

Time judgments in global temporal contexts

MARI RIESS JONES

Ohio State University, Columbus, Ohio

and

J. DEVIN MCAULEY

Bowling Green State University, Bowling Green, Ohio

With three experiments, we examined the effects of global temporal context on time judgments as gauged by constant errors (CEs) and estimates of a preferred period (P). In Experiment 1, participants in seven different conditions listened to sequences of a given rate (with interonset intervals ranging from 200 to 800 msec) and judged the relative duration of a final (comparison) time interval. No P emerged. In Experiments 2 and 3, we embedded the same rates in different global (session) contexts that varied according to (1) mean session rate, (2) standard deviation, (3) range, and (4) number of different rates in a session. Evidence from CEs indicated that P varied primarily as a function of mean session rate and range of tempi. The best predictor of errors involved a measure termed *relative range* (RR = range/mean session rate). A general algorithm incorporating RR successfully predicts P, and the implications of this algorithm are discussed.

Recent research on short-interval timing suggests that temporal context systematically affects the perceived duration of an embedded time interval (Barnes & Jones, 2000; Drake & Botte, 1993; Jones & Yee, 1997; Large & Jones, 1999; McAuley & Jones, 2003; McAuley & Kidd, 1998; Monahan & Hirsh, 1990). Figure 1 illustrates one type of time-judgment task in which this occurs: A *local context sequence* consists of a series of brief tones, which mark out an isochronous sequence of interonset intervals (IOIs), denoted as the base IOI. This sequence precedes a final pair of to-be-judged time intervals, referred to as the *standard* and the *comparison*. In this paradigm, the rate of the local context sequence can alter the perceived duration of the standard IOI whenever the base IOI (i.e., sequence rate) differs from the standard IOI (Barnes & Jones, 2000; McAuley & Jones, 2003). McAuley and Jones suggested that constant errors in the perceived duration of the standard IOI occur because the local context sequence induces an internal periodicity that distorts participants' perception of the standard IOI in the direction of the base IOI. In their framework, the perceived duration of a standard IOI, obtained by measuring a point of

subjective equality (PSE), estimates the current period of an internal oscillator tracking the local context sequence. However, Barnes and Jones found evidence that rates of other sequences within a session also affected the perceived duration of a standard IOI, hence the period of the tracking oscillator. This suggests a role for *global temporal context*.

The focus of this article is on the effects of global temporal context on time judgments. The study reported here had two aims. First, we aimed to assess how distributional properties of global temporal contexts affect patterns of constant errors in time judgments. Second, we sought to determine which effects of global distributional properties are mediated by the dynamic trial-to-trial changes in local sequence rate. To address these issues, we performed three time-judgment experiments, adapting the paradigm outlined in Figure 1. A local context sequence, consisting of three tones marking out a single base IOI, was followed by a standard IOI and then a comparison IOI where the standard IOI always equaled the base IOI. Across experiments, and in different multi-rate sessions, global temporal context was manipulated through distributional properties involving base IOIs. Constant errors (CEs) were measured by subtracting each base IOI from an obtained PSE ($CE = PSE - \text{base IOI}$); CEs were then compared across conditions (sessions) to assess the impact of global context on time judgments.

Classical Issues in Timing: Indifference Intervals

In the literature, patterns of constant errors in time perception are well established. Historically, a general finding has been that short standard intervals tend to produce positive CEs, whereas long ones tend to produce negative CEs, with an *indifference interval* specified as

A portion of this research was conducted while J.D.M. was a music cognition postdoctoral research fellow at Ohio State University. It was sponsored, in part, by grants awarded to M.R.J. from the International Foundation for Music Research and from the National Science Foundation (BCS-9809446). We are indebted to Ralph Barnes, Steven Finney, Noah Mackenzie, Rosalee Meyer, Nate Miller, J. P. Miller, Peter Pfordresher, Amandine Penel, Jennifer Hoffman Puente, and members of the Timing Research Group at Bowling Green State University for helpful comments on an earlier version of this manuscript. We are also grateful to Bruno Repp, who commented on an earlier version of the manuscript. Correspondence should be addressed to M. R. Jones, Department of Psychology, 142 Townshend Hall, Ohio State University, Columbus, OH 43212 (e-mail: jones.80@osu.edu).

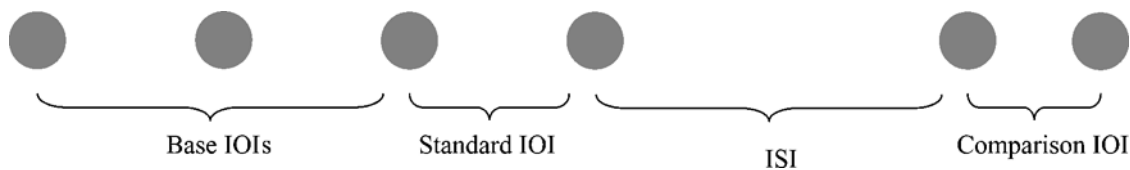


Figure 1. Standard and comparison interonset intervals (IOIs) preceded by a local context sequence that marks out a constant base IOI. The standard IOI was always equal to the base IOI. The interstimulus interval (ISI) was the time interval separating the onset of the tone ending the standard IOI and the onset of the tone beginning the comparison IOI. Listeners judged whether the comparison IOI was shorter, equal to, or longer than the standard IOI.

an intermediate time interval that produces a null constant error (Vierordt, 1868; Woodrow, 1934, 1951; Woodworth & Schlosberg, 1954). One issue concerns enormous variability in the reported location of the indifference interval; estimates range from a low of around 400 msec to a high of 5.0 sec (Fraisse, 1978; Wallin, 1911a, 1911b; Woodrow, 1951). Although some variability in indifference interval estimates is attributable to the range of time intervals involved (e.g., Jamieson & Petrusic, 1975; Woodrow, 1951), it is not known what distributional properties of global temporal contexts predict patterns of CEs and whether the source of the CEs is perceptual, memorial, or resides in decision processes (Allan, 1979). Moreover, firm conclusions about the source of CEs and the indifference interval are likely to be tied to a particular theoretical framework (Allan, 1977, 1979; Goldstone, Lhamon, & Boardman, 1957; Hellstrom, 1985; Helson, 1964; Hollingsworth, 1910; Treisman, 1963; Turchioe, 1948).

Much of the classical research on the indifference interval involves a theoretical framework that assumes an interval model of time perception with three separate stages: clock, memory, and decision stages (e.g., Gibbon, 1977; for a review, see Grondin, 2001). Applied to indifference interval phenomena, this approach assumes that CEs in timing reflect inadequate codes or memory traces for the standard IOI, the comparison IOI, or both. One assumption is that the indifference interval represents a duration code, which functions as a long-term internal referent for time judgments. From this perspective, both assimilation and contrast are considered aspects of indifference interval phenomena; *assimilation* refers to a perceptual distortion of an interval (e.g., a standard IOI) toward this global referent interval (e.g., based on some measure of central tendency), whereas *contrast* refers to an opposing drift—namely, a distortion away from the global referent interval (e.g., Hellstrom, 1985).

In contrast to an interval model interpretation, Fraisse used the construct of a *preferred internal periodicity* to explain indifference interval phenomena (Fraisse, 1963; cf. Hirsh & Watson, 1996). In this framework, the construct of a biologically preferred internal period (or tempo) referred to a dominant rhythmic action (e.g., walking) that displayed a period between 600 and 750 msec (Fraisse, 1963, 1978, 1984; see also Wallin, 1911a, 1911b). Early estimates of the indifference interval seemed to provide

initial support for such links in that they often fell within this range. Nevertheless, the possibility that session context exerts a strong influence on this phenomenon presented a significant problem for any hypothesis of a biologically preferred periodicity that is unresponsive to different stimulus time intervals. Although Fraisse acknowledged this, he offered no solution to the dilemma of how an indifference interval, on the basis of biological periodicity, might shift—that is, adapt—in response to changes in the rate of some external stimulus and the context in which it occurs (Fraisse, 1963, p. 22).

Current Issues: Internal Periods and Contextual Timing

Recent interest in entrainment timing models affords a reexamination of problems associated with biological periodicities. Entrainment models are not stage models. Along with beat-based timing models, they have developed within a different tradition that is congenial to hypotheses about internal periodicities and that offer the potential to explain the extrapolation of periodically timed responses to sequences (Boltz, 1994; Drake, Jones, & Baruch, 2000; Jones, 1976, 2004; Large & Jones, 1999; McAuley, 1995; McAuley & Jones, 2003; McAuley & Kidd, 1998; Povel & Essens, 1985; Schulze, 1978; cf. Keele, Nicoletti, Ivry, & Pokorny, 1989). Entrainment models differ from beat-based timing models (e.g., Povel & Essens, 1985; Schulze, 1978, 1989) in that the internal periodicities posited by entrainment models *adapt* to changes in local rate, whereas those proposed by beat-based models *do not adapt*. Entrainment models often involve one or more internal periodicities that adjust their phase and period in response to the timing of stimulus onsets in sequences. Internal periodicities are formalized as self-sustaining (limit cycle) oscillators that, within limits, selectively tune to and track changes in sequence rate. As such, they can supply a momentary internal standard for judgments about relative duration. McAuley and Jones (2003) showed that although both entrainment and beat-based models respond to sequences of different (fixed) rates, entrainment models are more successful in explaining time-judgment behavior, because they permit internal oscillators to adapt to various changes in rate, including those that ensue from trial to trial in a session.

Experimental Paradigms and Their Implications

Although both classical and current approaches to timing address context effects due to session properties, they have relied on different experimental paradigms. A common paradigm used to evaluate classical hypotheses about the indifference interval is the *two-interval paradigm*, which requires a response (e.g., shorter, longer) to a comparison time interval relative to a single standard IOI on each trial in a session. The two-interval paradigm does not precede the standard–comparison pair with a local context sequence. Instead, over trials the magnitude and presentation order of the standard and comparison IOIs are varied. In addition to CEs, time distortions due to session context have come to be assessed using a time-order error (TOE) measure. A TOE is evident when standard and comparison IOIs are physically identical but are consistently judged to differ in a given direction (Hellstrom, 1985). TOEs, which are calculated in various ways, can also index time distortions where the two intervals involved are physically different by using an asymmetry in the proportion of shorter and longer responses; in this case, the resulting values of TOEs depend on both the presentation order of these intervals (e.g., standard, comparison)¹ and their durations (Allan, 1979; Hellstrom, 1985; Michels & Helson, 1954). As with CEs, short durations tend to produce positive TOEs, whereas long durations tend to produce negative TOEs.² In brief, using the two-interval paradigm, an indifference interval can be estimated over trials as the hypothetical standard time interval that produces a TOE of zero, given some session (global) context.

A different paradigm, the *sequence paradigm*, precedes a standard–comparison pair with a local context sequence on each trial in a session as suggested in Figure 1 (Barnes & Jones, 2000; McAuley & Jones, 2003; McAuley & Kidd, 1998; Schulze, 1978; see Jones, 2004, for a review). In this paradigm, which is used in the present research, participants may be offered three judgment responses (shorter, same, longer) to prevent strategies based on listening only for departures from a local (sequence) rate. In this research, the local context sequence is a necessary element, because the goal is to understand how a larger session context affects participants' responses to the rate of individual sequences. A PSE is useful in estimating the response to a local rate because it assesses the difference between an individual's current sense of a sequence rate and the objective rate (base IOI) in the form of a CE. Thus, we assume that a PSE reflects not only the influence of a local context on listeners' perception of sequence rate, but also the influence of the prior global (session) context. In this respect, a PSE is assumed to provide a valid estimate of a participant's perceived duration of a local rate *at the moment of judgment*.³ In short, we rely on PSE, as well as CE, scores to assess the role of various contextual factors in a listener's perception of local context rate.

A distinctive aspect of this paradigm relates to the absence of order reversals of the to-be-judged intervals—

for example, standard and comparison IOIs (i.e., as in the two-interval paradigm). Although order reversals are essential to calculating TOE, they do not make sense in a study of context rate because the situation of interest has a built-in asymmetry between the standard IOI, which is a sequence of specified rate, and the comparison IOI, which is often a single interval (see Barnes & Jones, 2000, and Jones, 2004, for details). Thus, such designs do not lend themselves either to reversals or to TOE calculations. Alternatively, CE and PSE scores offer a means of evaluating context-sensitive responses to important aspects of everyday auditory environments where time intervals of speech and music commonly arrive embedded within distinctive rates and rhythmic contexts. Consequently, in this paradigm the assessment of global context effects requires a different strategy, one that permits a focus on sequence rate.

Such a strategy implies that we may rephrase questions about global context to ask: "Does a multirate session instill in participants some overriding *pace* that systematically distorts their judgments about individual sequence rates?" Hypotheses about sequence rate describe the internal referent in terms of a hypothetical pace or an *internal periodicity*; here, we assume that if a general sense of pace applies to an entire session, it must somehow emerge from the collective local rates of constituent sequences within that session. This is in contrast to an interval model where the internal referent (i.e., corresponding to an indifference interval) is conceived as a discrete interval elicited by constituent session time intervals. Thus, an important implication is a theoretical one concerning the *internal representation of time*: From an interval model perspective, the indifference interval is a memory trace of a single discrete interval, whereas from an entrainment perspective a global context induces one or more internal periodicities that contribute to an overall sense of pace that may be expressed as a reverberating circuit or emergent internal period (e.g., McAuley & Jones, 2003).

Entrainment models suggest that over a series of trials in multirate sessions, a listener's judgments will be increasingly influenced by one or several internal periodicities assumed by entraining oscillators. However, current entrainment models have not addressed global effects concerned with a general sense of pace in sessions that contain a variety of sequence rates (i.e., multirate sessions). In this article, we investigate aspects of multirate sessions that may predict a listener's development of a global pace, if one exists. This research is based on several assumptions. First, we assume that for any local sequence rate, the best estimate of the period of an entraining oscillator, as it responds to recurrences of that sequence over trials, is given by the PSE for that base IOI (collapsed over trials). Second, we assume that in multirate sessions, the best gross estimate of some global pace is given by the internal periodicity that is assumed most often by different oscillators activated over the course of a session. We propose that this global pace can be esti-

mated from observed data by examining the pattern of CEs over all base IOIs within a session; this estimate yields a contextually determined *preferred period* (P). Note that, although we assume that P is a biologically based internal periodicity, as Fraisse suggested, it is not fixed. Finally, we assume that because oscillators are adaptive, P can be “pushed around” by local sequence rates within a session context, and hence trial-to-trial rate changes will be important. That is, hypotheses based on adaptive oscillators have some potential for addressing unresolved issues concerned with effects of session context on P. We return to the relationship between P and the indifference interval, as classically described, in the General Discussion section.

Plan of Experiments

In three time-judgment experiments, using the paradigm of Figure 1, we manipulate statistical properties of session context to assess their impact on CEs and P. On each trial, a sequence of constant local rate (base IOI = standard IOI) is followed by a comparison IOI.⁴ To ensure that participants attend to both similarities and differences between comparison and base IOIs, we used the prospective time-judgment task employed by Barnes and Jones (2000) in which participants judge a comparison IOI using three response categories: *shorter*, *same*, or *longer*. When a comparison differed from the base IOI, the difference exceeded threshold values (.04–.07) established in time-discrimination tasks using sequences of similar lengths and rates (e.g., Drake & Botte, 1993; Halpern & Darwin, 1982; Jones & Yee 1997; Large & Jones, 1999).

Experiment 1 was a baseline study; participants in different conditions encountered sequences of the same rate (base IOI). Experiment 2 combined various base IOIs (local contexts) in a single session to create eight multirate conditions that differed in distributional properties (e.g., central tendency, heterogeneity of sessions). Experiment 3 further investigated multirate contexts by orthogonally manipulating two session heterogeneity properties, *n* (the number of different base IOIs presented within the session) and *R* (the range of the base IOIs). In all experiments, the overall magnitude and direction of errors was assessed using the average CE at each base IOI. An estimate of P was then derived under the assumption that the preferred oscillator period is one that minimizes failures to adapt; thus, P is estimated as the periodicity given by a CE of zero. We also assess overall accuracy (proportion correct, or PC), as well as conditional accuracy scores—that is, conditional on trial-to-trial changes in local sequence rate.

EXPERIMENT 1 Single-Rate Baseline Study

The primary goal of Experiment 1 was to establish stable baseline performance levels for base IOIs of 200, 300, 400, 500, 600, 700, and 800 msec in single-rate ses-

sions using the task shown in Figure 1. This design permits a direct test of an absolute P hypothesis. If a fixed biological predisposition for a particular sequence rate exists, a determinate value of P should emerge in spite of the fact that each of the $n = 7$ rates is presented to a different group of listeners. For example, if P corresponds to an internal P of 600 msec, a value given most often by proponents of this view (e.g., Fraisse, 1963), CE scores obtained from different groups of subjects, each exposed to a single repeated base IOI, should nonetheless reveal systematic over- and underestimates of the base IOI relative to 600 msec.

Method

Participants. A total of 113 students from an introductory psychology class at Ohio State University participated in the experiment for course credit. All reported normal hearing. They were randomly assigned to one of seven base IOI conditions (200 msec, $n = 18$; 300 msec, $n = 18$; 400 msec, $n = 16$; 500 msec, $n = 16$; 600 msec, $n = 16$; 700 msec, $n = 15$; 800 msec, $n = 14$). These numbers represent 96.6% of recruited individuals; data for the remainder were excluded for either failure to comply with instructions ($n = 2$) or for below-chance ($p = .333$) performance levels overall ($n = 2$).

Equipment and Stimulus Construction. All aspects of stimulus generation and response collection were controlled by the Midilab software package (Todd, Boltz, & Jones, 1989). Individual tone markers, each lasting 60 msec, were generated on an IBM-compatible computer interfaced with a Roland MPU-401 Midi processing unit that controlled a Yamaha TX81Z FM tone generator set to a sine wave voice at 440 Hz. Tone sequences were transmitted to a separate experimental room and amplified using a Rane HC-6 headphone console. Each participant listened to stimuli over AKG-K270 headphones at a comfortable listening level (adjusted to levels between 60 and 70 dB SPL according to individual preference).

Design. Experiment 1 implemented a single-factor between-subjects design. The participants were randomly assigned to one of seven base IOI conditions (200, 300, 400, 500, 600, 700, or 800 msec) and asked to judge the duration of a comparison IOI relative to the base IOI over the course of a single experimental session. For each base IOI, there were five possible comparison IOIs (24% shorter, 12% shorter, equal to the base IOI, 12% longer, and 24% longer).

Procedure. Each trial was initiated by a 500-msec high-pitched (5274-Hz) warning sound, followed by a 500-msec silent interval prior to the onset of the first context tone. On each trial, an isochronous four-tone context sequence (three base IOIs) was followed by a comparison IOI. The time interval between the onset of the tone terminating the base IOI sequence and that of the first tone of the comparison IOI was always equal to twice the base IOI. The same base IOI occurred on all trials. During a 2.5-sec response interval, the participants judged the duration of a comparison IOI relative to the base IOI by pressing one of three buttons (on a response box) labeled *shorter*, *same*, or *longer*. They were told to respond quickly and accurately and to guess if unsure. In each of the seven base IOI conditions, a session consisted of a series of 12 practice trials with corrective feedback, followed by five blocks of 36 test trials, without feedback. There were an equal number of shorter, equal, and longer comparisons included in each block. Over the 180 trials, 60 comparison intervals were equal to the base IOI; each of the four comparisons that differed from the base IOI occurred 30 times. Test trials were presented in one of three different random orders. Short rest breaks occurred between blocks; a longer break (ca. 5 min) occurred midway through the approximately 90-min session. Finally, listeners completed a musical background questionnaire (e.g., years of musical training, etc.).

Data analysis. Data were analyzed in terms of the probability of a time-judgment error, using proportion of correct responses (PC),

and in terms of the sign and magnitude of these errors, using CE scores. The former was measured by combining the five comparison IOIs (two shorter [12%, 24%], two longer [12%, 24%], and one same) and then determining the overall proportion of correct responses. The latter was estimated for each base IOI using the formulation: $CE = PSE - \text{base IOI}$.

With respect to the measurement of PSE, various proposals differ, depending on task (Gescheider, 1997; Hellstrom, 1979; Treisman, 1963). Conventionally, the PSE in a two-response task is defined as the stimulus continuum value (base IOI) in a psychometric function where $P(\text{short}) = P(\text{long}) = .50$. For the three-choice (shorter, same, longer) task, Greenberg (1965, pp. 111–112) described three methods, yielding similar results. The first effectively reduces the three-choice task to a two-choice task by assigning the same responses to the shorter and longer categories in equal proportions. The second method is to allocate the same responses to the shorter and longer categories in proportion to the observed $P(\text{short})$ and $P(\text{long})$. The third method is a modification of Thurstone's law of comparative judgment (Thurstone, 1927), which produces a result somewhere in between the first two methods.

The method we used to estimate PSE for the three-choice task is similar, but not identical, to Greenberg's (1965) third method (a modified Thurstone model). Although our method is somewhat unconventional, the results using it and the corresponding Greenberg method were highly correlated ($r = .97$) and in some cases our method provided a slightly better fit to the data (see the Appendix, Table A1, for comparative fits). We refer to this as a *difference method*.⁵ First, we determined [$P(\text{short}) - P(\text{long})$] for each of the five comparison IOIs (including the *same* comparison) and defined PSE as the base IOI value that produced $P(\text{short}) - P(\text{long}) = 0$. To estimate PSE, we linearly rescaled the difference scores for each comparison so that they were bounded between 0 and 1, rather than between -1 and 1. The resulting psychometric curves approximated a cumulative normal distribution (see the Appendix for exemplar psychometric functions). Transforming the resulting values to z coordinates yielded linear psychometric functions, validating normality assumptions (Gescheider, 1997). A regression line was then fit through the transformed points, and PSE was given by the intersection of the line with the abscissa ($z = 0$); see Macmillan and Creelman (1991, pp. 219–220) for a description of the z transform method for estimating PSE. On the basis of average PSE in each condition, we estimated P by regressing the constant error ($CE =$

$PSE - \text{base IOI}$) over the continuum of base IOI values and then determining the time interval that produced $CE = 0$.

Results and Discussion

Figure 2 shows mean CE with 95% confidence intervals for each of the seven base IOI conditions. Individual subject estimates of CE were mainly small and positive, indicating high accuracy with a tendency to produce more *shorter* than *longer* responses. On average, CE values ranged from a low of 0.6 msec (700-msec base IOI) to a high of 8.3 msec (300-msec base IOI). Overall, the CE data did not reveal a pattern of over- and underestimation errors suggestive of a P (i.e., positive CEs at short base IOIs and negative CEs at long base IOIs). This was confirmed by a one-way between-subjects analysis of variance (ANOVA) on CE, which showed no significant effect of base IOI [$F(6, 106) = 0.91, MS_e = 151.27, p = .49$] and a regression analysis of CE and base IOI, which produced a nonsignificant r squared value of .28. The corresponding one-way between-subjects ANOVA on PC also showed no significant effect of base IOI [$F(6, 106) = 1.92, MS_e = 0.006, p = .083$].⁶ Collapsed over base IOI, mean PC was 0.82 ± 0.01 . In summary, the results of Experiment 1 yield little support for the concept of an absolute P .

EXPERIMENT 2 Multiple-Rate Context Study

In Experiment 1, participants experienced the same base IOI on all trials in a session, and we found little evidence for an absolute P , independent of global context. However, local and global (session) rates in Experiment 1 were identical in the seven base IOI conditions. Consequently, although Experiment 1 provides useful baseline data, it tells us little about global contextual determi-

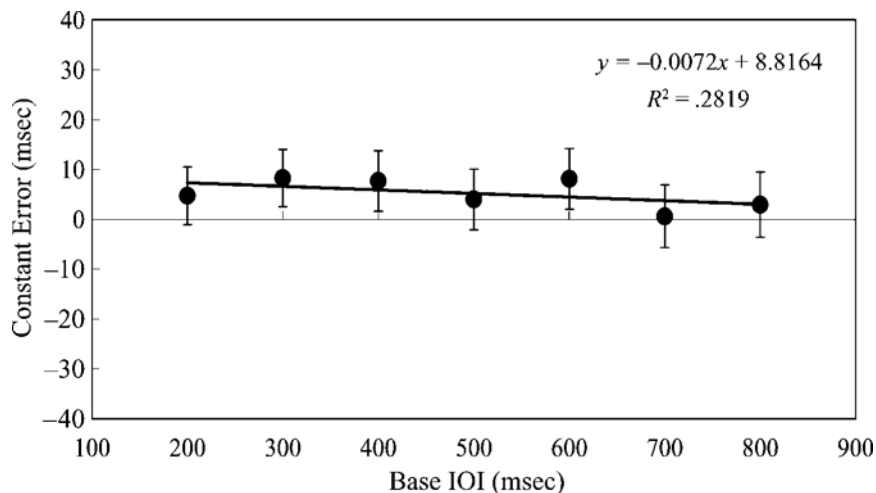


Figure 2. Constant error scores (averaged over participants) as a function of base IOI in Experiment 1. The solid line shows the regression line used to estimate the location of P (see text). Bars represent 95% confidence intervals.

nants of a P. In Experiments 2 and 3, we pursue this issue by decoupling local and global context by using multi-rate session contexts.

In Experiment 2, the same seven base IOIs of Experiment 1 were combined to form the eight multirate global (session) contexts shown in Table 1. The conditions differed with respect to the base IOIs included in each session (column 2). Two questions can be addressed with these manipulations. First, does a P emerge when the base IOI varies from trial to trial within a session, rather than remaining constant (as in Experiment 1)? Second, if a P does emerge, what global statistical properties of a session predict it? To these ends, Table 1 summarizes a number of important session properties of the set of *n* base IOIs for each condition (columns 3–7). These include measures of central tendency (arithmetic and geometric means) and heterogeneity (i.e., *n*, *SD*, *R*), where *R*, range, refers to the difference between the slowest (largest base IOI) and fastest (smallest base IOI) sequence in a session.

In terms of central tendency, if a P is due to an averaging process, this may involve either the arithmetic or geometric session mean of the *n* equiprobable base IOIs. As Table 1 indicates, the arithmetic mean in this design assumed only two values (IOIs of 400 or 600 msec); this variable, termed *set*, identified whether the mean base IOI conveyed, overall, a fast (*F* = 400 msec) or slow (*S* = 600 msec) global rate. By contrast, the geometric means (column 4 of Table 1; hereafter distinguished as *GM*) differed for each of the eight conditions. Thus, if P is determined by the arithmetic mean, *set* (*F*, *S*) should predict its value. As a check on statistical averaging, half of these conditions contained a reinforcing base IOI—that is, an IOI that was identical to the arithmetic mean rate—and half did not. If the arithmetic mean correctly describes the averaging process, it will be the better predictor of P, regardless of whether or not the mean base IOI actually occurs as a reinforcing local rate within the session. Finally, if P is determined by the *GM*, the eight individual geometric means will better predict P estimates than will the arithmetic mean. Hereafter, the term *mean* is synonymous with arithmetic mean.

We know little about the role of session heterogeneity. Speculation in the classical time-judgment literature associated with the indifference interval suggests that relevant heterogeneity measures may involve an uncertainty measure based on the number of different rates, *n*, the standard deviation of rates, *SD*, and/or the range of rates, *R* (Hellstrom, 1985; Jamieson & Petrusic, 1975; Woodrow, 1951). These statistics describe different distributional properties of base IOIs of the four conditions within each set (see Table 1). Within each set, a condition can be uniquely identified by *n* ($2 \leq n \leq 5$) when taken together with set type (*F* or *S*): *F2*, *F3*, *F4*, *F5*, and *S2*, *S3*, *S4*, *S5*. On the one hand, manipulations of *n*, *SD*, and *R* may have little effect on estimates of P if central tendency dominates time judgments. On the other hand, one of these statistics may affect either estimates of P or the reliability of these estimates. For instance, as session heterogeneity of local rates increases (due to larger *n*, *SD*, or *R*), overall accuracy may decline and/or estimates of a P (due to CE scores) may be more variable. To gain a more complete picture, we supplement estimates of P in three ways by assessing how different heterogeneity measures affect (1) overall error rates, inversely related to PC, (2) conditional accuracy as a function of trial-to-trial changes in base IOI, and (3) accuracy levels within one *diagnostic* rate condition (the base IOI of 500 msec) that occurred in all of the eight context conditions of Experiment 2.

Method

Participants. A total of 83 students from an introductory psychology class at Ohio State University (all with normal hearing) participated for course credit. This subject number reflects 94.3% of the individuals recruited; data for the remainder were excluded for either failure to comply with instructions (*n* = 1) or for below-chance (*p* = .33) performance levels overall (*n* = 4).

Equipment and stimulus construction. The equipment and stimulus construction were identical to those in Experiment 1.

Design. We employed a 4 × 2 mixed-factorial design. Four multiple rate conditions (*n* = 2, 3, 4, 5) were crossed with two mean rates (*F*, *S*) to generate eight between-subjects context conditions, as shown in Table 1. Half the contexts resulted in a mean session rate of 400 msec (*F*), and half realized a mean session rate of 600 msec (*S*). Half the conditions contained a base IOI at a rate equivalent to

Table 1
Global Distributional Properties of the Eight Session Context Conditions Constructed for Experiment 2

Context	Base IOIs (msec)	Central Tendency		Heterogeneity		
		Arithmetic Mean	Geometric Mean	<i>n</i>	<i>SD</i>	<i>R</i>
F5	200, 300, 400, 500, 600	400	372.8	5	158	400
F4	200, 300, 500, 600	400	366.3	4	183	400
F3	300, 400, 500	400	391.5	3	100	200
F2	300, 500	400	387.3	2	144	200
S5	400, 500, 600, 700, 800	600	582.7	5	158	400
S4	400, 500, 700, 800	600	578.5	4	183	400
S3	500, 600, 700	600	594.4	3	100	200
S2	500, 700	600	591.6	2	144	200

Note—Bold font indicates the conditions that included the arithmetic mean as one of the base IOIs.

mean session rate (i.e., F3, F5; S3, S5), and half did not (F2, F4; S2, S4; see column 4, Table 1). The number of participants in each context condition was F2, $n = 10$; F3, $n = 11$; F4, $n = 11$; F5, $n = 9$; S2, $n = 10$; S3, $n = 11$; S4, $n = 10$; S5, $n = 11$. As in Experiment 1, there were five possible comparison IOIs for each base IOI (24% shorter, 12% shorter, equal to the base IOI, 12% longer, and 24% longer).

Procedure. The procedure was identical to that of Experiment 1 with the exception that base IOI varied randomly from trial to trial. We equated the number of shorter, same, and longer comparisons across the n base IOIs in each condition, across conditions (using 12 to 18 practice trials and from 120 to 300 test trials, depending on n), and within blocks (see below). As in Experiment 1, test trials were presented in five blocks with an equal number of shorter, equal, and longer comparisons included in each block. In test trials, in all conditions, each base IOI always occurred 60 times. Thus, the number of trials in each of the five blocks necessarily differed with n ; trials per block were 24, 36, 48, and 60, for $n = 2, 3, 4$, and 5 conditions, respectively. Finally, the total number of test trials was 120, 180, 240, and 300, for $n = 2, 3, 4$, and 5 conditions, respectively. The total length of a session varied from 45 to 120 min.

Results

The results are presented in four sections. First, we examine the sign/magnitude of CE scores for each base IOI to obtain estimates of P for each condition; here, the contribution of central tendency to estimates of P is assessed. Second, we examine overall accuracy across conditions, using PC; here, the contribution of session heterogeneity is assessed. Third, we examine conditional (trial-to-trial) accuracy and consider its impact on CEs and PC. Finally, to confirm global context effects on responses to local rate, we analyze a single, diagnostic, rate condition (the 500-msec base IOI); by holding this rate constant, we can assess responses to it that result in different session contexts.

Constant errors and central tendency measures. Figure 3 summarizes CE data for the eight context conditions of Experiment 2. The 28 CE scores of Figure 3 were averaged over participants and plotted against base IOI. We note that individual CE scores depend on a fixed temporal order in which a sequence base IOI is always followed by one of five comparison IOIs. Nevertheless, these observed CE scores correlate highly ($r = .84$) with a related measure that precludes temporal order reversals—namely, responses to a comparison IOI that is identical to the standard [$P(\text{short/same}) - P(\text{long/same})$]; $p < .01$, two-tailed].⁷

In general, the pattern of CEs, shown in Figure 3, suggests the presence of a P : CEs were positive for sequence rates that were faster than the mean session rate and negative for rates that were slower than the mean session rate. In addition, the absolute magnitude of CEs appeared to be greater for sequence rates more remote from the session mean than for those close to the mean. Consistent with these observations, the CE data in Figure 3 were best described by two separate regression lines: one for fast sets ($R^2 = .72$) and one for slow sets ($R^2 = .67$). These lines identify, respectively, two preferred period values ($P = 426$ msec for fast sets, and $P = 615$ msec for

slow sets). Both estimates are relatively close to the respective mean of each set. This is in sharp contrast to the findings of Experiment 1, where P was indeterminate ($R^2 = .28$).

Next, we consider which central tendency measure, the arithmetic or the geometric mean, provides a better estimate of the location of P for each of the conditions. We estimated P separately for session contexts with $n > 2$ (i.e., conditions with sufficient points for a regression analysis). The resulting P values appear in Table 2. The arithmetic mean of base IOIs is constant within each set (either 400 or 600 msec), but the estimates of P are not. However, the observed variation of P is not explained by the GM (Table 1). The root mean-squared error of approximation (RMSEA) between P estimates and the arithmetic mean of F and S sets was approximately 53 msec, whereas the corresponding RMSEA involving the different values of the GM was greater (i.e., 65 msec in both F and S sets).

Aside from obvious nonsystematic factors, such as individual differences, the variation of P estimates raise questions about the role of additional variables that differentiate the four context conditions within each set. One candidate is the base IOI that reinforces the mean IOI of the session. For example, condition F5 has a session mean of 400 msec and also contains a sequence with a reinforcing base IOI of 400 msec. In general, conditions with a reinforcing base IOI (F5, F3, S5, S3) had smaller RMSEA values between P estimates and mean session rate (47 msec) than did those (F4, F2, S4, S2) that lacked one (68 msec). This suggests an influence of a reinforcing local rate on P , and it implies that simple statistical averaging does not uniquely determine P .

Other factors may also contribute to CEs and hence to variability of P estimates. A major one is session heterogeneity, which distinguishes the conditions within each set. The data of Table 2 do not rule in or out a specific heterogeneity measure (n , SD , or R). In the next section, we consider whether context heterogeneity affects the likelihood of errors by assessing overall accuracy. The rationale for this analysis is based on two observations: (1) P estimates depend upon regression of scores indicating average error magnitude (CE values) over an ordered set of base IOIs; and (2) overall accuracy, indexed by PC, reflects error frequency, which turns out to be related to the variability in P estimates by virtue of its correlation with absolute CE scores ($r = -.41$, $p < .05$).

In sum, analyses of CEs revealed two different P s for fast and slow sets that approximated the session mean, respectively. Within each set, part of the variation in P estimates appeared to be related to the presence or absence of a base IOI that reinforced the session mean.

Accuracy and session heterogeneity. Figure 4 shows mean PC for each base IOI within each context condition. Data are grouped across the four panels (A–D) of Figure 4 according to set (F , S) and R (narrow, wide); within each panel, conditions differ as a function of n and base IOI. Error bars represent 95% confidence in-

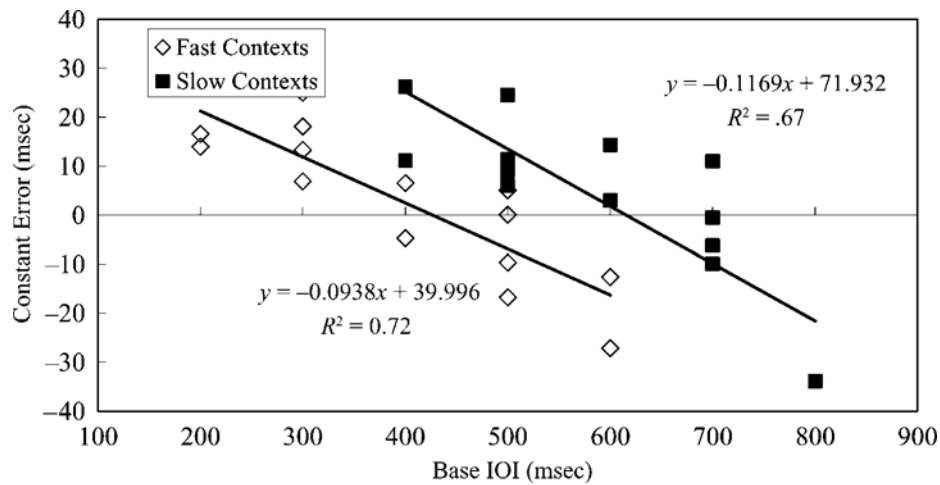


Figure 3. Constant errors as a function of base IOI and set (fast, slow) in Experiment 2. Solid lines show the regression lines for each set. For fast contexts, this line intersects the zero constant error line at 426 msec (P for fast set); for slow contexts, the intersection point is 615 msec (P for slow set).

tervals. Also shown (solid line) are the corresponding means from Experiment 1. In general, overall accuracy is lower in Experiment 2 than in Experiment 1. Global context effects due to multiple versus single rates in a session are most evident in PC scores for specific individual rate conditions of Experiment 2 versus Experiment 1. For instance, the PC scores are often lower for the same base IOIs in Experiment 2 than in Experiment 1. Over the eight Experiment 2 conditions, both the mean

rate (fast vs. slow) and session heterogeneity contribute to PC differences. A 4×2 ANOVA on overall PC scores showed a significant main effect of n [$F(3,75) = 8.23, MS_e = 0.007, p < .01$] and a significant main effect of mean rate [$F(1,75) = 11.61, MS_e = 0.007, p < .01$].

Although this omnibus analysis indicates that session heterogeneity affects error probabilities, it does not identify which of the three heterogeneity measures best predicts PC. Because n , SD , and R generate different predictions about the rank order of the four conditions within each set (see Table 1), we used planned comparisons to test these orderings, under the assumption that greater heterogeneity implies lower PC values (Kirk, 1995). Of these measures, SD was the worst predictor. This measure predicted the order (from highest to lowest PC): $3 > 2 > 5 > 4$ (in terms of n). This was not found. In fact, the range (R) was the best predictor of the observed ordering, which was: $2 = 3 > 4 = 5$. Within each set, those conditions with equivalent R values did not differ in mean PC (all $ps > .05$); PC was higher in the narrow range ($n = 2, 3$) than in the wide range ($n = 4, 5$) conditions in both fast sets [$F(1,39) = 9.79, MS_e = 0.01, p < .01$] and slow sets [$F(1,40) = 11.92, MS_e = 0.005, p < .01$]. Taken together, such findings are not consistent with predictions based on either SD or n . Although SD is unsatisfactory, we cannot fully discount uncertainty as a factor, because R and n covary. However, we can conclude that R appears to play the most important role in determining error rates, because these are reflected by PC values.

Because session mean rate also appeared to affect overall PC (lower PCs were found in fast sets than in slow sets), we combined mean rate and R to form a more economical predictor of PC termed *relative range* (RR). RR is calculated by dividing the range of rates for each

Table 2
Observed Location of P for All Session Context Conditions in Experiments 2 and 3

Context Condition	P' (msec)	P (msec)	Slope	R ²
Experiment 2				
F5	416	390	-0.13	.88
F4	416	480	-0.06	.82
F3	433	436	-0.14	.99
F2	433	n.a.	n.a.	n.a.
S5	624	620	-0.16	.87
S4	624	546	0.11	.90
S3	649	686	-0.03	.99
S2	649	n.a.	n.a.	n.a.
Experiment 3				
F5-WR	416	390	-0.13	.88
F5-NR	433	498	-0.11	.94
F3-WR	416	398	-0.13	.99
F3-NR	433	436	-0.14	.99
S5-WR	624	620	-0.16	.87
S5-NR	649	612	-0.20	.87
S3-WR	624	553	-0.16	.94
S3-NR	649	686	-0.03	.99

Note—P', predicted estimate; P, observed estimate; WR, wide range; NR, narrow range. Bold font indicates overlapping conditions in Experiments 2 and 3. No estimates were reported for F2 and S2 because the calculation of P involved only two points.

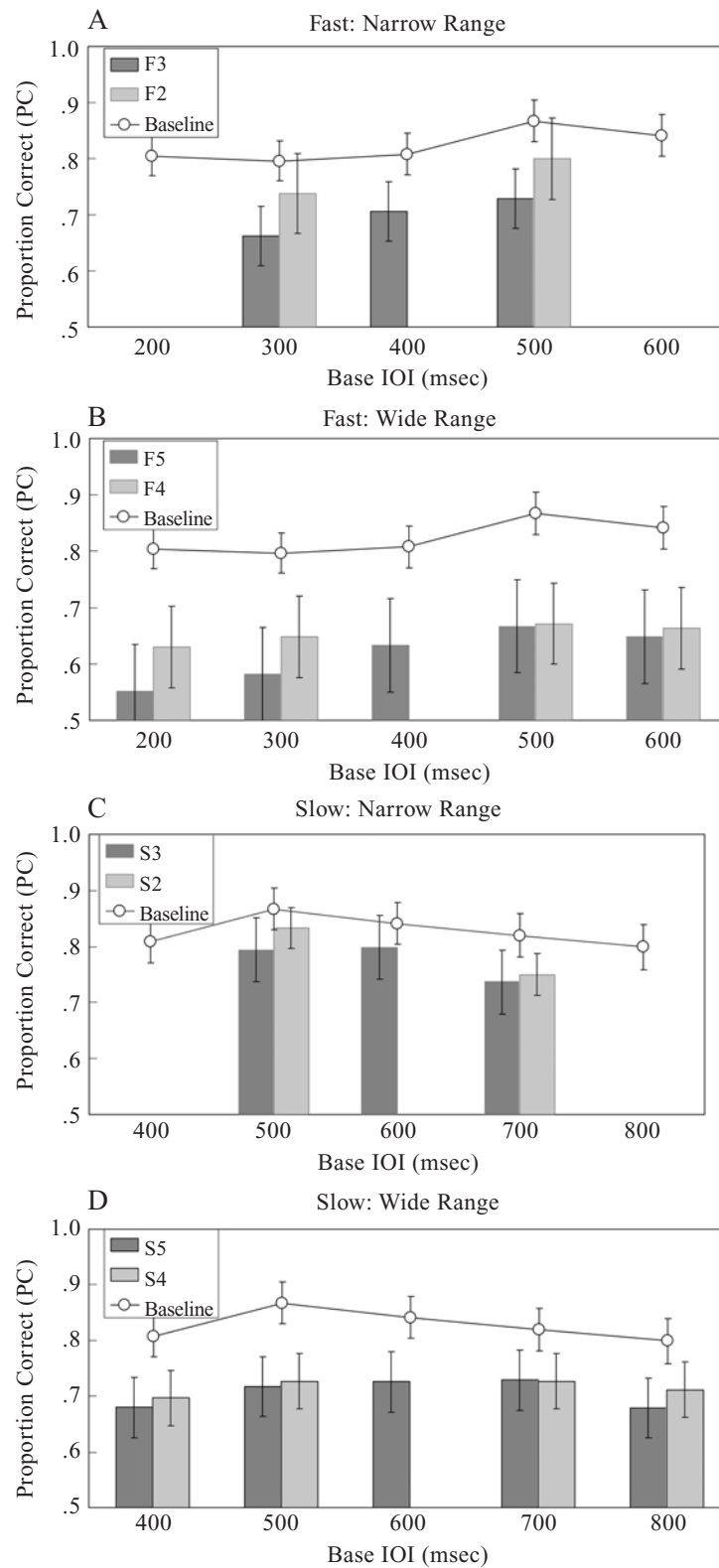


Figure 4. Mean proportion of correct (PC) responses for each base IOI within each of the eight session context conditions of Experiment 2. The solid line in each panel presents mean baseline PC from Experiment 1. (A) Fast sets with a narrow range (RR = .50) for $n = 2, 3$. (B) Fast sets with a wide range (RR = 1.00) for $n = 4, 5$. (C) Slow sets with a narrow range (RR = .33) for $n = 2, 3$. (D) Slow sets with a wide range (RR = .67) for $n = 4, 5$.

session by mean session rate. In Experiment 2, this yielded four RR values ranging from a minimum RR of 0.33 (narrow range, slow rate) to a maximum RR of 1.00 (wide range, fast rate). A one-way ANOVA on PC with the four levels of RR shows a significant main effect of RR [$F(3,79) = 10.68, MS_e = 0.007, p < .01$]. Accuracy over the eight context conditions rank orders extremely well with RR values: People are most accurate when RR = 0.33 ($M = 0.78$), next most accurate when RR = 0.5 ($M = 0.73$) and 0.67 ($M = 0.71$), and least accurate when RR = 1.0 ($M = 0.64$).

Conditional accuracy and trial-to-trial changes in rate. Why does R, and especially RR, successfully predict accuracy/error rates? One possible reason is that RR reflects the largest change in sequence rate that can occur between adjacent trials in a session. Such changes may be highly disruptive if listeners become accustomed to a certain “average pace” over several trials and are then forced to adapt to an extreme departure from this pace. To assess this, we determined PC in the context of different trial-to-trial rate changes in each session. We let PC_n represent the probability of a correct response on trial n . We hypothesized that, on average, more errors would occur after large than after small changes in base IOI from trial $n - 1$ to trial n , hence PC_n should be lower on these trials.

The rate change hypothesis is evaluated by modifying the RR metric to accommodate trial-to-trial rate changes. We defined a *rate change* metric (RC) as $RC = (|\text{local rate change}|) / \text{mean session rate}$, where absolute local rate change values were 0, 100, 200, 300, and 400 msec. Figure 5A shows PC_n regressed against the RC values in Experiment 2. As predicted, accuracy on a given trial, PC_n , is inversely related to the relative size of the RC ($r = -.62$), from trial $n - 1$ to trial n . Alone, the RC metric explains a significant proportion (.38) of the variance in PC_n [$F(1,28) = 17.16, MS_e = 0.014, p < .01$].

The preceding analysis suggests that the global metric, RR, predicts overall accuracy in part because it *capitalizes on detrimental effects of large trial-to-trial changes in local rate*. However, RR and RC are necessarily correlated ($r = .41$). Therefore, the most revealing analyses with respect to the rate change hypothesis are those that examine the impact of local rate changes from an immediately preceding trial when session range (hence RR) is held constant. We assessed RC separately for each of the two wide range context conditions ($R = 400$ msec) that include the most extreme local rates (insufficient points exist in narrow range conditions). In the fast–wide range condition (RR = 1.00), RC and PC_n yielded a correlation of $-.66$; these data are shown in Figure 5B (for comparison with data in Figure 5A where RR varies). We also found a significant correlation between RC and PC_n in the slow–wide range condition (RR = .67): $r = -.50$. Thus, in sessions containing the most extreme base IOIs, 25% to 44% of the variance in errors on a given trial is explained strictly by changes in local rates from the immediately preceding trial. When we combined RR with

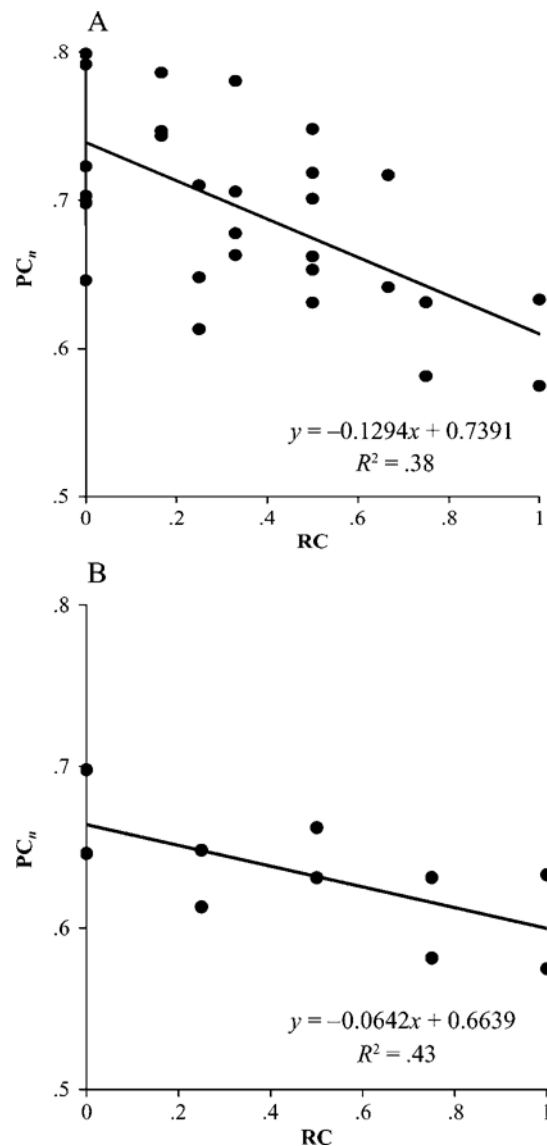


Figure 5. Conditional probabilities of a correct response on trial n (PC_n) as a function of the magnitude of rate change between trial $n - 1$ and trial n . Plots show PC_n as a function of the rate change metric (RC) for (A) all conditions of Experiment 2 and (B) the F5 (RR = 1.0) condition of Experiment 2.

RC in a step-wise regression to predict PC_n and included other heterogeneity predictors (n, SD) for comparative purposes, only RR and RC contributed significantly to the predictability of PC_n [$F(1,28) = 57.89, MS_e = .037, p < .01$, and $F(2,27) = 44.26, MS_e = .032, p < .01$]; together, RR and RC accounted for 76.7% of the variance in PC_n .

In sum, analyses of conditional accuracy allow us to conclude: (1) trial-to-trial changes in base IOI (local rate) significantly affect PC_n , with the largest absolute changes in local rate corresponding to the most errors (lowest PC_n) on a given trial; and (2) effects of RR re-

ported in the analyses of overall accuracy are partially, but not entirely, due to having large trial-to-trial rate changes in session contexts with large relative ranges.

Diagnostic 500-msec base IOI condition. Finally, we examined performance for the 500-msec base IOI only because this sequence rate appeared in all eight context conditions. These analyses permit us to confirm our conclusions that session contexts indeed do distort the time perception of a given rate. Previous analyses included all base IOIs; thus, heterogeneity effects may have resulted from inclusion of a particular base IOI in the analysis.

First, we focused upon CE scores for this condition. A 2×4 ANOVA, with set and n as factors, indicated only a significant main effect of set [fast vs. slow; $F(1,79) = 17.12$, $MS_e = 403.4$, $p < .01$] but no other main effects or interactions. As expected, mean CE scores were negative ($M = -4.9$ msec) in fast contexts where the 500-msec rate was slow, relative to the session mean and positive ($M = 12.9$ msec) in slow contexts where the 500-msec rate was fast, relative to the session mean. Next, we assessed error rates in this condition by performing a one-way ANOVA on PC scores using RR as the single factor. As anticipated, we found a significant effect of RR [$F(3,79) = 8.20$, $MS_e = 0.009$, $p < .01$]; the lowest PC occurred for $RR = 1.0$ ($M = 0.67$) and the highest for $RR = 0.33$ ($M = 0.81$), with intermediate PC values for RR values of 0.5 ($M = 0.76$) and 0.67 ($M = 0.72$).

Finally, we also confirmed general findings involving PC_n for this diagnostic condition. For the 500-msec base IOI condition, the separate correlations between RR and RC with PC_n , respectively, were $r = -.75$ and $r = -.62$ ($ps < .01$). Negative r values confirm the hypothesis that larger trial-to-trial changes in local rate are more detrimental than smaller ones. A step-wise regression with RC and global predictors n , SD , and RR indicated that, as before, only two predictors were significant: RR and RC [$F(1,26) = 34.39$, $MS_e = 0.039$, $p < .01$, and $F(2,25) = 29.49$, $MS_e = 0.034$, $p < .01$, respectively]. The RR metric accounted for 56.9% of the total variance, with RC contributing an additional 13.3%; the total variance accounted for was 70.2%.

In sum, all of the contextual effects evident when session context was evaluated by including all base IOIs are also found when we consider only performance for the single 500-msec base IOI that appeared in every session context condition. Thus, the context effects we report do not appear to be due to differences in the base IOIs comprising the different context conditions, but rather illustrate systematic distortions in how people perceive a particular sequence rate.

Discussion

Overall, Experiment 2 shows that multirate sessions instill in listeners a sense of overall pace, or preferred periodicity (P) that affects time-judgment accuracy. Evidence for P is based upon the distinct pattern of errors that emerged for each session context. With regard to

gross contextual determinants of P, two major candidates emerge: (1) The mean of base IOIs within a multirate session predicts the location of P reasonably well; CEs for sequence rates slower and faster than P tend to be positive and negative, respectively. (2) The session range of local rates—namely, the RR—predicts error rates, as indexed by PC, and has the potential to explain some of the variability in P estimates. In this section, we develop a simple algorithm for predicting P that is based on these properties.

We begin by proposing that the CE for a given base IOI in a particular session context can be predicted using a slightly modified RR metric. Because CE for a given base IOI is signed, we transform RR into a *directional relative range metric* (dRR) by multiplying it by d (the signed deviation score of a base IOI from the session mean). This new predictor variable (dRR) incorporates direction, session heterogeneity, and a session's mean rate. Information from d is associated with the value of a given error (magnitude, sign), whereas the information from RR is associated with the likelihood of such an error within a session context. This metric can be used to predict the location of P for all eight Experiment 2 conditions combined. Figure 6 summarizes CE scores as a function of dRR; open circles (○) represent Experiment 2 data, and ✖ denotes Experiment 3 data. For Experiment 2, dRR accounts for 64.9% of the variance in these data [$F(1,26) = 48.2$, $MS_e = 10.26$, $p < .001$].

A general formula for predicting the location of P as a function of global session context derives from solving the obtained regression equation for the (signed) value of dRR—namely, dRR_0 —that yields $CE = 0$. For any session context, the predicted location of P—namely, P' —is given by:

$$P' = dRR_0/RR + \text{mean session IOI}, \quad (1)$$

where we assume that dRR_0 is a constant that can assume positive or negative values. Thus, Equation 1 indicates that P' is inversely related to RR; it predicts that as the relative range of a session increases, P' gravitates to the session mean (i.e., in the limit $dRR_0/RR \rightarrow 0$). Table 2 shows the predicted (P') and observed (P) estimates of the preferred period, assuming that dRR_0 is under +16.3. Over six of the eight conditions (excluding $n = 2$), the mean absolute difference between predicted (Equation 1) and observed (Table 2) values is 34 msec (under 10% of the mean base IOI).

Equation 1 provides a global predictor of average session PSE, based on gross statistical properties of a session. Because of this, it does not explicitly incorporate two important factors that undoubtedly play a role in its efficacy. One factor involves the likelihood of a condition containing a reinforcer rate. For example, the 34-msec difference between P and P' is reduced to 17.5 msec in sessions that contain a local sequence rate that reinforces the global mean rate. A second factor involves moment-to-moment changes in local rate as gauged by RC (cf. Figure 5). Equation 1 only indirectly incorpo-

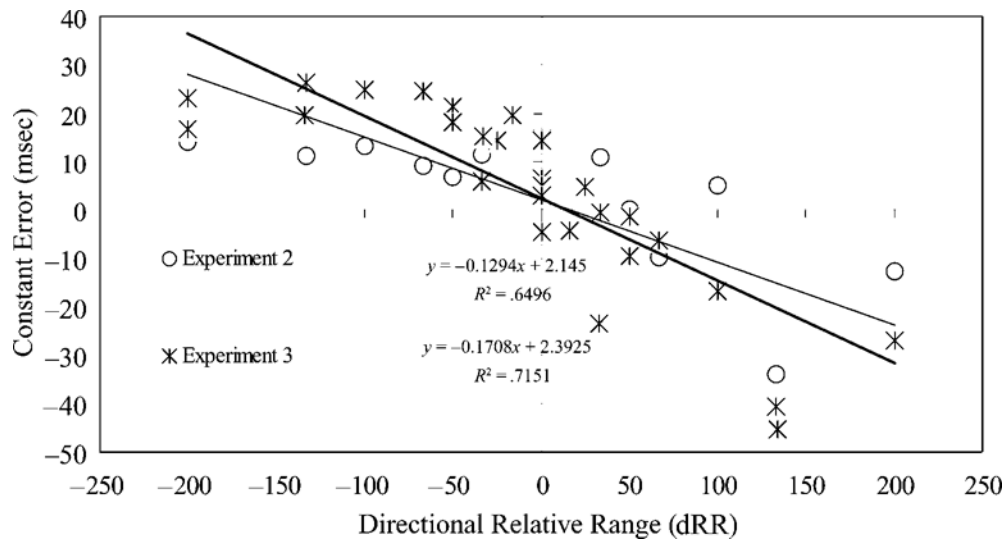


Figure 6. Constant error plotted as a function of directional relative range (dRR) for Experiments 2 and 3. Separate regression lines are shown for Experiments 2 and 3 data.

rates this fine-grained session information; for instance, RC enters the picture through its links to the global metric (RR). However, it is possible to assign a more explicit role for one or both of these factors by assuming that they affect the value of dRR_0 . Thus, dRR_0 may take a smaller value when a reinforcing rate is present than when one is absent. The potential import of such factors instills a cautionary note: A singular emphasis on gross statistics and global determinants of P can mask important local and moment-to-moment influences on the underlying process.

In summary, the most economical global predictor of CEs in time judgments in extended temporal contexts combines RR with a deviation score based on individual base IOIs to form a predictor variable (dRR). The development of the dRR metric leads to a general algorithm that relies on global (not local) distributional properties, such as session mean and range, to predict the location of P.

EXPERIMENT 3 Range and Uncertainty

Experiment 2 data indicate that a metric based on RR accurately predicts effects of global context on CEs, and an algorithm based on this metric predicts the location of P. All of this implies that the range of a session is an important determinant of the preferred period. However, range and uncertainty were partially correlated in Experiment 2. One aim of Experiment 3 was to tease apart the effects of R and n . The second aim was to provide an evaluation of the proposed algorithm, Equation 1.

Experiment 3 introduces new conditions that yield an orthogonal manipulation of R and n when supplemented with certain conditions from Experiment 2. We use the

same mean session rates of 400 msec (Experiment 3A) and 600 msec (Experiment 3B) as in Experiment 2. However, design constraints in Experiment 3 necessitated a denser distribution of local rates within two of the new conditions to achieve the requisite combination of narrow range and large n . Finally, unlike Experiment 2, in Experiment 3 all session contexts contain a reinforcer rate—that is, a base IOI that matches the mean IOI of the session.

Method

Participants. A total of 47 students from an introductory psychology class at Ohio State University volunteered for Experiments 3A ($n = 25$) and 3B ($n = 22$). These numbers constitute 92% of the individuals recruited; data for the remainder were excluded due to failures to comply with instructions, extremely poor accuracy levels, or equipment failure.

Equipment and stimulus construction. Equipment and stimulus construction were identical to those for Experiments 1 and 2.

Design. Mean session rate was either fast (F, 400 msec; Experiment 3A) or slow (S, 600 msec; Experiment 3B). The common design for Experiments 3A and 3B included relevant conditions from Experiment 2; it was a $2 \times 2 \times 2$ mixed factorial design. The between-subjects variables were mean rate (400 msec, 600 msec), n (3, 5), and range (200 msec, 400 msec). Statistical properties of the eight context conditions used in Experiment 3 are summarized in Table 3; note that in Experiment 3, *all* conditions contain a reinforcer rate.

Procedure. The procedure was identical to that of Experiments 1 and 2.

Results and Discussion

The findings of Experiment 3 parallel those of Experiment 2. When we decouple the heterogeneity measures, R and n , we find that range is the primary factor that determines performance. This is evident in analyses of absolute CE scores from Experiment 3 where an ANOVA showed a significant main effect of R [narrow vs. wide; $F(1,81) = 12.83$, $MS_e = 200.3$, $p < .01$], but no effect of

Table 3
Global Distributional Properties of the Eight Session Context
Conditions Constructed for Experiment 3

Context	Base IOIs (msec)	Central Tendency		Heterogeneity		
		Arithmetic Mean	Geometric Mean	<i>n</i>	<i>SD</i>	<i>R</i>
F5-WR	200, 300, 400, 500, 600	400	372.8	5	158	400
F5-NR	300, 350, 400, 450, 500	400	393.6	5	50	200
F3-WR	200, 400, 600	400	363.4	3	200	400
F3-NR	300, 400, 500	400	391.5	3	100	200
S5-WR	400, 500, 600, 700, 800	600	582.7	5	158	400
S5-NR	500, 550, 600, 650, 700	600	595.8	5	50	200
S3-WR	400, 600, 800	600	576.5	3	200	400
S3-NR	500, 600, 700	600	594.4	3	100	200

Note—Bold font indicates conditions that overlap with Experiment 2.

either mean session rate [F vs. S; $F(1,81) = 1.94$, $MS_e = 200.3$, $p = .17$] or n [3 vs. 5; $F(1,81) = 0.06$, $MS_e = 200.3$, $p = .8$]. Mean absolute CE scores were lower in the narrow range ($M = 15.47$ msec) than in the wide range ($M = 26.28$ msec) conditions. The PC data lead to a similar conclusion; significantly fewer errors also occurred in the narrow than in the wide range condition [mean PCs were .72 vs. .66; $F(1,81) = 12.24$, $MS_e = 0.008$, $p < .01$], but no significant differences were due to n [$F(1,81) = 0.31$, $MS_e = 0.008$, $p = .58$]. Session rate also had a significant effect, with faster sets producing lower PC values than slower ones [$F(1,81) = 8.80$, $MS_e = 0.008$, $p < .01$]. Taken together, these data accomplish our first aim—namely, to confirm the findings of Experiment 2 with respect to the role of R . Combined, the two experiments eliminate n and SD as predictors of time-judgment errors.

In the service of our second aim—that is, evaluation of Equation 1, we performed separate regression analy-

ses of CEs for each of the eight context conditions to determine the location of P . Resulting P values appear in Table 2 (conditions from Experiment 2 are indicated in bold). Again, the CE data were combined for the fast and slow sets, respectively, to arrive at two corresponding estimates of the location of P , 430 msec ($R^2 = .83$ for fast session contexts) and 606 msec ($R^2 = .80$ for slow session contexts). These data, and best-fit regression lines, appear in Figure 7. Overall, Experiment 3 regression analyses agree with corresponding ones of Experiment 2 (Figure 3) in showing prominent influences of the arithmetic mean on these P values.

Next, we evaluated dRR as a predictor of the pattern of constant errors and hence of P . The relevant CE data are shown as a function of dRR in Figure 6 (data points of Experiments 2 and 3 are denoted by \circ and \times , respectively). This figure indicates that dRR accounts for 71.5% variability in the CE scores over all eight Experiment 3 conditions. When applied to all CE data—that is,

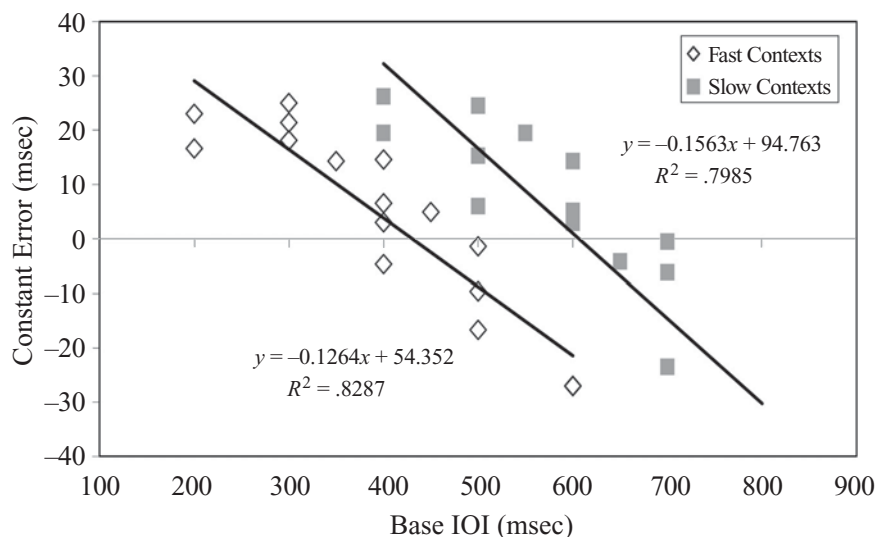


Figure 7. Constant error scores as a function of base IOI and set (fast, slow) in Experiment 3. Solid lines show the regression line for each set. For fast contexts, this line intersects the zero constant error line at 430 msec (P for fast set); for slow contexts, the intersection point is 606 msec (P for slow set).

from Experiments 2 and 3—the dRR metric accounts for 64.2% of the variance in constant errors.

Finally, we used Equation 1 to predict the observed location of P for each of the four new session contexts in Experiment 3. For generality, we assumed the same dRR_0 value (16.3 msec) as for Experiment 2. The resulting P' values appear in Table 2; the mean absolute deviation between observed and predicted preferred period (P vs. P') was 33 msec, which is very similar to that observed for Experiment 2.

In summary, Experiment 3 confirms that R , and not n , is the best heterogeneity measure. In addition, Experiment 3 data provide converging support for Equation 1; they support the hypothesis that the location of P in any session context is determined by a combination of mean rate and session range.

GENERAL DISCUSSION

Three experiments examined the effects of context in multirate sessions on the development of a preferred period. In the introduction, we proposed two broad goals of this research. One was to examine global (distributional) effects of session context on a listener's sense of pace. The second was to assess the influence, if any, of trial-to-trial rate changes on listeners' time judgments. In this research, we identified general statistical properties of a session (e.g., mean and range of rates) that contributed to listeners' overall sense of pace in multirate sessions. In addition, we found that momentary trial-to-trial changes in local rate significantly influence accuracy and contribute to the potent effects of session range on the emergence of a preferred period. In this section, we discuss these findings in terms of the spectrum of context factors that appear to influence time judgments and then consider the broader implications of this research.

The Spectrum of Context Factors

Three levels of context appear to contribute to an individual's response to a given time interval. These involve, respectively, the global distributional properties of a session, trial-to-trial influences, and the local rate of a sequence on a given trial.

Global context. The major goal of the present research was to isolate principal distributional properties of a temporal context that contribute to a listener's sense of pace. Most relevant to this goal were Experiments 2 and 3, in which local rate was decoupled from global session rate in multirate sessions. In these experiments, it was clear that session context had striking effects on overall accuracy (PC) and constant errors (CE). Indeed, the changing pattern of CE scores in the different session context conditions provided the strongest evidence that the emergence of a preferred periodicity (P) depends on the global distributional properties of a session.

We considered two possible predictors of P associated respectively with different measures of central tendency: the arithmetic mean and the geometric mean. In both the

fast and slow context conditions, we found that the preferred period (P) was best predicted by the arithmetic mean, not the geometric mean, of the base IOIs of the session. Overall, P was closer to 400 msec in sessions with mean rates of 400 msec and closer to 600 msec in those with mean rates of 600 msec. Although the present data clearly favor the arithmetic mean, we do not rule out the possibility that sessions with broader (or different) distributions of session rates might tilt toward the GM (Wearden & Ferrara, 1995, 1996). Moreover, the arithmetic mean was not a perfect predictor. Although it provided very good estimates of P in some conditions (i.e., within 2 to 12 msec), in others this mean was wide of the mark by as much as 98 msec. In short, the arithmetic mean is part of the story, but it is not the whole story.

Session heterogeneity provides another piece of the story. Heterogeneity effects were evident in both the magnitude and likelihood of errors. Here, session range appears to play an important role. With respect to likelihood of errors, Figure 4 revealed higher PC levels for the 500-msec base IOI in the two narrow-range conditions (F, S), with the lower levels for the same base IOI in corresponding wide-range conditions of Experiment 2. Not only was overall accuracy much lower (and error likelihood higher) in these multirate conditions than in the single-rate conditions (Experiment 1), it was also much lower in those multirate conditions where session range was wide (holding mean session rate constant).

When session range was normalized by mean session rate, the resulting relative range metric (RR) explained almost 70% of the variance in overall accuracy. People made the most errors when a session had a wide range of rates that were, on average, fast and the fewest errors when a session manifested a narrow range of local rates that were, on average, slow. Similarly, taking account of the magnitude and direction of these errors, the best predictor of CE scores involved RR, weighted by deviation scores (i.e., the dRR metric). The dRR metric predicted 60%–70% of the variance in CE scores in Experiments 2 and 3. Ultimately, CE errors, regressed over base IOIs, revealed the signature pattern of over- and underestimations that we identify with listeners' reliance on a preferred period (or periods) where base IOIs with relatively fast local rates tend to be overestimated and those with relatively slow rates tend to be underestimated. These errors, which are symptomatic of failures to adapt to a local rate, were not random; rather, they reflected an overall bias toward those base IOIs that were central in each multirate condition.

Equation 1, which was developed to explain the observed pattern of constant errors, predicts the preferred internal periodicity for any condition (session) as a function of RR, dRR (i.e., dRR_0), and the session mean. It predicts that P' varies inversely with RR and has two limiting conditions. For a given (signed) value of dRR_0 : (1) If RR is arbitrarily small, due to a narrow range, P' assumes an arbitrarily large absolute value, and a flat regression line should emerge over CE scores. This ap-

proximates the case of Experiment 1, where RR was zero. Indeed, in Experiment 1, Equation 1 correctly predicts an indeterminate P and a flat regression line. (2) As RR becomes arbitrarily large, due to gross widening of the range (relative to mean rate), according to Equation 1, P' converges to the mean session rate. In the limit, with large RR, the arithmetic mean is the best global estimator of P. Thus, in Experiment 3 the two conditions with the largest RR values (i.e., RR = 1.00 in F3-WR and F5-WR due to wide range and fast mean rate) yielded predicted and observed P estimates quite close to the session mean rate (i.e., with deviations, on average, within 6 msec of 400 msec).

Finally, we have cautioned that Equation 1 summarizes primarily global properties of a session and their influence on P. However, other session properties, which may be only indirectly captured by Equation 1, may also play a role. One of these involves trial-to-trial rate changes, which we consider next.

Trial-to-trial context. The second goal of this research was examination of the effects of trial-to-trial rate changes. In terms of errors, it is clear that the likelihood of making a mistake in judging a comparison time interval is greater following a large change in local rate from one trial to the next than following a small change. Evidently, the perception of sequence rate on one trial depends upon the lingering effects of the rate that is experienced on an immediately preceding trial (cf. Jungers, Palmer, & Speer, 2002). Our findings in this respect are noteworthy because they offer a glimpse of systematic moment-to-moment changes in accuracy. Moreover, they suggest that one reason the range, which is determined by endpoints of an abstract time continuum of base IOIs, is important in Equation 1 is that it reflects the largest rate change in *real time* from one trial to the next that a listener encounters in a given session.

Others have noted similar trial-to-trial effects (Treisman, 1963; Warren, 1985). Warren suggests that a dynamic criterion shift rule is required to describe what seem to be continual adjustments to novel stimulation, although he does not spell out a trial-to-trial application of such a rule. A useful feature of entrainment models is that they do tackle this issue (e.g., Large & Jones, 1999). In the latter vein, the present findings are compatible with the basic assumption of entrainment models that adaptation to a new rate should be differentially compromised when a relatively large rate change is involved.

This interpretation of a dynamically shifting preferred period is supported by analyses of trial-to-trial responding. Over all sessions, 38% of the variance in error frequency on a given trial (PC_n) was explained by the difference between the local sequence rate on one trial and that of the preceding trial. Thus, a large tempo change, such as from a slow local rate to a much faster one (on trials $n-1$ and n), renders a listener more likely to judge a shortened comparison on trial n to be "same" or "shorter" than its base IOI than would be the case with a smaller rate change. The importance of moment-to-moment

changes is evident from analyses that hold constant relative range and assess only effects due to local rate changes within a session (indexed by RC). For instance, with RR constant, between 25% and 44% of the variance in errors on any given trial can be explained by the trial-to-trial rate change.

Local context. Finally, in a task that asked listeners to judge the duration of a comparison interval relative to the duration specified by the local rate of the preceding sequence, it is not surprising to find that listeners did what they were asked. Listeners' reliance on local rate was readily inferred from the PSE scores for each base IOI. Indeed, although the analysis of the diagnostic 500-msec condition confirmed that the context of multirate sessions systematically distorted perception of individual local rates, these data also speak to the dominance of the local context rate on judgments. On average, the mean CE values in the 500-msec context condition ranged from -17 msec to +24 msec, reflecting error magnitudes of less than 5%. Such findings converge with those of McAuley & Jones (2003), as well as others (e.g., Barnes & Jones, 2000; Drake & Botte, 1993; McAuley & Kidd, 1998) who found that listeners tended to rely on local context rate even in cases where they are told to ignore it.

Because local rate is such an important determiner of performance, it appears to offer an intriguing link to global context effects in the form of the local rate that reinforces the mean rate of the session. In sessions containing a sequence with a local rate that reinforced the mean rate of a session, the predicted value of P (i.e., P') tended to be closer to observed estimates than in conditions lacking a local rate that reinforced the mean rate. Such findings imply that explanations of globally induced time distortions must ultimately accommodate the primacy of local context effects.

Broader Implications

Our interpretations speak to situations where people must respond to time intervals embedded in contexts where they experience a variety of pattern rates. As such, these data have implications for understanding the development of sensitivities to both local and global tempi in music and other natural contexts. They suggest that in listening to musical events, where tempo routinely changes over time, listeners should exhibit systematic errors in judging an individual local tempo. Tempo distortions should arise from lingering effects of prior tempi (Madison, 2001). In addition, a listener's emerging sense of an overall pace (i.e., a prevailing beat) in a musical event will be drawn to those local tempi that are centrally located within the distribution of experienced tempi. Together, these local tempi will dominate in a listener's memory to specify the stable pace of the piece. Finally, however, this construct of a global tempo, as indexed by P, should not obscure the fact that it rests on real-time experience.

Although entrainment models are designed to describe tempo tracking by explicitly addressing trial-to-trial rate

changes, none have addressed the emergence of a sense of overall pace based on the larger context. In an entrainment framework, the global sense of pace might emerge from the activity of a single dominant adaptive oscillator, or, alternatively, from reorganizing properties of an ensemble of many adaptive oscillators. Determining which of these explanations is correct is a problem for theoretical modeling and is beyond the scope of this article. Nevertheless, applications of such entrainment constructs to time judgments suggest some potential for predictions about historically familiar phenomena, such as *assimilation* (adaptation) and *contrast* (failure to adapt) and gravitation to a general adaptation level (e.g., Helson, 1964; Warren, 1985). Although entrainment theories appear to converge with general adaptation-level approaches, we do not claim a broader application of our findings to dimensions other than time.

We return to an unanswered question regarding the relationship between P (or P') and the indifference interval. For Fraisse, a biological periodicity is somehow responsible for indifference interval phenomena. Experiment 1 seems to cast doubt on this explanation; moreover, Experiments 2 and 3 suggest the importance of an environmental shaping of some internal periodicity or periodicities. However, recent research suggests that the periods of active biological oscillations may differ with age. Drake et al. (2000) found that young children preferred sequence rates that were significantly faster than those preferred by older children. Thus, in the present research wherein the ranges of local rates included those most comfortable for young adults, we may not see large influences of biologically determined preferred periods.

As we have seen, classical approaches often reject the idea that periodic processes explain the indifference interval in favor of interval time constructs using adaptation-level calculations (e.g., Hellstrom, 1985; Helson, 1964). For reasons outlined earlier, we have refrained from equating a preferred periodicity with an indifference interval. Yet, as already noted, it is difficult to ignore certain conceptual resemblances between adaptive oscillators, preferred periods, and various constructs used in adaptation-level theories. Moreover, with respect to the indifference interval, which is calculated as a null error point within a set of individual time intervals, using the two-interval paradigm, an undeniable counterpart is found in the preferred periodicity, which is taken as a null error point within a set of sequence rates, using the sequence paradigm. However, as conventionally calculated, the indifference interval often rests on TOEs, which requires a temporal reordering of the two individual time intervals involved; TOE calculations putatively assess the relative likelihood of errors in internal time codes for one or both of the intervals. By contrast, P is calculated from constant errors for a given sequence rate (base IOI). This methodological/computational difference is accompanied by a corresponding interpretative difference: in the latter, the sign and magnitude of a CE theoretically reflect the degree to which an entraining oscillator fails

to adapt to a given sequential rate. Note that the practice of reordering individual intervals, intrinsic to TOE calculations, inevitably changes the rate and/or rhythm of the local context in the sequence paradigm, thereby obscuring an assessment of adaptation. Hence, possible theoretical differences or commonalities between the indifference interval and P are obscured by methodological/computational differences.

One way of adjudicating these differences entails a broader interpretation of temporal context, one that incorporates between-trial timing (local rate change) as well as within-trial timing (e.g., local sequence rate). A case can be made for this interpretation from the present research, where we found that the effective context on trial n appears to include the sequence on trial $n-1$. Given a broader view of temporal context, we can ask whether it is possible to adapt Equation 1, which predicts P', to predict values of an indifference interval. The two-interval paradigm, which is typically used to estimate the indifference interval, can now be viewed as a limiting case of the sequence paradigm where within-trial timing involves a short sequence of length one (i.e., having a single base IOI), followed by a comparison interval. This limiting case has been studied by McAuley and Jones (2003), who successfully applied an entrainment model to it. In this interpretation, a listener may track a series of individual time intervals—that is, short sequences—over trials in a session. All other things being equal, if entrainment occurs in such a situation, adaptation should be less efficient than in sessions comprising longer, homogeneous sequences; this is due to the increased density of moment-to-moment time changes associated with within- and between-trial timing in sessions with short sequences (cf. Large & Jones, 1999). However, because Equation 1 ignores trial-to-trial information, it can be applied to such two-interval sessions to ascertain whether P' (derived from CE scores) predicts the value of an indifference interval that emerges over these trials (i.e., as measured by TOE scores). The most important consequence will involve the value of R, the session range. In a true sequential paradigm, R is constrained by the asymmetry of standard and comparison IOIs in that it reflects the range of only base IOIs. By contrast, in a typical two-interval paradigm, where order reversals of standard and comparison IOIs occur, R will necessarily encompass comparison as well as the standard IOI values. Therefore, the value of the range will generally be larger in the two-interval paradigm than in the asymmetrical (sequential) case. Nevertheless, if a commonality obtains between P and the indifference interval, following Equation 1, we expect that in sessions with randomly arranged pairs of intervals, the indifference interval will be predicted by Equation 1 and will manifest gravitation to the session mean with increases in RR as implied by Equation 1.

In sum, recent experimental investigations inspired by entrainment theory have concentrated on the effects of manipulations of local context on time judgments. The present research represents an attempt to understand the

role of an extended (i.e., global) context on listeners' reaction to local sequence rates. In this regard, we find lawful effects of global context, reflected in the emergence of preferred periods; these effects can be understood in terms of limits on the adaptive functions of an entraining oscillator. Finally, they lead to an enlarged interpretation of the context within which a time judgment occurs.

REFERENCES

- ALLAN, L. G. (1977). The time-order error in judgments of duration. *Canadian Journal of Psychology*, **31**, 24-31.
- 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.
- BOLTZ, M. (1994). Changes in internal tempo and effects on learning and remembering of event durations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1154-1171.
- DRAKE, C., & BOTTE, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics*, **54**, 277-286.
- DRAKE, C., JONES, M. R., & BARUCH, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, **77**, 251-288.
- 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.), *Handbook of perception VIII: Perceptual coding* (pp. 203-254). New York: Academic Press.
- FRAISSE, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, **35**, 1-36.
- GESCHIEDER, G. (1997). *Psychophysics: The fundamentals* (3rd ed.). Mahwah, NJ: Erlbaum.
- GIBBON, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, **84**, 279-325.
- GOLDSTONE, S., LHAMON, W. T., & BOARDMAN, W. K. (1957). The time sense: Anchor effects and apparent duration. *Journal of Psychology*, **44**, 145-153.
- GREENBERG, M. G. (1965). A modification of Thurstone's law of comparative judgment to accommodate a judgment category of "equal" or "no difference." *Psychological Bulletin*, **64**, 108-112.
- GRONDIN, S. (2001). From physical time to the first and second moments of psychological time. *Psychological Bulletin*, **47**, 22-44.
- HALPERN, A. R., & DARWIN, C. J. (1982). Duration discrimination in a series of rhythmic events. *Perception & Psychophysics*, **31**, 86-89.
- HELLSTROM, A. (1979). Time errors and differential sensation weighting. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 460-477.
- HELLSTROM, A. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin*, **97**, 35-61.
- HELSON, H. (1964). *Adaptation-level theory*. New York: Harper & Row.
- HIRSH, I. J., & WATSON, C. S. (1996). Auditory psychophysics and perception. *Annual Review of Psychology*, **47**, 461-484.
- HOLLINGSWORTH, H. L. (1910). The central tendency of judgement. *Journal of Philosophy, Psychology & Scientific Methods*, **7**, 461-468.
- JAMIESON, D. G., & PETRUSIC, W. M. (1975). The dependence of time-order error direction on stimulus range. *Canadian Journal of Psychology*, **29**, 175-182.
- JONES, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, **83**, 323-355.
- JONES, M. R. (2004). Attention and timing. In J. G. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 49-85). New York: Academic Press.
- JONES, M. R., & YEE, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 693-703.
- JUNGERS, M. K., PALMER, C., & SPEER, S. R. (2002). Time after time: The coordinating influence of tempo in music and speech. *Cognitive Processing*, **1**, 21-35.
- 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.
- KIRK, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole.
- LARGE, E. W., & JONES, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, **106**, 119-159.
- MACMILLAN, N. A., & CREELMAN, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- MADISON, G. (2001). Variation in isochronous tapping: Higher order dependencies as a function of intertap interval. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 411-422.
- MCAULEY, J. D. (1995). *Perception of time phase: Toward an adaptive oscillator model of rhythmic pattern processing*. Doctoral dissertation, Indiana University.
- MCAULEY, J. D., & JONES, M. R. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short interval timing. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 1102-1125.
- 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 & Performance*, **24**, 1786-1800.
- MICHELS, W. C., & HELSON, H. (1954). A quantitative theory of time-order effects. *American Journal of Psychology*, **67**, 327-334.
- MONAHAN, C. B., & HIRSH, I. J. (1990). Studies in auditory timing: 2. Rhythm patterns. *Perception & Psychophysics*, **47**, 227-242.
- POVEL, D. J., & ESSENS, P. J. (1985). Perception of temporal patterns. *Music Perception*, **2**, 411-440.
- 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.
- THURSTONE, L. L. (1927). A law of comparative judgment. *Psychological Review*, **34**, 273-286.
- 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**, 1-31.
- TURCHIOE, R. M. (1948). The relation of adjacent inhibitory stimuli to the central tendency effect. *Journal of General Psychology*, **39**, 3-14.
- VIERORDT, K. (1868). *Der Zeitsinn nach Versuchen*. Tübingen: Laupp.
- WALLIN, J. E. W. (1911a). Experimental studies of rhythm and time. *Psychological Review*, **18**, 100-131.
- WALLIN, J. E. W. (1911b). Experimental studies of rhythm and time: II. The preferred length of interval (tempo). *Psychological Review*, **18**, 202-222.
- WARREN, R. M. (1985). Criterion shift rule and perceptual homeostasis. *Psychological Review*, **92**, 574-584.
- WEARDEN, J. H., & FERRARA, A. (1995). Stimulus spacing effects in temporal bisection by humans. *Quarterly Journal of Experimental Psychology*, **48B**, 289-310.
- WEARDEN, J. H., & FERRARA, A. (1996). Stimulus range effects in temporal bisection by humans. *Quarterly Journal of Experimental Psychology*, **49B**, 24-44.
- WOODROW, H. (1934). The temporal indifference interval determined by the method of mean error. *Journal of Experimental Psychology*, **17**, 167-188.
- WOODROW, H. (1951). Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1224-1236). New York: Wiley.
- WOODWORTH, R. S., & SCHLOSBERG, H. (1954). *Experimental psychology* (rev. ed.) New York: Holt.

NOTES

1. There are a number of theory-dependent formulas for measuring TOE that assess asymmetries in shorter/longer judgments as a function

of the presentation order of the standard and comparison IOIs. A common equation for the calculation of TOE in the two-interval task is $TOE = 1/2 [P(\text{correct}/SL) - P(\text{correct}/LS)]$, where S and L refer to short and long intervals, respectively.

2. The interpretation of TOE has a long, distinguished, and challenging history (for an excellent review, see Hellstrom, 1985). In general, the causes of TOE are probably multiple and remain a topic of debate.

3. At this stage of theorizing, the separation of a perception stage from a subsequent criterion setting in a decision making stage is virtually impossible. These distinctions rest upon a number of assumptions that are not met in the dynamic environments we consider. Furthermore, even if such stages could be justified, teasing them apart on a moment-to-moment basis is certainly problematic both empirically and theoretically. However, stage distinctions become less problematic within an entrainment framework where moment-to-moment changes in attending are described (e.g., Large & Jones, 1999).

4. In the present design, order reversals of each (single) comparison with the final sequence IOI, for example, for each local rate condition, would confound both order and comparison magnitude with degree of rhythmic disruption of the context sequence.

5. We also analyzed the distribution of "same" responses to all comparisons for a given base IOI. Following an adaptation of the method of adjustments for determining PSE (Gescheider, 1997), we computed the mean of this distribution and determined the comparison IOI corresponding to this value for all conditions of each experiment. These PSE estimates were very close to those generated by our difference method.

6. ANOVAs were also performed that included trial block and random order (I, II, III) as factors. Neither factor produced significant main effects or interactions ($p > .05$).

7. A restricted range analysis that involves only "same" responses to *same* comparisons also yields a significant correlation ($r = -.42$) between abs (CE) and P ("same"*|same*; $p < .05$, two-tailed).

APPENDIX

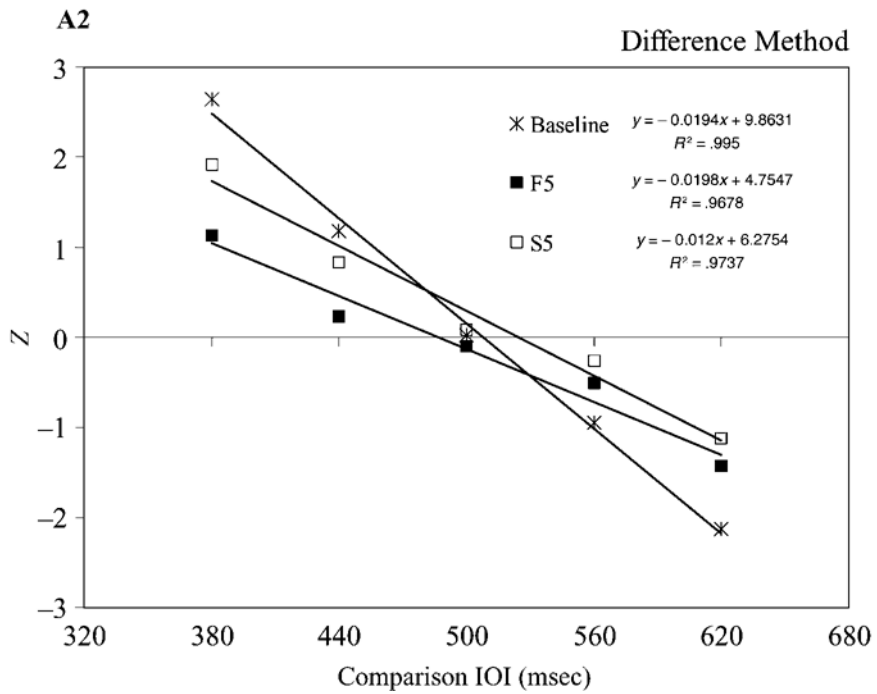
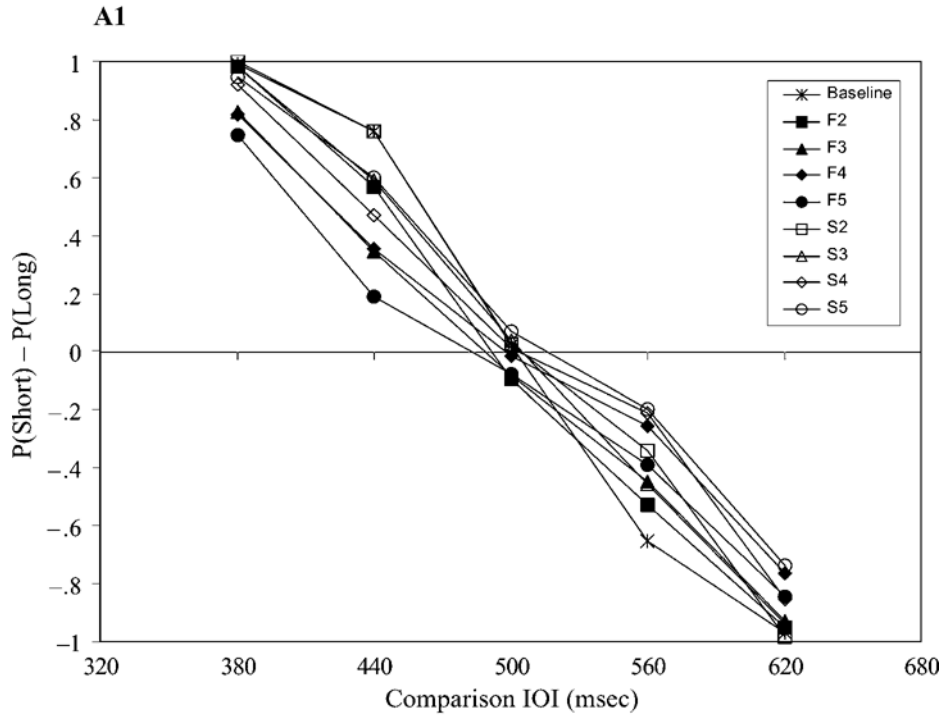
Psychometric functions for the three-response category (shorter, same, longer) task used to estimate point of subjective equality (PSE) and hence constant errors. Table A1 compares PSE estimates using our difference method with those obtained using the method proposed by Greenberg (1965) for three response category tasks. Three figures follow. Figure A1 shows $P(\text{short}) - P(\text{long})$ for the 500-msec base IOI for Experiment 1 (baseline) and all eight session context conditions in Experiment 2. Figures A2 and A3 compare the z transforms for our difference method and Greenberg's method for the baseline, F5, and S5 conditions. All graphs report data averaged over subjects.

Table A1
Estimates of PSE for the 500-msec Base IOI Using the
Proposed Difference Method and the Modified Thurstone
Model Proposed by Greenberg (1965)

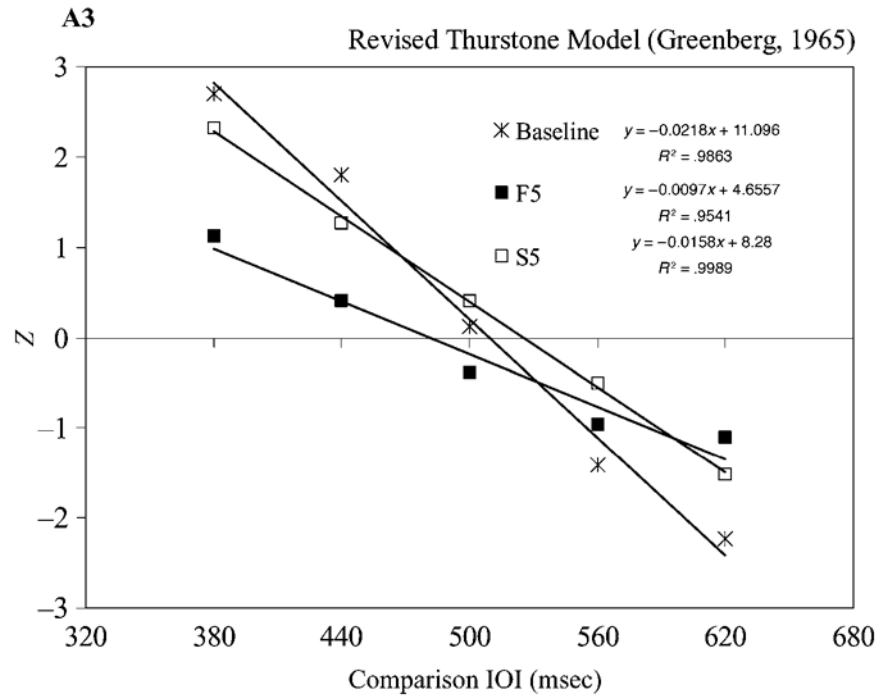
Context Condition	Estimate of PSE for 500-msec Base IOI	
	Difference Method	Modified Thurstone Model
Baseline	504	508
F5	483	480
F4	505	500
F3	490	482
F2	500	500
S5	524	524
S4	509	508
S3	506	506
S2	511	509

Note—Estimates from the two methods were significantly correlated ($r = .97$).

APPENDIX (Continued)



APPENDIX (Continued)



(Manuscript received November 3, 2003;
 revision accepted for publication June 10, 2004.)