described assimilation of the standard. The running-average model (phase-reset RA2) with two parameters, one that produces a weighted average of the standard-plus-context intervals and another that permits the exclusion of highly deviant IOIs, was quite accurate in predicting patterns of time judgment errors in Experiments 1 and 2.

In spite of these predictive successes, a real-time running-average model, which includes parameters that effectively "turn off" the pacemaker at strategic points in the sequence, introduces a logical problem that ultimately threatens parsimony. In principle, a pacemaker's job is elegantly simple: It automatically responds independently to each successive physical marker and cannot contemplate the "what," "when," or "how" of the next marker. But models such as the phase-reset RA2 effectively endow the pacemaker with a dependency among time intervals that presumes foreknowledge of a future time interval (e.g., the gap). These models assume that a pacemaker can "know" in advance that a gap will follow a forthcoming tone onset and can further decide to exclude it on the basis of the value of another future time interval (the standard). Foreknowledge is also required in the phase-reset RA3 model, to implement constraints that determine whether two (or more) independently estimated intervals must be concatenated to yield an interval estimate that is twice the duration of a context IOI. Logically, it is unclear how all of this happens. One possible explanation appeals to gestalt patterning: A lengthened gap, in segregating groups of tones, triggers switching off of a pacemaker prior to the gap. However, gestalt groupings notoriously operate after the fact (e.g., after the gap); gestalt grouping principles, too, are not endowed with foreknowledge. Moreover, gestalt contingencies do not figure into current accounts of either running averages or of pacemaker activity; indeed, such a contingency for pacemakers raises issues beyond the scope of this article. The main point of this discussion is that any rationale for stopping a pacemaker strategically within a series of homogeneously marked IOIs is ad hoc. Such a rationale also fails to square with the logic of independence on which current interval model rests. Ultimately, this undermines claims about parsimony.

Entrainment explanations of assimilation build directly on presumed dependence among successive time intervals. Predictions of both quadratic ending profiles and flat beginning profiles derive from these assumptions. The ending profile reflects assimilation that is mainly the result of incomplete period correction, meaning that the impact of a prior IOI on an oscillator's period lingers to affect a participant's perception of a current time interval. Thus, a quadratic ending profile occurs when an unexpected standard is assimilated into a preceding context because of incomplete period correction. This profile sharpens as values of W_p approach zero, implying greater serial dependency. Predictions of beginning profiles (both flat and asymmetric) also reflect serial dependencies, but in this case they draw on interactions between phase and period correction at Times i and $i + 1$. In the best-fitting entrainment model, estimates of W_ϕ and W_p reflect this interaction; they are responsible for this model's ability to predict a quadratic expectancy profile in some cases (in which the timing manipulation affects the duration of a to-be-judged time interval) and a flat expectancy profile in other cases (in which the timing manipulation does not affect the duration of a to-be-judged time interval). Finally, although these real-time corrections of period and phase bear some similarity to a running average, here averaging is filtered through a current expected period and measured in relative (phase) rather than absolute terms.

Although the entrainment-model approach does not appeal to foreknowledge, one strength of this approach is that it nonetheless explains anticipations about forthcoming time intervals on the basis of assimilative responses to prior stimulation. Even when participants are explicitly told to ignore prior context intervals, they appear to be implicitly influenced by them. Does this mean that attending is involuntarily controlled by stimulus timing properties? Although the present research was not designed to fully answer this question, we think that some portion of participants' attention is probably automatically responding to context timing (Jones, in press). However, the degree to which this is so will depend on estimated values of phase- and period-correction parameters. We feel this is an important avenue of future research.

Overall, our results highlight the importance of quantitative (over qualitative) modeling of behavior and lead us to conclude that the most parsimonious account of rhythmic effects on perceived duration involves a self-sustaining oscillator that is phase- and period-coupled with stimulus markers. This view converges, in some respects, with recent research on synchronized tapping (e.g., Repp, 2002a; but see Repp, 2002b) and thus provides additional support for the conclusion that the perception and production of temporal intervals involve common timing mechanisms (Ivry & Hazeltine, 1995; Keele et al., 1985; Kristofferson, 1984).

Conclusions

Finally, we offer several general conclusions from this research pertinent to the debate between theorists investigating interval and entrainment views of time perception and production (Barnes & Jones, 2000; Drake & Botte, 1993; Ivry & Hazeltine, 1995; Jones, in press; Keele et al., 1989; Large, 1994; Large & Jones, 1999; McAuley, 1995; McAuley & Kidd, 1998; Pashler, 2001; Schulze, 1989; Vos et al., 1997). Overall, our findings suggest that people perceive and remember durations much as they pick up an underlying periodicity (i.e., a beat period) in music. On an intuitive level, the process we study here relates to a listener's sense and use of an induced but flexible periodicity. This beat, whether it is slow or fast, carries attention through certain lengthier intervals. Small variations in stimulus timing of individual tones do not throw a listener far off track, but these unexpected time intervals will be interpreted within the prevailing periodic framework and so may be distorted. Indeed, some large time changes can be disruptive.

Although this intuitive interpretation of entrainment seems to suggest that our findings relate only to musical events, in which rhythmic structure is explicit, it has broader applicability. With greater temporal variability in a sequence, whether music, speech, or other sound patterns, one or more internal oscillations can continue to adaptively track any temporally irregular sequence, in spite of the fact that the respective periodic expectancies become less precise (see, e.g., Large & Jones, 1999; McAuley, 1995). This fact casts doubt on claims that an interval model can accomplish what an entraining oscillator does, merely by operating in a "rhythmic mode," but that such an oscillator cannot operate in an interval mode. In fact, the interval-timer models examined here represent the most sophisticated adaptations of these models to a rhythmic mode that we have found; yet, they sometimes fail to "get the beat." By contrast, the entrainment models not only get the beat when timing is regular, they can "forget the beat" when timing is irregular by operating in an interval mode. That is, entrainment models may exhibit erratic expectancies in variably timed sequences (e.g., Large & Jones, 1999) or they may perfectly describe performance with the isolated time intervals used in classic two-interval time-judgment tasks that lack a rhythmic context, as shown here in Experiment 2 for control participants.

References

Allan, L. G. (1979). The perception of time. *Perception & Psychophysics, 26,* 340–354.

Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology, 41,* 254–311.

Bharucha, J. J., & Pryor, J. H. (1986). Disrupting the isochrony underlying rhythm: An asymmetry in discrimination. *Perception & Psychophysics, 40,* 137–141.

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review, 4,* 184–197.

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.

Drake, C., & Botte, M. C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics, 54,* 277–286.

Fraisse, P. (1963). *The psychology of time.* New York: Harper & Row.

Fraisse, P. (1978). Time and rhythm perception. In E. C. Carterette & M. P. Friedman (Eds.), *Perceptual coding: Handbook of perception* (Vol. 8, pp. 203–254). New York: Academic Press.

Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

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

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.

Grondin, S. (2001a). Discriminating time intervals presented in sequences marked by visual signals. *Perception & Psychophysics, 63,* 1214–1228.

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

Grondin, S. (2001c). A temporal account of the limited processing capacity. *Behavioral & Brain Sciences, 24,* 122–123.

Grondin, S., Ivry, R. B., Franz, E., Perreault, L., & Metthé, L. (1996). Marker's influence on the duration discrimination of intermodal intervals. *Perception & Psychophysics, 58,* 424–433.

Halpern, A. R., & Darwin, C. J. (1982). Duration discrimination in a series of rhythmic events. *Perception & Psychophysics, 31,* 86–89.

Hellstrom, A. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin, 97,* 35–61.

Hinton, S. C., & Meck, W. H. (1997). The "internal clocks" of circadian and interval timing. *Endeavour, 21,* 3–8.

Hirsh, I. J., Monahan, C. B., Grant, K. W., & Singh, P. G. (1990). Studies in auditory timing: I. Simple patterns. *Perception & Psychophysics, 47,* 215–226.

Hollingworth, H. L. (1910). The central tendency of judgment. *Journal of Philosophy, 7,* 461–468.

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 and Performance, 21,* 3–18.

James, W. (1890). *The principles of psychology* (Vols. 1–2). New York: Holt.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83,* 323–335.

Jones, M. R. (1981). A tutorial on some issues and methods in serial pattern research. *Perception & Psychophysics, 30,* 492–504.

Jones, M. R. (in press). Attention and timing. In J. Neuhoff (Ed.), *Ecological psychoacoustics.* New York: Academic Press.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96,* 459–491.

Jones, M. R., & Yee, W. (1993). Attending to auditory events: The role of temporal organization. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 199–230). Oxford, England: Oxford University Press.

Keele, S. W., Nicoletti, R., Ivry, R. I., & Pokorny, R. A. (1989). Mechanisms of perceptual timing: Beat-based or interval-based judgments? *Psychological Research, 50,* 251–256.

Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. I. (1985). Do perception and motor production share common timing mechanisms: A correlation analysis. *Acta Psychologica, 60,* 173–191.

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

Kristofferson, A. B. (1984). Quantal and deterministic timing in human duration discrimination. In J. Gibbon & L. Allan (Eds.), *Timing and time perception: Annals of the New York Academy of Sciences* (Vol. 423, pp. 3–15). New York: New York Academy of Sciences.

Large, E. W. (1994). *Dynamic representation of musical structure.* Unpublished doctoral dissertation, Ohio State University.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106,* 119–159.

Luce, R. D. (1959). *Individual choice behavior.* New York: Wiley.

Lustig, C., & Meck, W. H. (2001). Paying attention to time as one gets older. *Psychological Science, 12,* 478–484.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

Matell, M. S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *BioEssays, 22,* 94–103.

McAuley, J. D. (1995). *Perception of time phase: Toward an adaptive oscillator model of rhythmic pattern processing.* Unpublished doctoral dissertation, Indiana University Bloomington.

McAuley, J. D., & Kidd, G. R. (1998). Effect of deviations from temporal expectations on tempo discrimination of isochronous tone sequences. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1786–1800.

Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Computation, 1,* 359–371.

Michels, W. C., & Helson, H. (1954). A quantitative theory of time-order effects. *American Journal of Psychology, 67,* 327–334.

Michon, J. A. (1964). Studies on subjective duration: I. Differential sensitivity in the perception of repeated temporal intervals. *Acta Psychologica, 22,* 441–450.

Monahan, C. B., & Hirsh, I. J. (1990). Studies in auditory timing: 2. Rhythm patterns. *Perception & Psychophysics, 47,* 227–242.

Pashler, H. (2001). Perception and production of brief durations: Beat-based versus interval-based timing. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 485–493.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2,* 411–440.

Rakitin, B. C., Penney, T. B., Gibbon, J., 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.

Repp, B. H. (2002a). Automatic and voluntary control of phase correction following event onset shifts in sensorimotor synchronization. *Journal of Experimental Psychology; Human Perception and Performance, 28,* 410–430.

Repp, B. H. (2002b). Perception of timing is more context sensitive than sensorimotor synchronization. *Perception & Psychophysics, 64,* 703–716.

Restle, R. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review, 77,* 481–496.

Schulze, H. H. (1978). The detectability of local and global displacements in regular rhythmic patterns. *Psychological Research, 40,* 173–181.

Schulze, H. H. (1989). The perception of temporal deviations in isochronic patterns. *Perception & Psychophysics, 45,* 291–296.

Todd, R., Boltz, M., & Jones, M. R. (1989). The MIDILAB auditory research system. *Psychomusicology, 8,* 17–30.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychological Monographs: General & Applied, 77*(13, Whole No. 576), 1–31.

Treisman, M., Faulkner, A., Naish, P. L. N., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception, 19,* 705–743.

Turchioe, R. M. (1948). The relation of adjacent inhibitory stimuli to the central tendency effect. *Journal of General Psychology, 39,* 3–14.

Vorberg, D. (1978). Problems in the study of response timing: Comments on Reece's (1976) "A model of temporal tracking." *Acta Psychologica, 42,* 67–77.

Vorberg, D., & Hamburg, R. (1984). Timing of two-handed rhythmic performance. In J. Gibbon & L. Allan (Eds.), *Timing and time perception: Annals of the New York Academy of Sciences* (Vol. 423, pp. 390–406). New York: New York Academy of Sciences.

Vorberg, D., & Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action, Vol. 2: Motor skills* (pp. 181–261). New York: Academic Press.

Vos, P. G., van Assen, M., & Franek, M. (1997). Perceived tempo change is dependent on base tempo and direction of change. Evidence for a generalized version of Schulze's (1978) internal beat model. *Psychological Research/Psychologische Forschung, 59,* 240–247.

Wallin, J. E. W. (1911). Experimental studies of rhythm and time. *Psychological Review, 18,* 202–222.

Wearden, J. H., Philpot, K., & Win, T. (1999). Speeding up and (. . . relatively . . .) slowing down an internal clock in humans. *Behavioural Processes, 46,* 63–73.

Winfree, A. T. (2001). *The geometry of biological time* (2nd ed.). New York: Springer-Verlag.

Woodrow, H. (1951). Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1224–1236). New York: Wiley.

Yee, W., Holleran, S., & Jones, M. R. (1994). Sensitivity to event timing in regular and irregular sequences: Influences of musical skill. *Perception & Psychophysics, 56,* 461–471.

Zakay, D. (1993). Time estimation methods: Do they influence prospective duration estimates? *Perception, 22,* 91–101.

(*Appendixes follow*)

Appendix A

Model Simulations

Formula for Temporal Contrast (C)

$$C = \begin{cases} (\phi_i + \mathrm{IOI}_i/P_i) \pmod 1 - 1.0 & \text{if } (\phi_{i-1} + \mathrm{IOI}_i/P_i) \pmod 1 > 0.5 \\ (\phi_i + \mathrm{IOI}_i/P_i) \pmod 1 & \text{otherwise} \end{cases},$$

where ϕ_i is the relative phase of the *i*th stimulus onset, IOI_i is the duration of the *i*th stimulus IOI, and P_i is the current oscillator period. In this equation, values of C range between -0.5 and 0.5.

Luce-Choice Rule

For all models (interval and entrainment), values of final temporal contrast, C_{final}, were converted to shorter, same, and longer response probabilities using a three-category (shorter, same, longer) Luce-choice rule (Luce, 1959): $\exp[\gamma\ f(C_i)]\ /\ \Sigma\ \exp[\gamma\ f(C_i)]$. In this equation, the numerator is determined for each category, i (in our case, shorter, same, and longer), with the denominator being obtained by summing across all three categories; $f(C_i)$ is determined separately for each response category, i, using a linear transformation of C_{final}. There was one scaling parameter, γ, which we permitted to vary between experiments. Estimates of γ are instructive in that they captured overall differences in performance levels across experiments (higher values of γ correspond to better overall performance). It is noteworthy that, by experiment, estimates of γ did not differ between the best-fitting interval and entrainment models, indicating that performance was scaled identically for both classes of model and that the observed RMSEA differences could be attributed entirely to differences in phase and period correction. A between-experiments comparison of γ highlights differences in overall task difficulty. For Experiments 1 and 2 (which differed only in the gap), estimates of γ were equivalent ($\gamma = 4.5$); for Experiment 3 (which involved halving the context IOI and slightly lower performance levels), the best estimate of γ was lower than that for the previous two experiments ($\gamma = 3.0$). In Experiment 4, which involved a two-choice task with a fixed standard, we converted values of C_{final} directly to *shorter* and *longer* responses by assuming that response probabilities approximated a normal distribution, with $C_{\mathrm{final}} < 0$ associated with *shorter* responses and $C_{\mathrm{final}} > 0$ associated with *longer* responses.

Appendix B

Mean Proportions of Shorter, Same, and Longer Responses for the 27 Conditions in Experiment 1 (3 Standard Endings × 3 Comparison Beginnings × 3 Types of Comparison)

Interonset interval (ms)		Mean response probability		
Standard	Comparison	Shorter	Same	Longer
		Early beginning of comparison		
540	486	.88	.10	.02
540	540	.35	.55	.10
540	594	.07	.54	.39
600	540	.84	.12	.04
600	600	.15	.73	.12
600	660	.06	.31	.63
660	594	.51	.43	.06
660	600	.16	.54	.30
660	726	.04	.17	.79
		On-time beginning of comparison		
540	486	.83	.15	.02
540	540	.34	.56	.10
540	594	.10	.50	.40
600	540	.82	.14	.04
600	600	.18	.72	.10
600	660	.07	.26	.67
660	594	.50	.44	.06
660	600	.12	.60	.28
660	726	.05	.16	.79
		Late beginning of comparison		
540	486	.83	.15	.02
540	540	.32	.56	.12
540	594	.11	.50	.39
600	540	.81	.17	.02
600	600	.17	.70	.13
600	660	.05	.25	.70
660	594	.52	.42	.06
660	600	.12	.58	.30
660	726	.06	.15	.79

Received September 27, 2001
Revision received April 2, 2003
Accepted May 8, 2003 ■