

Contextual relative temporal duration judgment: An investigation of sequence interruptions

MELODY S. BERENS and RICHARD E. PASTORE
State University of New York, Binghamton, New York

How do listeners judge relative duration? There currently are three primary classes of proposed timing mechanisms: interval based (judgments of discrete events), beat based (judgments of beat synchrony), and oscillator based (judgments of relative phase or synchrony). In the present research, these mechanisms were examined in terms of predictions both about how an induction sequence of preceding intervals affects relative temporal duration comparison in a simple, two-interval task and about how a pause (e.g., an interstimulus interval) within the presented sequence affects relative temporal duration judgment. Results indicated that a relative temporal judgment was best when all intervals preceding the to-be-judged interval were equal in duration to the to-be-judged interval. Results also indicated that whereas short pauses significantly impair a relative temporal duration judgment, long pauses generally do not reduce the effectiveness of the induction sequence. The results are not entirely consistent with current conceptualizations of any of the proposed mechanisms but can be fully accommodated with simple modifications to the oscillator-based mechanism.

A variety of evidence seems to indicate that context may affect the ability to judge the relative duration of intervals (e.g., Barnes & Jones, 2000; Ivry & Hazeltine, 1995; Large & Jones, 1999; McAuley & Jones, 2003; Pashler, 2001), but the different findings do not provide a clear understanding of the benefits of redundant temporal information and of the cost or benefits of pauses in temporal sequences. In the present study, we examine context effects for a simple, ordered *same-different* task involving two judgment intervals: a *target* (T) interval followed by a *comparison* (C) interval (see Figure 1A). Context is provided either by altering the temporal spacing of these intervals (Experiment 1; see Figures 1A and 1B) or by manipulating an added contextual *induction sequence* that precedes the judgment intervals (Experiments 2 and 3; see Figures 1C and 1E). With such an induction sequence, Barnes and Jones (2000, Experiment 2) found that a relative temporal duration judgment was most accurate when the duration of the contextual induction sequence interval was equal to the T interval. Three different types of timing mechanisms have been proposed to account for such results: interval, beat, and oscillator.

The present research had three major goals. Our first goal was to evaluate the alternative timing mechanisms and to test models based on these timing mechanisms; the mechanisms and, thus, the timing models differ in both whether and how induction sequence information

can alter the effectiveness of the processing components involved in a relative temporal duration judgment. Because there is some evidence that listeners with musical training may process rhythmic sequences more accurately than do listeners without musical training (e.g., Snyder & Krumhansl, 2001), our second goal was to examine the importance of musical experience in performing the basic relative temporal judgment task both with the presence and with the absence of contextual information. Finally, pauses (e.g., an interstimulus interval, or ISI) have been inserted between the induction sequence and the judgment intervals (e.g., Barnes & Jones, 2000; Large & Jones, 1999; McAuley & Jones, 2003; see Figure 1E), between the judgment intervals (e.g., Barnes & Jones, 2000; Large & Jones, 1999; McAuley & Jones, 2003; Pashler, 2001; see Figures 1B and 1D), or both (e.g., Barnes & Jones, 2000; Large & Jones, 1999; McAuley & Jones, 2003; not depicted). Such pauses may affect relative temporal duration judgment accuracy, such as by disrupting a cyclic timing mechanism or, possibly, by creating the functional equivalence of a brief contextual induction sequence. Our third goal was to separately investigate the effect of inserting a pause between the judgment intervals (Experiment 1) and a pause between the induction and the judgment intervals (Experiment 3). The results from these investigations are evaluated in terms of the timing models.

The Timing Mechanisms

The timing mechanisms and, thus, the timing models based on these mechanisms differ with respect to the manner in which initial induction information can be used to alter the relative duration judgment of the T and C intervals. In order to understand any timing mechanism, we first need to understand the requirements for perform-

This study was supported in part by Grant 25990 from the National Science Foundation to the second author. Correspondence concerning this article should be addressed to M. S. Berens, Department of Psychology, State University of New York, P.O. Box 6000, Binghamton, NY 13902 (e-mail: mberens@binghamton.edu).

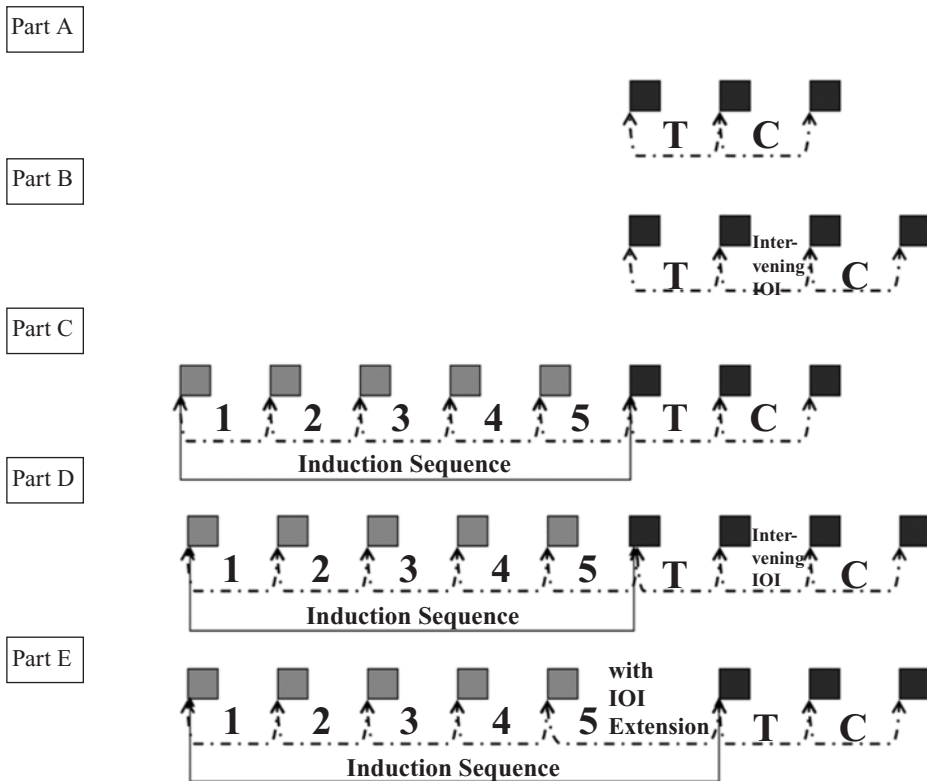


Figure 1. (A) Adjacent judgment intervals (target, T, and comparison, C) are periods of silence that are marked by sound with no interstimulus interval (ISI) between the judgment intervals. (B) Separated intervals are also periods marked by sound, but with an ISI separating the judgment intervals; these stimuli are compared in Experiment 1. (C) The simplest contextual stimulus configuration, with an adjacent induction sequence followed immediately by the two adjacent judgment intervals, used in Experiment 2. (D) Adjacent induction sequence followed by separated judgment intervals, used in other studies (see the text). (E) Adjacent induction sequence with the interonset interval (IOI) of the fifth induction sequence interval extended, followed by adjacent judgment intervals, used in Experiment 3. Judgment interval marker tones are depicted by dark gray squares, and induction sequence marker tones are depicted by light gray squares.

ing the basic two-interval comparative temporal judgment task, as well as the nature of limitations to performing this basic task in the absence of a contextual induction sequence. We believe that the basic (two-interval) relative temporal judgment task (Figure 1A) involves (1) the encoding of duration information about the T interval, (2) the storage of this information, (3) the encoding of duration information about the C interval, and (4) a decision based on a comparison of the encoded information about the T and C intervals. What are the limits to performing this task? One possible answer, which we call the *processing variability* hypothesis, is that there is noise (imprecision, jitter, or variability) in the encoding of the two temporal intervals (e.g., Creelman, 1962; Gibbon, 1977). The induction sequence might then provide information that will enable the timing mechanism to overcome this limitation.

Interval-based mechanism. In terms of our task conceptualization, the interval-based timing mechanism estimates, encodes, and stores the duration of the interval(s) preceding the C interval and then compares some

aspect of the stored information with the estimated duration of the C interval (see Table 1; e.g., Keele, Nicoletti, Ivry, & Pokorny, 1989; Pashler, 2001; Schulze, 1978). There currently are two distinct models of the interval-based mechanism, which Keele et al. call the *adjacency model* and the *memory model*. In the adjacency model (e.g., Keele et al., 1989; Schulze, 1978), duration information about an interval is estimated, is stored briefly, and is available for comparison only with the next interval (see Table 1). Because the stored duration information about any interval is replaced by duration information about the next presented interval, there can be a temporal duration comparison only between adjacent intervals. In the present task (e.g., Figure 1C), because all timing information (e.g., induction sequence intervals) prior to the T interval is lost (or replaced), the induction sequence should have no effect on temporal duration judgment performance; the adjacency model thus represents a null hypothesis for context.

The interval-based memory model is a generic alternative hypothesis, in which an induction sequence can

Table 1
Summary of Timing Mechanisms and Models Based on Mechanisms

| Mechanism | Model | Summary |
|----------------------------|---------------------------------|---|
| Interval-based mechanism | Adjacency model | Compares intervals measured using an interval metric Only compares duration of adjacent or consecutive intervals |
| | Original memory model | Compares last interval with average standard |
| | Multiple-look model | Original memory model + average standard incorporates variability |
| | Multiple-look/exclusion model | Multiple-look model + exclusion criterion |
| Beat-based mechanism | Rigid beat-based model | Internal beat established; duration judgment based on synchrony/asynchrony distinction between internal beat and end marker of C interval Internal beat is rehearsal of duration of first interval |
| | Beat-based/phase correction | Rigid beat-based model + full phase correction |
| | Beat-based/adaptive model | Beat-based/phase correction model + period correction |
| | Beat-based/suspension model | Beat-based/adaptive model + exclusion criterion |
| Oscillator-based mechanism | | Internal oscillator entrains to external sequence and duration judged on the basis of synchrony/asynchrony between internal oscillator and the external sequence |
| | Oscillator-based model (single) | Internal oscillator can exhibit phase and period correction; attentional pulse width reflects redundancy of intervals in the sequence |
| | Oscillator-based model (double) | Oscillator-based model + a second internal oscillator entrains when interval is asynchronous by a magnitude that exceeds the exclusion criterion |

alter temporal duration judgment performance. In the *original memory model*, the duration of the induction sequence intervals (and the T interval) is averaged¹ (see Table 1). The stored average, termed the *standard*, is then compared with the duration of the C interval. Thus, if the average duration of the induction sequence intervals is different from the T interval duration, the standard will provide incorrect information for the comparison of the T and the C intervals. Because the standard incorporates information from all of the intervals preceding C, it might be reasonable to assume that, if all of the contextual induction sequence intervals are equal in duration to T, temporal duration judgment performance should be more accurate than the (no-context) comparison of *only* the T and the C intervals. Descriptions of the original memory model (e.g., Keele et al., 1989; Schulze, 1978), however, did not provide an explicit explanation for how or why an induction sequence would cause an improvement in performance. The *multiple-look model* (Drake & Botte, 1993), a version of the memory model, estimates both the mean and the variability of the intervals, and this accounts for improved performance on the basis of the processing variability hypothesis² (see Table 1). Specifically, if, as we conjecture, there is considerable variability in the encoding of any stimulus interval (e.g., T), use of the average of encoded (physically identical but perceptually variable) intervals on each trial, rather than the encoding of the single T interval, will reduce the variability (a statistical advantage) across trials of the standard, one of the two comparison components. This would produce a steeper psychometric function, which would result in a lower threshold relative to the basic relative temporal judgment condition summarized in Figure 1A.

Beat-based mechanism. In beat-based models, an *internal beat* is established with a period based on the interonset time (marking tone onset to marking tone

onset) of intervals in the *external sequence* (e.g., Keele et al., 1989; Povel & Essens, 1985; Schulze, 1978; see Table 1). The internal beat is a form of stored encoded information. In principle, instead of a comparison of direct estimates of (T and C) interval durations, a temporal duration judgment can be accomplished by the detection of synchrony versus asynchrony (or magnitude of asynchrony) between the internal beat and the onset of the C interval ending marker tone.³ In the most literal interpretation of the beat-based mechanism, the *rigid beat-based model*, the internal beat is defined only by the duration of the first presented interval, with the internal beat being essentially a rehearsal of that first interval's duration (e.g., Keele et al., 1989; Schulze, 1978; see Table 1). For this model, if the duration between the first induction sequence interval and the T interval is not an integer multiple of the duration of that first induction sequence interval (i.e., first induction sequence interval duration * integer), the internal beat will not be in synchrony with the onset marker of the C interval, and the synchrony/asynchrony distinction will be irrelevant for comparing the T and the C intervals.

For this rigid beat-based model, the assumption that the internal beat is based on the first interval means that none of the induction sequence intervals presented after the first induction sequence interval should alter the internal temporal representation. Thus, when all of the induction sequence intervals are equal to the T interval, the rigid beat-based model should suffer from the same problem of encoding variability as a single interval that is faced by the comparison of only the T and the C intervals, but with the added handicap that the compared intervals (Intervals I and C in Figure 1C) are separated in time by the added intervening intervals. As a result, other than judgment's possibly being improved because it is based on a synchrony versus asynchrony distinction, it is unclear

how or why this mechanism's performance would ever be improved by an induction sequence.

Oscillator-based mechanism. For oscillator-based models, there are conjectured to be a number of *internal oscillators* with different periods (Large & Jones, 1999; McAuley & Jones, 2003). When the onsets of the external auditory interval markers occur at approximately equal phases of adjacent cycles of an internal oscillator, that internal oscillator is selected, or entrained, and becomes an internal representation of the external sequence. The entrained internal oscillator thus specifies both how interval duration is measured and how it is retained.

The internal oscillator sets up an expectation about when the next marker tone should occur, with this expectation manifested as an *attentional pulse*. The attentional pulse defines a range in temporal space in which the next external marking tone should occur, with the center of the attentional pulse equal to the zero phase of the internal oscillator. A judgment of C interval equality/inequality, relative to the preceding interval(s), is based on whether or not the ending C interval marker falls within the span of the attentional pulse. A judgment about the directionality of an unequal C interval is determined by the relative phase of the internal oscillator when the ending C marker tone occurs (see Table 1). The width of the attentional pulse narrows after a number of induction sequence intervals with the same duration have been presented, thus focusing attention on a small range of temporal space in which the marking tone is expected (Large & Jones, 1999). Therefore, duration judgment accuracy should be better (threshold lower) when there is an induction sequence and all of the intervals in the sequence are equal to the T interval, as compared with when only the T and C intervals are presented (see Figure 1A).

When intervals in the induction sequence differ in duration, oscillator-based models can exhibit full or partial phase and period corrections to maintain synchrony between the internal oscillator and the external sequence (e.g., Large & Jones, 1999; McAuley & Jones, 2003). Full phase correction resets the zero phase (center of the attentional pulse) to align with the onset marker tones. Period correction adjusts the timing of the internal oscillator so that the onsets of the marker tones occur at adjacent cycles. Each phase and period correction causes the temporal precision of the expectancy for the next pulse but also widens the attentional pulse (Large & Jones, 1999). Full period correction should always equate the period of the entrained oscillator to the most recently presented external interval. With partial period correction, the period of the internal oscillator at the end of a sequence (e.g., when the C interval is presented) should be equivalent to a short-term average of recently presented intervals (e.g., 4, 5, and T in Figures 1C and 1E),⁴ resulting in a judgment that is made relative to an adaptation level standard (Helson, 1947). The phase and partial period correction properties of oscillator-based models should typically enhance the validity (and thus the accuracy) of the internal oscillator for judgments of syn-

chrony/asynchrony, but the phase and period corrections also involve increases in the width of the attentional pulse, thus reducing relative temporal duration judgment accuracy. The oscillator-based model we tested has all of the described properties and capabilities (see Table 1).

Phase and Period Correction

Phase and period correction are extra properties or capabilities that increase the predictive power but reduce the parsimony of oscillator-based models. One legitimate research strategy is to compare only alternative models that have been proposed, independently of the nature and number of added properties. An alternative, and probably more productive, research strategy is to focus on comparing the basic mechanisms, considering models of each type of mechanism operating with equivalent additional capabilities. The advantage of the latter strategy lies in the ability to evaluate both the basic underlying conceptual mechanism (interval, beat, or oscillator) and the importance of the capabilities that have been posited for specific models. In the models we have described, phase and period may be parameters of the oscillator-based mechanism, but the basic functionality of allowing adjustments to these parameters is not limited to internal oscillators, and the alternative mechanisms should not be dismissed because of the added capabilities posited for models based on another mechanism.

The cyclic nature of a beat allows the use of concepts equivalent to period (e.g., interbeat interval) and phase (offset from beat defined as proportion of interbeat interval). Thus, the capabilities of adjusting phase and period are relevant to the beat-based mechanism, and adding these capabilities to models based on this mechanism would result in a model whose added capabilities make it comparable to the versatile oscillator-based models. Phase correction would serve to line up the internal beat and the marker onsets of the external sequence, making synchrony versus asynchrony judgments valid under a broader range of conditions. A beat-based/*phase-correction* model has all of the properties of the rigid beat-based model, plus the capability of full phase correction (see Table 1). Partial period correction would adjust the rate of the internal beat on the basis of the external sequence and could transform the beat-based mechanism from one for which the only possible advantage is a synchrony/asynchrony judgment task (essentially a *straw man*) to a model with the added advantage of a type of short-term statistical averaging. The beat-based/*adaptive* model has the properties of the rigid beat-based model with full phase correction (i.e., the same as the beat-based/*phase-correction* model), but with the added property of partial period correction (see Table 1); this beat-based/*adaptive* model could be considered a special case of an oscillator-based model. For the beat-based/*adaptive* model, a judgment of the C interval would be based on a synchrony/asynchrony comparison with an internal beat that represents a statistical average, with the corresponding reduction of encoding and storage variability across trials. The beat-based/

adaptive model, therefore, has the functional (statistical) advantage of Drake and Botte's (1993) multiple-look version of the interval-based mechanism, but with the variant that judging temporal duration can be based on synchrony versus asynchrony, instead of a direct duration comparison of encoded interval information. Because of the similar functionality between these two models, some predictions for performance that are based on the beat-based/adaptive model are similar to those for the multiple-look model; threshold should be lower when all of the induction sequence intervals are equal to the T interval, as compared with when only the T and C intervals are presented. The addition of phase and period correction parameters, however, make the beat-based/adaptive model less parsimonious than the existing rigid version.

An interval-based mechanism lacks the cyclic regularity that, for the other mechanisms, defines period and phase. The interval-based timing mechanism, however, restarts at each onset marker, and this is somewhat analogous to phase being reset to zero (e.g., McAuley & Jones, 2003). For models based on the other mechanisms, partial period correction results in an averaged timing parameter against which to judge the duration of the C interval, and this is functionally equivalent to the role of the standard in the interval-based memory models. Thus, the functional role of partial period correction and resetting the phase are already present in memory models.

Methodological Issues

The various timing mechanisms have been tested using a wide variety of paradigms with different tasks, and conclusions have not always been consistent across studies. Because different paradigms and different tasks often require very different types of processing, one must be careful when trying to draw solid conclusions from results across different paradigms and tasks (e.g., Woodworth & Schlosberg, 1959). In the present study, a very simple unidirectional *same-different* task was used; when the duration of the C interval was different from the duration of the T interval, the C interval was always longer. This task requires only the detection or recognition of the presence of a difference in duration. Conceptually, the participant sets a single criterion that separates equality and inequality decisions.⁵ This task does differ from other typical tasks in terms of what is required for participants to perform the judgment task.

For tasks in which the C interval is (1) greater than, (2) equal to, or (3) less than the T interval (e.g., Barnes & Jones, 2000; Large & Jones, 1999; McAuley & Jones, 2003), the participant has three decision alternatives that define three responses. This task requires the participant to set two criteria: One criterion contrasts *greater than* with *equal to* plus *less than*; the other criterion contrasts *less than* with *equal to* plus *greater than*. One must be careful in both analyzing and interpreting the results from this task. First, whether simple accuracy (e.g., Woodworth & Schlosberg, 1959) or signal detection theory

(e.g., Pastore, Crawley, Berens, & Skelly, 2003) measures are used, the data must be partitioned to reflect the criterion contrasts we have indicated. Second, not only is it rare to find an equal difference limen for less than and greater than (the point of subjective equality seldom equals the point of objective equality; see, e.g., Woodworth & Schlosberg, 1959), but also it is possible for any bias to shift the two criteria in the same, rather than the opposite, direction. Finally, relative to the requirements of our task, there is a conceptual difference in the nature of the judgment being made; the processing requirements for detecting the presence of a difference may not be the same as those required for identifying the nature of the difference. For example, there is a 10-dB difference in the threshold for detecting an onset asynchrony of otherwise simultaneous auditory stimuli and for identifying which of the stimuli had the earlier onset (e.g., Hirsh, 1959), and a 10- to 12-dB difference in threshold for detecting speech in noise and recognizing what was said (Licklider, 1951; see Allan, 1975, Miller, Wier, Pastore, Kelly, & Dooling, 1976, and Pisoni, 1980, for other temporal judgment examples).

The basic premise of the task we have just discussed is that there are three perceptual states (less than, equal to, and greater than). Another common task is judgment of directionality (shorter than vs. longer than). With two response choices, only a single criterion is set, somehow bisecting the three perceptual categories. The criterion for this task represents subjective equality plus bias and does *not* directly correspond to a threshold (e.g., Woodworth & Schlosberg, 1959). This analysis applies to judgments comparing both pairs of intervals (e.g., Barnes & Jones, 2000; Pashler, 2001) and pairs of sequences (e.g., Drake & Botte, 1993; Ivry & Hazeltine, 1995; Keele et al., 1989, Experiment 2; McAuley & Kidd, 1998). Our claim is not that one paradigm is inherently better than another but, rather, that conclusions about timing mechanisms and models need to take into consideration the nature of the task employed. In the three experiments in the present study, we used a common task, with at least one set of common judgment stimuli, which enabled us to explicitly examine how induction sequence intervals and pause duration affect relative temporal duration judgment for this paradigm and to evaluate the timing models, using these results.

EXPERIMENT 1

Studies of temporal judgment have differed in the temporal proximity of the intervals being evaluated. In addressing the importance of temporal proximity, we focus only on the two judgment intervals in the basic task used throughout the present study. In an *adjacent-interval* configuration (Figure 1A), the T and the C intervals are directly adjacent to each other and are defined by three equal-duration auditory stimuli: a beginning, a separating, and an ending marking tone. In the *separated-interval* configuration (Figure 1B), the T and the C intervals are

separated by an ISI; because the T and the C intervals are not directly adjacent to each other, four marking tones are required: a beginning and ending marking tone for each interval. In contrast to there being only a single possible version of the adjacent-interval condition, there is a range of separated-interval conditions that are distinguished by ISI duration. One goal of Experiment 1 was to provide an adjacent-interval threshold to use as a *no induction sequence baseline* for the comparison of adjacent-judgment intervals when no induction sequence was presented, thus allowing an evaluation of the effects of ISI presence and duration (Experiment 1) and of the various induction sequence properties in Experiments 2 and 3.

Historically, the separated-interval configuration has been used to investigate relative temporal duration judgments of two intervals (e.g., Abel, 1971; Creelman, 1962; Divenyi & Danner, 1977; Divenyi & Sachs, 1978; Grondin, 1993; Grondin, Ivry, Franz, Perreault, & Metthé, 1996; Henry, 1948; Rammsayer, 1994; Rammsayer & Leutner, 1996; Small & Campbell, 1962; Westheimer, 1999). Many researchers testing alternative timing mechanisms with contextual sequences used configurations with an induction sequence preceding separated judgment intervals (illustrated in Figure 1D; see, e.g., Barnes & Jones, 2000; Ivry & Hazeltine, 1995; Large & Jones, 1999; McAuley & Jones, 2003; Pashler, 2001) or with an ISI between two compared sequences (e.g., Keele et al., 1989; McAuley & Kidd, 1998). When memory models were tested, the purpose of the ISI often was to alter the duration of the standard (e.g., McAuley & Jones, 2003; Pashler, 2001), but the ISI may add a memory load to the task requirements. When beat-based and oscillator-based mechanisms were tested, the purpose of the ISI often, but not always, was to potentially disrupt synchrony between the internal representation (i.e., the internal beat for beat-based models and the internal oscillator for oscillator-based models) and the external sequence (e.g., McAuley & Jones, 2003; McAuley & Kidd, 1998; Pashler, 2001). For example, in their Experiment 1, McAuley and Jones used the stimulus configuration in Figure 1D, disrupting synchrony by manipulating the ending marker of the T interval and the beginning marker of the C interval (for simplicity, we use our terminology to describe their stimuli).⁶ McAuley and Jones's results indicated that manipulations of the temporal position of the ending marker of the T interval and, thus, ISI duration affected duration judgment accuracy. In contrast, Pashler also used the stimulus configuration in Figure 1D and manipulated the duration of the ISI, but concluded that the duration of the ISI did not affect duration judgment accuracy. Ivry and Hazeltine (Experiment 4) also investigated the effects of the presence (Figures 1C vs. 1D) and duration of an ISI between the judgment intervals, concluding that the presence of an ISI did not affect duration judgment accuracy. We suspect that these different patterns of results may reflect differences in the processing requirements of the tasks employed.

In light of these somewhat conflicting conclusions, the second goal of Experiment 1 was to compare performance on adjacent and separated intervals (Figures 1A and 1B) to test whether the presence of an ISI between the T and the C intervals for the separated intervals affects relative temporal duration judgment accuracy and, if so, whether ISI duration is important. Because cyclic (i.e., beat-based and oscillator-based) mechanisms measure the marking tone onset to marking tone onset duration, the relevant measure of ISI for these mechanisms is *intervening interonset interval* (IOI). The distinction between an ISI and an intervening IOI is that the ISI value refers *only* to the duration between the two marking tones, but the intervening IOI value refers to the time from the onset of the beginning marking tone to the onset of the ending marking tone—thus, the ISI plus the duration of the beginning marking tone. The interval-based mechanism can measure either the ISI or the intervening IOI. As will soon be evident, there are similarities and differences in how the timing models should process adjacent and separated intervals.

For the most literal interpretation of the adjacency model, only the two most recent intervals are compared; thus, in the separated-interval condition, the intervening IOI and the C interval are compared. Except when the intervening IOI is equal in duration to the T interval, the C interval will be compared with an interval with the wrong duration, and the threshold should be higher than that in the adjacent-interval condition. Because the actual T interval is irrelevant to the separated-interval task for the adjacency model, when the intervening IOI is equal to T, the threshold should be equal to, but not better than, the threshold for the adjacent-interval condition comparison of the T and the C intervals.

In memory models (original and multiple look), the duration of the C interval is compared with the average of the T interval and the intervening IOI. The average standard should be equal to T only when the T interval and the intervening IOI have identical durations (Figure 1B only) and should differ from T for all other nonzero IOI durations. Because there is no mechanism for a reduction of encoding and storage variability in the original memory model, the threshold should be identical for the adjacent-interval condition and the separated-interval condition with the intervening IOI equal to T (Figures 1A and 1B). In contrast, because the average should have a small statistical advantage (slightly reduced variability across trials) for the multiple-look model, the threshold for the separated-interval condition with the intervening IOI equal to T should be lower than (improved performance) the threshold for the adjacent-interval condition. For all other values of the intervening IOI, the threshold for both the original memory and the multiple-look models should be higher than that for adjacent intervals; this is because the average standard never will be equal to the duration of the T interval.

The rigid beat-based model does not correct phase, and thus, synchrony between the internal beat and the ex-

ternal sequence will be maintained only when the duration of the intervening IOI is equal to the duration of the T interval or, possibly, equal to an integer multiple of the T interval. Therefore, the threshold for the rigid beat-based model should be equal to that in the adjacent-interval condition when the intervening IOI is equal to an integer multiple of T and higher for all other values of the intervening IOI. Because a rigid internal beat is established on the basis of the first interval (T for separated-interval stimuli; see Figure 1B) for the beat-based/phase-correction model, the duration of the intervening IOI should not be important, since phase correction should align the internal beat with the onset marker tone of the C interval. Therefore, for the beat-based/phase-correction model, the threshold for all the conditions should be similar to that in the adjacent-interval condition.

Because the beat-based/adaptive and oscillator-based models can correct phase and period of the internal representation (i.e., internal beat for beat-based models and internal oscillator for the oscillator-based model), synchrony between the internal representation and the external sequence can be maintained for all the intervening IOI values. Thus, when the intervening IOI is *not* equal to an integer multiple of T, phase and period should correct, but with a loss of precision in the internal representation (i.e., a more variable internal beat for beat-based models and a widening of the attentional pulse for oscillator-based models), resulting in a threshold that is equal to or higher than that in the adjacent-interval condition. When the intervening IOI is equal to an integer multiple of T, there might be a slight statistical advantage for the beat-based/adaptive model and a narrowing of the attentional pulse for the oscillator-based model, resulting in a threshold that is at least as good, if not better, than the threshold in the adjacent-interval condition (Figure 1A). Thus, the presence and duration of an intervening IOI should affect relative temporal duration judgment accuracy for all of the proposed timing models, and the pattern of results will enable some conclusions about consistency with the timing model(s).

Method

Participants. The participants in all three of the present experiments were from the psychology department human participants pool at Binghamton University, with participation compensated by partial fulfillment of a requirement for an undergraduate psychology course. The participants were required to have self-reported normal hearing. Either 1 participant was run alone, or 2 participants were run simultaneously in a commercial sound chamber. In each condition of every experiment, there were two equal-sized groups of participants, a musically trained group and a musically untrained group. The criterion for inclusion in the musically trained group was at least 4 years of formal musical training, with all the other participants placed in the musically untrained group. In Experiment 1, there were 16 participants (8 in each musical training group) in each of seven conditions. Experiment 1 took approximately 30 min per participant.

Stimuli. Marker stimuli (dark bars in Figures 1A and 1B) that defined interval onsets were 50-msec segments of 1000-Hz tones, which started and ended at a waveform zero-crossing. The stimuli

were created using 12-bit pulse code modulation (PCM) synthesis at a 10k sample rate. The stimuli were low-pass-filtered at 2 kHz, then presented binaurally through Sennheiser Model 450II headphones at 66 dB (SPL). The silent intervals were created as part of the PCM code and, thus, were timed by the digital-to-analog converter, resulting in an accuracy of 0.1 msec.

In Experiment 1, there were two basic stimulus configurations—the adjacent intervals (Figure 1A) and the separated intervals (Figure 1B)—with each configuration having two different values of T: 450 and 650 msec. These two values of T, coupled with different values of the intervening IOI, created five separated-interval conditions. The three separated-interval conditions with a T interval duration of 450 msec had intervening IOI values that were shorter than (350 msec), equal to (450 msec), or an integer multiple of (900 msec) the T interval. The two separated-interval conditions with a T interval duration of 650 msec had an intervening IOI that was either shorter than (350 msec) or equal to (650 msec) the T interval. The two adjacent- and five separated-interval conditions defined the seven between-subjects conditions.

The possible durations of the C interval were created using the following equation:

$$C = T + (n * \Delta T); n = 0 \text{ to } 8 \quad (\Delta T = 15 \text{ msec}). \quad (1)$$

In all the experiments, 13 trials defined a single trial set: 5 trials with $n = 0$ (C duration = T), and 1 each for $n = 1$ to 8 (eight durations of C that are longer than T). Thus, if trials with a low value of n are at or below threshold, the C interval will be perceived as being equal in duration to T on approximately half of the trials, and the C interval will be perceived as longer in duration than T on approximately half of the trials.

Procedure. In all three experiments, the participants performed a unidirectional fixed *same-different* task, with the task being to judge whether the C interval was equal to or longer than the duration of the T interval. Each trial had a 1,500-msec response interval. Responses were made by pressing one button (left) to indicate that the stimuli had the same duration or a different button (right) to indicate that stimuli had a different (longer) duration. Feedback indicating the correct answer was given at the end of each trial, with the words *same* or *different* appearing above the correct response button. In each condition of Experiment 1 and, thus, for each participant, the value of the T interval and, for the separated-interval conditions, the intervening IOI was constant.

In all the experiments, prior to performing the task, the participants were given both a verbal explanation of the task and a visual diagram of the stimuli that they would be judging (in Experiment 1, the diagram was similar to Figure 1A or 1B). They were informed that if the durations of the T and the C intervals were different, the C interval would always be longer, and never shorter, than the T interval. Each participant began with one practice repetition of a single 13-trial set. The goals of this simple practice set were to assist the participants in understanding the task, to give them some experience with the stimuli, and to evaluate task understanding before beginning data collection. If either self-report or responding during the practice indicated a failure by the participant to understand the task, the task was explained again, and a second set of practice trials was given. An a priori protocol was to discard the data from any participant still unable to perform the task, but this never occurred in any of the experiments. In Experiment 1, there were 18 repetitions of the 13-trial set, with all 234 trials randomly intermixed by the computer for each session. To combat participant fatigue, two breaks were given in all the experimental sessions.

Results and Discussion

For all the experiments, only mean (i.e., not trial-by-trial) data for each C value were collected. In all the experiments, probabilities of hits (correctly identified differ-

ences) and false alarms (incorrectly reported differences) were converted to d' values. Linear regression of the d' psychometric function (between asymptotic levels of performance) was used to estimate the ΔT threshold [amount of difference for $d' = 1$, or unbiased $P(c) = .69$] for each participant.⁷ Testing for threshold outlier values was done using an a priori criterion of two standard deviations beyond the mean. Outlier values were discarded and replaced by data from new participants so that there were an equal number of participants in each condition. In Experiment 1, seven threshold values were classified as outliers and were replaced by new data, with no more than two data points removed from any condition. Analyses for Experiment 1 were conducted on Weber fractions ($\Delta T/T$), thus correcting for differences in threshold across values of T that merely reflected Weber's law. In Figure 2, mean Weber fractions are plotted for the seven conditions, allowing comparison of sensitivity across values of T. The left and right ordinates are scaled in terms of the actual threshold values in milliseconds for the 450-msec target and the 650-msec target stimuli, respectively. Threshold means separated by musical training group are provided in the Appendix (Table A1).

Unless otherwise indicated, significance is based on a difference that exceeds two standard errors. A determination of significance thus can be made using the standard error bars shown in all the figures. To investigate the question of whether there is an advantage (i.e., a sta-

tistical advantage for the multiple look and the beat-based/adaptive models or a narrowing of the attentional pulse for the oscillator-based model) when an intervening IOI is equal to the T interval and, thus, whether the intervening IOI contributes to the internal representation for the timing models (i.e., the standard for memory models, the internal beat for beat-based models, and the internal oscillator for oscillator-based models), we separately compared Weber fractions for the adjacent-interval conditions with those for the separated-interval conditions in which the intervening IOI was equal to T. Specifically, in a 2 (musical training) \times 2 (T duration) \times 2 (stimulus configuration) analysis of variance (ANOVA), the adjacent-interval conditions were compared with the separated-interval condition in which the intervening IOI was equal to T (a 450-msec intervening IOI for the 450-msec T and a 650-msec intervening IOI for the 650-msec T). The only significant factor was stimulus configuration [$F(1,56) = 15.67, p < .0005$], demonstrating that Weber fractions, or relative thresholds, were significantly smaller (performance was better) for the separated intervals (see Figure 2). This improvement in performance can be interpreted as indicating that the intervening IOI does contribute to the internal representation. The improved performance is consistent with models that can take advantage of redundancy with T—thus, the multiple-look, beat-based/adaptive, and oscillator-based models. The improved performance is not consistent with the adja-

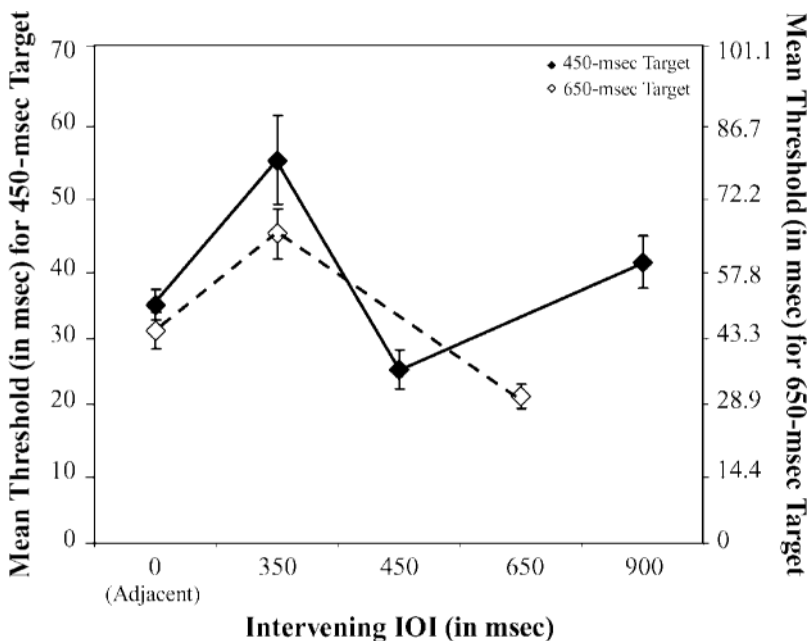


Figure 2. Weber fraction threshold means and standard errors for the adjacent- and the separated-interval conditions in Experiment 1 for T = 450 msec (black squares) and T = 650 msec (white squares). The left axis shows the threshold in milliseconds for T = 450 msec conditions, and the right axis shows the threshold in milliseconds for T = 650 msec conditions. IOI, interonset interval.

gency model, because, by limiting comparison to the immediately preceding interval, the adjacency model has no means for an improvement in performance for separated intervals when the intervening IOI is equal to T. These results also provide evidence against strict interpretations of the original memory, the rigid beat-based, and the beat-based/phase-correction models, because, in these models, there is no reason why the relative temporal duration judgment should have been assisted by redundant temporal information (see Table 1).

In the second 2 (musical training) $\times 2$ (T duration) $\times 2$ (stimulus configuration) ANOVA, the adjacent-interval conditions were compared with the separated-interval conditions, when the intervening IOI was *not* equal to T (intervening IOI = 350 msec for 450- and 650-msec values of T, and intervening IOI = 900 msec for T = 450 msec). Again, the only significant factor was stimulus configuration [$F(1,72) = 12.36, p < .001$]. The pattern of results indicated that Weber fractions (proportional thresholds) for adjacent intervals were smaller (performance was better) than the thresholds in the separated-interval conditions with intervening IOI durations different from the T interval (see Figure 2). These results rule out the possibility that performance is improved when any intervening IOI is presented (e.g., when time is allowed for the encoding or consolidation of interval information).

Musical training. Musical training was not significant in any of the analyses of the full set of conditions. Because, in our subsequent experiments, we used the 450-msec adjacent-interval condition thresholds as a *no induction sequence baseline*, however, we more closely examined performance for this one condition and found that the musical training group performed significantly better than the group without musical training [$t(14) = 3.05, p < .01$]. Whereas the significant musical training results for the 450-msec adjacent-interval condition could be spurious, these results may be an indication that musically trained listeners can establish an internal representation based on one interval more quickly and more accurately than musically untrained listeners can. This difference in musical training groups, however, disappears when the C interval is preceded by two consecutive equal-duration intervals (intervening IOI equal to T and T; see Table A1). We will return to this musical training issue after we consider the results for the remaining experiments, in which the effects of initial induction sequences were evaluated.

EXPERIMENT 2

The various timing models differ, to varying degrees, in whether or not, and how, a preceding induction sequence will alter temporal judgment ability. Prior research, in which a number of different paradigms have been used, has provided mixed results about the influence of contextual timing information on temporal judgment ability (e.g., Barnes & Jones, 2000; Drake & Botte, 1993; Ivry & Hazeltine, 1995; Keele et al., 1989; Jones

& Yee, 1997; Large & Jones, 1999; McAuley & Jones, 2003; McAuley & Kidd, 1998; Pashler, 2001; Rousseau & Rousseau, 1996; Schulze, 1978; Sorkin, 1990). To evaluate the influence of an induction sequence, one needs to compare performance (e.g., threshold or accuracy) in a basic task both in isolation and in the presence of an induction sequence (e.g., Figure 1C vs. 1A). With few exceptions (e.g., the present study; McAuley & Jones, 2003), induction sequence studies have not provided this comparison. Experiment 1 provided a basic understanding of the ability to process different configurations of (nominally) single pairs of intervals and established baseline estimates of the ability (threshold) to perform the ordered *same-different* task for a single pair of judgment intervals (Figure 1A). In Experiment 2, we examined the influence of various five-interval isochronous induction sequences on the threshold for performing this same basic task, allowing us to examine several important questions. Can the presentation of additional intervals result in improved relative temporal duration judgment accuracy? If so, does the number of intervals in the contextual sequence affect relative temporal duration judgment accuracy? Finally, if the context provided by an isochronous induction sequence of intervals equal to T results in improved performance (lower threshold), what will be the effect of an isochronous induction sequence whose intervals differ significantly from T?

The separated-interval condition in Experiment 1 with an intervening IOI equal to T is functionally equivalent to an adjacent-interval condition (T and C intervals) with the addition of a preceding single induction interval; because they are equal in duration, we can conceptually exchange the labels T and intervening IOI in Figure 1B. Thus, Experiment 1 also provided a measure of the influence of a single induction interval (equal to T) on ability (threshold) to perform the same task. The thresholds for these separated-interval conditions (Figure 2 for intervening IOI equal to T with T = 450 and 650 msec) were significantly lower than the threshold for performing the adjacent-interval task. Thus, the addition of a single induction sequence interval does significantly improve performance. This answers our first question from above in the affirmative and leads to our second question about number of induction sequence intervals.

If there is a statistical advantage for initial redundant interval information, variability should decrease as a function of the number of intervals. In studies evaluating the judgment of two intervals preceded by a contextual sequence, Ivry and Hazeltine (1995), McAuley and Jones (2003), and Madison (2004) reported improved temporal duration judgment accuracy with a greater number of contextual intervals, whereas Pashler (2001) reported no difference for one versus five contextual intervals. In two-sequence paradigms, there is some evidence that an increase in the number of intervals can improve temporal duration judgment accuracy (e.g., Drake & Botte, 1993; McAuley & Kidd, 1998; Schulze, 1978). We have already

expressed the concern that, with the differences across studies in task, procedure, and stimulus configuration (including presence, location, and duration of the ISI), it is difficult to draw simple general conclusions about the importance of the amount of redundant temporal information.

Method

Participants. Participant information is the same as that in Experiment 1, except that 20 participants (10 musically trained and 10 musically untrained) were used in each of the three conditions. The experiment took approximately 1 h for each participant.

Stimuli. Stimulus information is the same as that in Experiment 1, except that the T = 450 msec adjacent judgment interval stimuli were used for every condition, and a five-interval isochronous (equal-interval) induction sequence directly preceded the judgment intervals on each trial (see Figure 1C). The three conditions were classified by the duration of the induction intervals, in which the IOI (beginning of onset marking tone to beginning of offset marking tone) was shorter than (250 msec), equal to (450 msec), or longer than (650 msec) the T interval. The induction sequence intervals were defined by five 50-msec 1500-Hz tones (thus, different from the 1000-Hz judgment interval markers). Each participant received only one duration of induction sequence intervals, which, therefore, was a between-subjects variable.

Procedure. As in Experiment 1, each participant began with one practice repetition of the adjacent-interval 13-trial set with no induction sequence. The participants in Experiment 2 also received a second version of practice with the full stimuli (i.e., the induction sequence, consisting of the five induction sequence intervals with the duration assigned to that condition, immediately preceding the T and C judgment intervals). Task, response, and feedback were the same as those in Experiment 1. The data for each participant were

based on 20 repetitions of the 13-trial set, with the 260 trials randomly intermixed by the computer for each session.

Results and Discussion

All analyses were conducted on the ΔT threshold data. In Experiment 2, two threshold values, each from a different condition, were classified as outliers and were replaced by threshold data from new participants. The threshold data were evaluated using a 2 (musical training) × 3 (induction sequence interval duration) ANOVA. The main effect of musical training was not significant [$F(1,54) = 0.46, p > .5$]. Figure 3 summarizes the threshold results for the Experiment 2 conditions, collapsed across musical training groups and the no induction sequence (from Experiment 1). Threshold means separated by musical training group are provided in the Appendix (Table A2). The effect of induction sequence interval duration on the listeners' temporal duration judgment threshold, or sensitivity, was significant [$F(2,54) = 12.53, p < .0005$], with threshold lowest (performance best) for the 450-msec induction sequence intervals.

Performance for the 250- and 650-msec induction sequence conditions was significantly worse than performance in the 450-msec *identical induction sequence* condition (see Figure 3). For the 250-msec induction sequence condition, threshold was similar to the no induction sequence baseline, which, taken in isolation, could be an indication that participants can ignore the induction sequence when it does not provide useful temporal

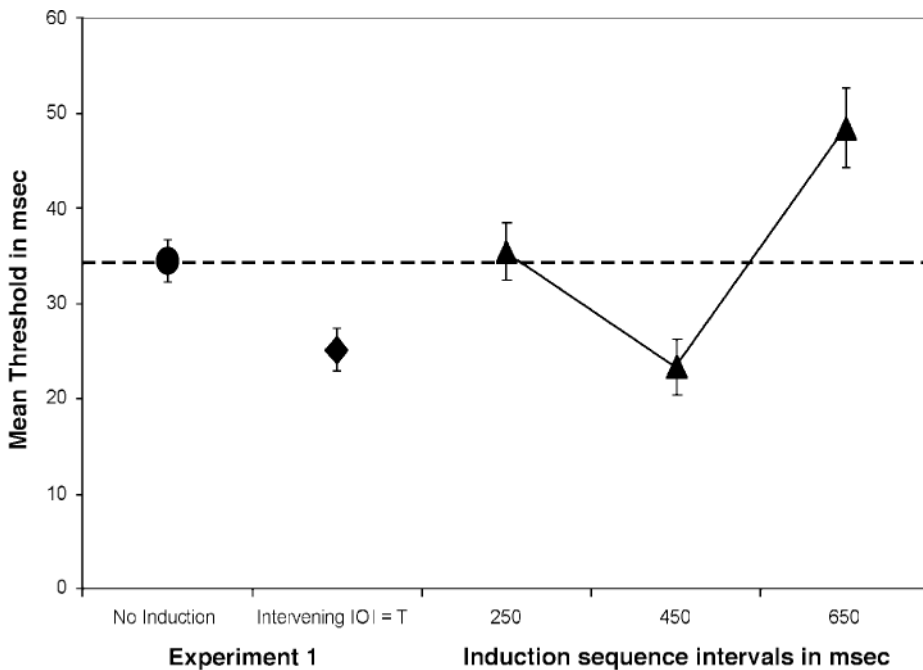


Figure 3. Threshold means in milliseconds and standard errors for the conditions in Experiment 2. The dashed line is the average no-induction sequence threshold (baseline) from Experiment 1. IOI, interonset interval; T, target.

information. The threshold for the 650-msec induction sequence condition, however, was significantly worse than that for the no induction sequence baseline, indicating that, at times, a mismatched induction sequence can adversely affect temporal duration judgment capabilities. The last results provide evidence that listeners sometimes cannot ignore (irrelevant) temporal information provided by the contextual induction sequence, which, in turn, is consistent only with temporal judgment models that account for contextual influences on the basis of the duration of induction sequence intervals (i.e., the multiple-look model, the beat-based/adaptive model, and the oscillator-based model). Although this impairment in performance for this longer (slower) induction sequence condition may be theoretically important, a further evaluation of this finding is beyond the scope of the present study.

The critical condition in Experiment 2 was when the induction sequence provided relevant redundant temporal information. The thresholds for this identical (450-msec) induction sequence condition were significantly lower than the threshold for the no induction sequence baseline [the 450-msec adjacent-interval condition from Experiment 1; $t(34) = 2.89, p < .01$; see Figure 3], which again is consistent only with temporal judgment models that account for contextual influences on the basis of the duration of induction sequence intervals. Specifically, for the multiple-look model and the beat-based/adaptive model, the largest number of same-duration intervals should produce the most statistically accurate standard and internal beat, respectively. For the oscillator-based model, the larger the number of same-duration intervals, the narrower the attentional pulse. There should be no context effects for all the other timing models (see Table 1). Because the results from Experiment 2 indicated a significant improvement when the judgment intervals were preceded by an induction sequence with intervals equal to the T interval, the results are not consistent with these four other models.

Number of induction sequence intervals. To test whether performance is different for one-interval and five-interval induction sequences, we compared thresholds for the identical induction sequence condition in Experiment 2 and the 450-msec intervening IOI with those for the 450-msec T separated-interval condition in Experiment 1. There was no significant difference between these two conditions [$t(34) = 0.42, p > .65$]. Furthermore, the overall mean thresholds were nearly identical (less than 2-msec difference); the difference was essentially zero for the musically untrained group and was under 4 msec (less than one standard error) for the musically trained group (see the Appendix). Thus, even one additional induction sequence interval provides enough temporal information to lead to an improvement in relative temporal duration judgment accuracy (see, e.g., Madison, 2004). Overall, the results from Experiment 2 indicated that a temporal context does affect relative temporal duration judgment accuracy, with accuracy best when the induction sequence intervals are equal to the T interval; these findings are consistent with the multiple-look, the beat-based/adap-

tive, and the oscillator-based models. The improvement when an identical temporal context is presented, however, is not affected by additional contextual intervals.

EXPERIMENT 3

A number of experiments have inserted pauses within sequences of stimuli to evaluate the effects of disrupting the timing of intervals on relative temporal judgment (e.g., Ivry & Hazeltine, 1995; McAuley & Jones, 2003; Pashler, 2001). Experiment 1 demonstrated that, for our unidirectional *same-different* task, a pause (intervening IOI) between the judgment intervals does alter temporal judgment threshold, with the nature of the change dependent on the duration of the intervening IOI relative to the T interval. Recently, McAuley and Jones addressed the question of whether a pause between the induction sequence and the judgment intervals alters relative temporal duration judgment accuracy. In their Experiment 2, McAuley and Jones compared judgment accuracy for stimulus configurations with (not depicted) and without (Figure 1D) a pause between the induction sequence and the judgment intervals. Finding that accuracy was not significantly impacted, they concluded that the introduction of the pause between the induction sequence and the judgment intervals did not “weaken participants’ reliance on the context sequence” (McAuley & Jones, 2003, p. 1113). As was noted earlier, because the various studies we have discussed differed in both stimulus configuration and the task required of participants, the results do not necessarily generalize to other tasks, such as the one in the present study, or to other conditions (e.g., other pause lengths). Experiment 3 was designed to systematically evaluate whether a pause between the induction sequence and the judgment intervals affects relative temporal duration judgment and whether the magnitude of any effect is a function of pause duration.

Because the addition of a pause was expected either to have no effect or to decrease performance (increase threshold), we chose for Experiment 3 the induction sequence that exhibited the best performance in Experiment 2, the 450-msec induction sequence stimuli (sequence intervals = T), adding a pause between the induction sequence and the judgment intervals. The pause was an extension of the IOI of the last interval in the induction sequence, thus an *IOI extension* (see Figure 1E), with pause duration varying across conditions. Because this final IOI extension should affect the operation of the critical internal timing mechanisms of the multiple-look, beat-based/adaptive, and oscillator-based models in both similar and different ways, the pattern of results should allow us to evaluate both the common and the different predictions of these timing models.

Method

Participants. Participant information is the same as that in Experiment 1, with 16 participants, 8 musically trained and 8 musically untrained, in each of the 11 conditions. There thus were 176 participants, 172 from the psychology department human partici-

pants pool at Binghamton University, plus 4 graduate student volunteers.

Stimuli. The stimuli were those used in the 450-msec induction sequence condition in Experiment 2, but with the duration of the final induction sequence interval lengthened by an IOI extension, as depicted in Figure 1E. There were 11 final IOI extensions: 50, 100, 200, 225, 300, 400, 450, 500, 900, 950, and 2,250 msec. Thus, for example, when the final IOI extension was 200 msec, the last induction sequence interval had a total IOI of (450 + 200 =) 650 msec. Each participant was exposed to only one IOI extension, with each session lasting approximately 1 h.

Procedure. Each participant received two versions of practice: The first version was the no induction sequence adjacent intervals (Figure 1A), and the second version was identical to the stimuli for the participant (i.e., induction sequence with a specific final IOI extension; Figure 1E). Task, response, and feedback were identical to those in the previous experiments. The data for each participant were based on 20 repetitions of the 13-trial set, with the 260 trials again randomly intermixed.

Results and Discussion

Thresholds were calculated in the same manner as in the previous experiments, with outliers identified, removed, and replaced. As a result, the data from 7 participants were discarded and replaced. No more than two data points were removed from any single condition. The threshold data were subjected to a 2 (musical training) × 11 (final IOI extension) ANOVA. The duration of the final IOI extension significantly affected temporal duration judgment sensitivity [$F(10,154) = 4.05, p < .0001$]. Musically trained participants also had significantly lower thresholds than did musically untrained participants [$F(1,154) = 5.32, p < .03$], but there was no sig-

nificant interaction between musical training and final IOI extension duration.

Thresholds separated by musical training group are provided in Table A3 in the Appendix, which indicates that the only large differences in thresholds between musically trained and untrained listeners were for the 100-, 225-, and 500-msec IOI extension conditions. The most salient attribute about the 225-msec IOI extension condition is that the IOI extension was exactly one half of a 450-msec induction sequence interval. Several studies have shown that listeners are able to use halved intervals in an induction sequence for accurate relative temporal duration judgment (e.g., Barnes & Jones, 2000; Berens & Pastore, 2005; McAuley & Jones, 2003), and other studies have shown that listeners with musical training often are more adept at certain types of rhythmic judgments (e.g., Jones & Yee, 1997; Snyder & Krumhansl, 2001), such as halving. Given the possibility that the difference in threshold for the 225-msec IOI extension condition might be spurious, and because the 225-msec IOI extension condition was the only condition that used a halved interval, a detailed discussion of listener halving capabilities and of this finding is beyond the scope of the present study. We will delay our discussion of the difference for the 100- and the 500-msec IOI extension conditions, since it is best understood in the context of the revisions for the timing mechanisms. The thresholds for the remaining IOI extension conditions were highly similar across groups.

Figure 4 shows the 11 thresholds for the different final IOI extension conditions with the relevant baseline comparisons for the no induction sequence and for the iden-

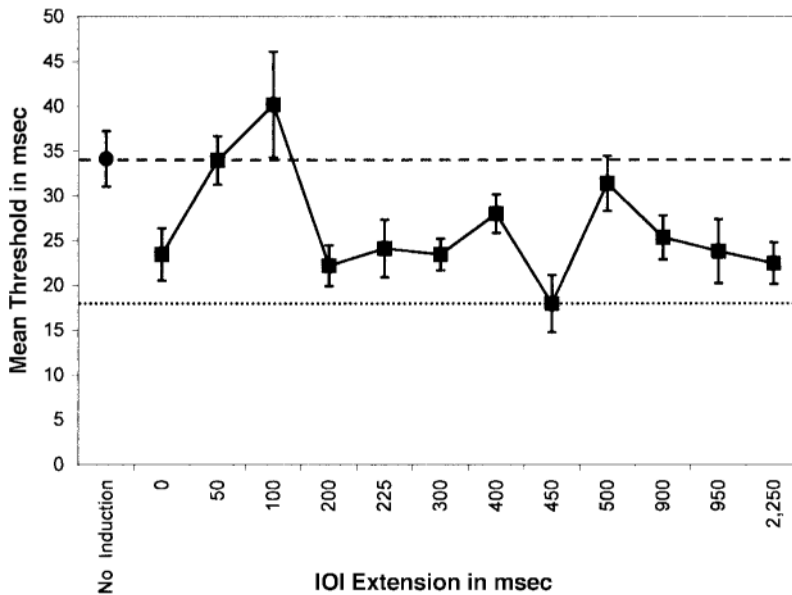


Figure 4. Threshold means in milliseconds and standard errors for the interonset interval (IOI) extension conditions in Experiment 3. The dashed line is the no-induction sequence baseline from Experiment 1, and the dotted line is the identical induction sequence baseline from Experiment 2.

tical induction sequence with no final IOI extension (taken from Experiments 1 and 2, respectively). Except for the 500-msec IOI extension condition, the thresholds for all final IOI extension values that were longer than 100 msec were similar to the identical induction sequence baseline. Because only a long pause was used in their study, these long IOI extension results are consistent with the findings of McAuley and Jones's (2003) Experiment 2. Our results, however, indicate that, depending on duration, the addition of a pause can adversely affect performance. The present pattern of thresholds will be discussed in terms of their consistency with predictions based on each of the three remaining timing models.

Multiple-look model. The multiple-look model computes an average standard of the initial intervals (e.g., the induction sequence and the T intervals), then uses this average standard to judge the relative duration of the C interval (see Table 1). If the extended final induction sequence interval is included in the average, there will be a discrepancy between the standard and the duration of the T interval, with the magnitude of this discrepancy a direct function of the duration of the final IOI extension. Thus, as final IOI extension increases, the standard will become increasingly dissimilar to the T interval (i.e., the standard duration will increase), resulting in an increasing number of errors and, thus, increases in threshold as a function of increasing final IOI extension. The actual pattern of results is almost the opposite, with the greatest interference found for the shortest, rather than the longest, IOI extensions (see Figure 4).

Cyclic models without phase and period correction. In models based on both the beat-based and the oscillator-based mechanisms, phase, or synchrony between the final marker of the C interval and the cyclic internal representation, is used to judge the relative duration of the C interval (see Table 1). The oscillator-based model can vary in the degree of phase and period correction (e.g., McAuley & Jones, 2003). If the phase and period correction parameters of the oscillator-based model are set to zero, resulting in no phase or period correction, the phase and period of the internal oscillator would be determined solely by the first induction sequence interval, and the cycle of the internal oscillator would be *rigid*. Except for the basic timing mechanism, this configuration of the oscillator-based model is functionally equivalent to the rigid beat-based model, or any other rigid cyclic model that, once started, continues in time (see Table 1). Ignoring consideration of the lack of statistical advantage, such rigid cyclic models should perform well under certain conditions of Experiment 3 and poorly under other conditions. Specifically, because of expected synchrony between the onset of marker stimuli and the rigid cyclic timing process, accurate judgment of the duration of the C interval based on synchrony/asynchrony or phase should be accurate only when the IOI extension is an integer multiple of the 450-msec T interval (i.e., 450, 900, and 2,250 msec). If, in the absence of any external marker stimuli, there is degradation of the cyclic internal timing process over longer periods of time,

threshold should increase for longer IOI extension (i.e., 2,250 msec and, possibly, 900 msec) and, thus, be worse than the identical induction sequence baseline. Although the average threshold for the 450-msec and, possibly, the 900-msec IOI extension conditions was slightly lower than any of the other thresholds in Experiment 3, the difference was not significant; essentially, the thresholds for these three integer multiple conditions were equal to each other and to the identical induction sequence baseline (0-msec IOI extension). Thus, if there is degradation of the cyclic interval process over time, the effect is minimal for the delays studied in this experiment. In isolation, the results for these integer multiple conditions are consistent with rigid cyclic models.

In contrast to the integer multiple IOI extension conditions, asynchrony, or phase, between the externally and internally defined intervals does not provide any useful information for a rigid cyclic model in the noninteger multiple of IOI extension conditions (i.e., 50, 100, 200, 225, 300, 400, 500, and 950 msec). Because any attempt to use a synchrony/asynchrony distinction to judge duration will result in an incorrect decision, threshold for these noninteger multiple IOI extension conditions should be elevated relative to the identical induction sequence baseline. It is here that the present results are not consistent with expectations for a rigid cyclic model. Specifically, except for the shortest IOI extension conditions (50 and 100 msec) and the 500-msec IOI extension condition, thresholds for all the IOI extension conditions were nearly equal to each other and were similar to the identical induction sequence baseline (see Figure 4).

Cyclic models with phase and/or period correction. The beat-based/adaptive model is characterized by full phase and partial period correction, and the oscillator-based model (McAuley & Jones, 2003) can exhibit any magnitude of phase and period corrections (see Table 1). Can phase and period correction allow cyclic models to account for the full pattern of results from Experiment 3? Full phase correction aligns the zero phase of the internal representation with the markers in the external sequence. As long as the period of the internal representation of these cyclic models is equal in duration to the T interval, this phase alignment should guarantee that synchrony/asynchrony, or phase, always will be a useful indicator of the relative duration of the C interval. There are, however, negative consequences for period correction. Specifically, partial period correction should add an error to the averaged period of the internal timing process, with the magnitude of error a monotonic function of the magnitude of the IOI extension for the beat-based/adaptive model. Phase and any (partial or full) period correction also should result in a widening of the attentional pulse for the oscillator-based model. These consequences of period correction should result in threshold that is elevated, as compared with the identical induction sequence baseline. Thus, because thresholds for most of the longer IOI extension conditions were similar to the identical induction sequence baseline, the overall pattern of results is not consistent with any of the cyclic models.

GENERAL DISCUSSION

How Does Temporal Context Affect Relative Duration Judgment?

For the task of comparing the duration of a pair of judgment intervals (i.e., T and C), the results from Experiments 1 and 2 (Figures 2 and 3, respectively) demonstrated that the difference threshold was always smaller, indicating greater sensitivity, when all the preceding intervals were the same duration as the T interval, relative to when any of the recent intervals was not equal to the T interval. These results provide a clear demonstration that the redundancy provided by a temporal context (i.e., either separated judgment intervals or the presence of an induction sequence) assists relative temporal duration judgment accuracy. This improved performance can be explained in terms of a statistical advantage in the operation of the multiple-look model, a functionally equivalent statistical advantage in the operation of the beat-based/adaptive model, and the narrowing of the attentional pulse (functionally, a statistical advantage) with redundancy in the oscillator-based model. In terms of reduction in variance, the statistical advantage for an average of six intervals (five induction sequence intervals and T interval) should result in far greater accuracy (less encoding and storage variability) than for only two intervals and, thus, a lower threshold. Therefore, the finding of equivalent thresholds with five same-duration induction sequence intervals (Experiment 2), as compared with one same-duration induction sequence interval (the IOI equal to T condition in Experiment 1), was not fully consistent with the multiple-look model, the beat-based/adaptive model, and the oscillator-based model. Some insight into the reason for this lack of difference can be gained by returning to the significant musical training difference for the 450-msec adjacent-interval condition in Experiment 1, which contrasts with the lack of such a difference in Experiment 2.

We introduced our adjacent-interval task with an analysis of processing requirements; we now will use those requirements to examine how the timing mechanisms, and the models based on these mechanisms, perform this task without, and thus with, the context provided by an initial induction sequence. For the interval-based mechanism, performance of the adjacent-interval task required a rapid and accurate estimate of the duration of the T and C intervals and then a comparison of the two encoded intervals. For the beat-based mechanism, an internal beat had to be established solely on the basis of T, allowing the synchrony/asynchrony of the ending C marker tone to be used to judge the relative duration of C. Likewise, an oscillator-based mechanism required entrainment of an internal oscillator solely to the T interval so that the synchrony/asynchrony, or phase, of the ending C marker tone could be used to judge the relative duration of C. Thus, for all of the timing mechanisms, the listeners had to establish an internal representation of duration on the basis of a single interval and then use that quickly established representation for the

comparison task. As part of the skills needed to perform music in any kind of ensemble, musicians need to be able to make rapid, accurate judgments of duration; there is no reason to believe that the typical nonmusician possesses a highly developed similar skill. This skill would account for the lower threshold exhibited by the group with musical training for the 450-msec adjacent condition in Experiment 1. When there are two equal-duration intervals preceding the C interval (i.e., the intervening IOI and the T interval duration are equal for the separated-interval condition in Experiment 1), the redundant timing information is probably sufficient for all listeners to establish a relatively precise internal representation (i.e., interval standard, internal beat, or internal oscillator), and this may be adequate to accurately perform the relative temporal judgment task; the thresholds for this last condition are roughly equivalent across groups (see the Appendix, Table 1A).

What Happens When There Is an IOI Extension of the Last Induction Sequence Interval?

The multiple-look, beat-based/adaptive, and oscillator-based models are equally consistent with the threshold results from Experiments 1 and 2. Experiment 3 provided the basis for both evaluating and beginning to differentiate among these models and the mechanisms they are based on. The critical findings from Experiment 3 are that (1) thresholds are elevated for short (50- and 100-msec) IOI extensions, (2) thresholds are lower and equal to the fully redundant isochronous induction sequence threshold for all longer (200-msec and longer) IOI extensions, with the one exception being (3) the elevated threshold for the 500-msec IOI extension. We will delay trying to address and, in fact, temporarily will ignore the final exception (3), focusing instead on the two major range effects (1 and 2).

Oscillator-based mechanism. On the basis of the finding that thresholds for the longer IOI extension conditions are similar to the identical induction sequence baseline, we can make three inferences about the functioning of the oscillator-based model for these conditions: (1) The internal oscillator must be in synchrony with the onset marker of the C interval, (2) the period of the oscillator must be a close approximation of the T interval, and (3) the attentional pulse must be considerably narrower than when no induction sequence interval precedes the judgment intervals. On the basis of these three inferences, the only reasonable conclusion is that the longer noninteger multiple IOI extensions did not affect the phase and period of the internal oscillator that was used for the relative temporal duration judgment task. How might the oscillator-based model exclude a long or deviant interval? Because there are multiple internal oscillators, each with a different period, it is conceivable that the longer IOI extensions entrained a different internal oscillator with a different period (see Large & Jones, 1999, for a detailed description of entraining multiple internal oscillators). As a result, an internal oscilla-

tor with a period approximately equal to T would be entrained by the first induction sequence interval. With the presentation of Induction Sequence Intervals 2, 3, and 4 (in Figure 1E), all equal to T , the period of that internal oscillator quickly converges to a period equal to T , and the attentional pulse narrows. When significantly longer than T , the last induction sequence interval (Interval 5 in Figure 1E) entrains a different internal oscillator, not altering anything about the initially entrained oscillator. The subsequent presentation of the T interval can only reinforce the operation of the initially entrained oscillator, and full phase correction aligns the zero phase of the first entrained oscillator to the T interval markers. The temporal judgment task then can be accomplished using the synchrony/asynchrony distinction, or phase, of the ending marker tone of the C interval with that first internal oscillator. With a narrow attentional pulse, temporal judgment accuracy should be high (threshold low). Thus, for relatively long IOI extension values (e.g., values clearly greater than 100 msec; see Figure 4), this model should function as effectively as when there is no extension of the IOI of the final induction sequence interval.

The critical assumption here is that, when the final induction sequence interval with IOI extension entrains to a second internal oscillator, the period of the initially entrained oscillator is not affected. If every interval in this external sequence interacts (e.g., entrains or results in a period correction) with the same oscillator, accuracy for this model will be reduced, because the phase and period correction caused the attentional pulse to widen and, thus, produce higher thresholds. We would expect that this will happen for relatively small values of IOI extension (e.g., for the stimuli in Experiment 3, typically, values clearly less than 200 msec; see Figure 4). Specifically, when the IOI extension is small, the final induction sequence interval with IOI extension does not entrain a different oscillator but, rather, alters the period (and phase) of the initially entrained oscillator. The distinction between interval durations that do and do not interact with a specific oscillator can be quantified by an *exclusion criterion*, whose value should be specific to each internal oscillator's dynamic range of periods and may differ somewhat across listeners. On the basis of the pattern of results from Experiment 3, an internal oscillator with a period of approximately 450 msec probably has a typical exclusion criterion that is approximately 150 msec (e.g., between 100 and 200 msec). This model accounts for all of the Experiment 3 results, except for the 500-msec IOI extension threshold.

The addition of the exclusion criterion with an approximate value of 150 msec for 450-msec induction sequence intervals allows us to address the second condition in which the listeners with musical training performed significantly better than the listeners without musical training—the 100-msec IOI extension condition. The musically trained group exhibited a significantly lower average threshold in the (no-induction) adjacent-interval condition in Experiment 1, which provided the no-induction sequence baseline for Experiments 2 and 3. The

musically trained group also exhibited a lower, but not significantly lower, threshold with the 450-msec induction sequence conditions in Experiment 2 (the 0-msec IOI extension condition plotted in Figure 4); the lack of significance may simply reflect a floor effect. If there is a tendency for differences in basic sensitivity across musical training groups, there may well be similar differences in the exclusion criterion across these groups of listeners, with musically trained individuals tending to have a lower exclusion criterion. We assume that for nearly all of the listeners, the 50-msec IOI extension should be below the exclusion criterion (resulting in elevated thresholds), and the 200-msec IOI extension should be above the exclusion criterion (resulting in low thresholds). The 100-msec IOI extension, however, is probably below the individual exclusion criterion (resulting in even more elevated thresholds) for a larger proportion of musically untrained, than musically trained, listeners. Following this logic, there should be greater total variability across listeners for the 100-msec IOI extension condition, and this can be seen in the size of the standard errors in Table A3. A significant between-groups difference, however, always should result in greater variability in the pooled data. More telling is that the distribution of thresholds across the two groups was bimodal in this condition and was unimodal in the other 10 conditions in Experiment 3. Furthermore, the size of the standard error for the thresholds of each group is larger in this 100-msec IOI extension condition than in any of the other conditions. This pattern of threshold distribution seems to be indicative of two modes of processing, such as the 100-msec IOI being below the exclusion criterion for some listeners and above the criterion for others, with a significant difference in proportion of listeners across the two musical experience groups.

With the addition of the exclusion criterion and the discussion of differences for the value of the exclusion criterion for listeners, we now can address the threshold for the 500-msec IOI extension condition. If the initially entrained oscillator continues to operate through the final induction sequence interval (through the IOI extension) even when a different oscillator is entrained by that longer interval (and there is no reason why it should not do so), there should be an expectation of an interval marker at the next zero phase of that oscillator and, thus, at a point in time equal to an IOI extension of 450 msec. This situation is equivalent to a missed beat or note; in essence, we here are again suggesting that, despite the exclusion criterion, integer multiples of the oscillator's period are important. For the 500-msec IOI extension condition, the interval marker is now asynchronous with the initially entrained oscillator external sequence by 50 msec, which is less than the exclusion criterion. This could lead to period correction, which would cause both a change in the accuracy of the internal oscillator in judging duration and a widening of the attentional pulse, resulting in a threshold that is higher than the identical induction sequence baseline. This possibility is logical and is consistent with the threshold for this condition

(see Figure 4). If period correction occurs for 50-msec asynchronies from the zero phase of the continuing internal oscillator (i.e., 500-msec IOI extension), why did period correction not also occur for the 950-msec IOI extension? We can only conjecture that, when more than one external interval marker is missing, there will not be period correction to the internal oscillator.

Beat-based mechanism. The key for the oscillator-based model being able to account for the results in Experiment 3 is the inability (exclusion) of highly deviant intervals (relative to the timing process's period) to alter the period. Keele et al. (1989) suggested such a capability for beat-based models, hypothesizing that, when a highly deviant interval is encountered, the internal beat can be suspended (excluding the interval) and restarted (a phase correction). On the basis of this Keele et al. hypothesis, we propose the beat-based/*restarting* model. This model has all of the properties of the beat-based/*adaptive* model, with the addition of an exclusion criterion that determines whether or not the internal beat is suspended and, thus, whether an interval contributes to the internal beat (see Table 1). Using the logic underlying our estimate of the exclusion criterion for the oscillator-based model, the typical exclusion criterion should again be around 150 msec for the stimuli used in Experiment 3. This added capability allows the beat-based/*restarting* model to account for the results in Experiment 3 in a manner similar to that of the oscillator-based model, again with the problem presented by the 500-msec IOI extension results.

Because the internal beat for the beat-based/*restarting* model suspends (i.e., does not continue) when the magnitude of asynchrony exceeds the exclusion criterion, the threshold for the 500-msec IOI extension condition should be similar to the identical induction sequence baseline. Although an oscillator should continue despite a missed external interval marker, should this be a property of a beat-based model? If the concept of suspension of the internal beat is taken literally, the internal beat should be suspended for *all* IOI extensions greater than 150 msec, and the results for the 500-msec IOI extension are not consistent with the beat-based/*restarting* model. If, however, the beat is allowed to continue for one missed beat, with the possibility of period adjustment based on a small asynchrony with the next beat, this model should make the same errors as the oscillator-based model. Although we now have a beat-based model that accounts for all of the present results, this post hoc model has additional properties that make it less parsimonious than the simple, rigid version of the beat-based mechanism (e.g., Schulze, 1978).

Interval-based mechanism. The lack of a monotonic relationship between threshold and IOI extension (see Figure 4) also is not consistent with the behavior of any interpretation of the memory model, including the multiple-look model. The only sensible, logical alternative to the memory and multiple-look models would require that the final extended induction sequence interval be excluded from calculation of the standard. Using the logic of al-

lowing equivalent capabilities across models based on the different timing mechanisms, we propose the multiple-look/*exclusion* model, which has the properties of the multiple-look model, plus an exclusion criterion that determines what intervals contribute to the calculation of the standard⁸ (see Table 1). The value of the exclusion criterion is equivalent in magnitude to the asynchrony for second oscillator entrainment for an oscillator-based model with that capability (and thus, when the internal beat is suspended and restarted for the beat-based/*restarting* model; see above), but the exclusion criterion for the multiple-look/*exclusion* model must be based on integer duration information. In the multiple-look/*exclusion* model, the extended final induction sequence interval is excluded from the calculation of the standard *only* when that interval deviates from the standard by a magnitude that exceeds the exclusion criterion (i.e., the extended final induction sequence interval is included in the calculation of the standard when that interval deviates from the standard by a magnitude that does not exceed the exclusion criterion). Again, the thresholds for all of the IOI extension conditions, except the 500-msec IOI extension, are consistent with the multiple-look/*exclusion* model (see Figure 4). Because the multiple-look/*exclusion* model is not cyclic in nature, there is no reason why the 500-msec IOI extension would be treated differently than other IOI extensions, and there is no reasonable modification that can be made to the multiple-look/*exclusion* model to account for the 500-msec IOI extension condition threshold. As with the other expanded models of the other mechanisms, the added properties of the multiple-look/*exclusion* model make it less parsimonious than the simpler models based on an interval-based mechanism (e.g., Keele et al., 1989).

Overall Summary

The present study was designed (1) to provide a clear understanding of basic temporal judgment abilities, as well as the effects of an initial contextual induction sequence and pauses on those abilities (as measured by thresholds), (2) to evaluate the three types of timing mechanisms proposed to explain such context effects and the importance of the specific capabilities incorporated in the models based on these mechanisms, and (3) to begin to evaluate whether the basic abilities and the relevance of the model attributes differ as a function of musical training. The results indicate that contextual information redundant with the judgment standard (T interval) improves performance but that the greatest contextual benefit is provided by only a small amount of redundancy; a single immediately preceding redundant interval provides the same benefit as five immediately preceding intervals. A redundancy benefit is consistent with there being some form of variability in the encoding and storage of interval information, but a conceptualization based on a statistical benefit of redundancy should result in equal benefits for each doubling of redundancy (e.g., one to two and two to four redundant intervals). It also would appear that adding a short pause in

a temporal sequence can disrupt timing accuracy but that once the effects of a relevant redundant induction sequence have been established, relatively long pauses (at least up to 2 sec) do not significantly interfere with temporal judgment ability. The simplest timing models, with limited capabilities, cannot explain the present patterns of results. The oscillator-based model is capable of providing an explanation of the present results, and the present results, in turn, provide insights into what model capabilities are important for the conditions we investigated. With capabilities equivalent to the oscillator-based model, a model based on a beat-based mechanism (the beat-based/restarting model) can explain most of the present results. With equivalent capabilities, a model based on an interval mechanism (the multiple-look/exclusion model) is less capable of explaining the full set of present results. With added processing capabilities and, thus, added parameters, these models, however, are less parsimonious than more rudimentary models. Finally, those listeners with at least some musical training typically exhibit better, but otherwise similar, abilities (typically, lower thresholds) than do listeners with little or no musical training; these differences can be accommodated by differences in model parameters (e.g., a lower exclusion criterion).

REFERENCES

- ABEL, S. M. (1971). Discrimination of temporal gaps. *Journal of the Acoustical Society of America*, **52**, 519-524.
- ALLAN, L. G. (1975). The relationship between judgments of successiveness and judgments of order. *Perception & Psychophysics*, **18**, 29-36.
- BARNES, R., & JONES, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, **41**, 254-311.
- BERENS, M. S., & PASTORE, R. E. (2003, June). *Temporal duration judgment with a variable induction sequence*. Poster session presented at the biannual meeting of the Society of Music Perception and Cognition, Las Vegas.
- BERENS, M. S., & PASTORE, R. E. (2005). *Is half of something equal to two of another? A comparison of halving and doubling capabilities*. Unpublished manuscript.
- CREELMAN, C. D. (1962). Human discrimination of auditory duration. *Journal of the Acoustical Society of America*, **34**, 582-593.
- DIVENYI, P. L., & DANNER, W. F. (1977). Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra. *Perception & Psychophysics*, **21**, 125-142.
- DIVENYI, P. L., & SACHS, R. M. (1978). Discrimination of time intervals bounded by tone bursts. *Perception & Psychophysics*, **24**, 429-436.
- DRAKE, C., & BOTTE, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics*, **54**, 277-286.
- GIBBON, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, **84**, 279-325.
- GRONDIN, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception & Psychophysics*, **54**, 383-394.
- GRONDIN, S., IVRY, R. B., FRANZ, E., PERREAULT, L., & METTHÉ, L. (1996). Markers' influence on the duration discrimination of intermodal intervals. *Perception & Psychophysics*, **58**, 424-433.
- HELSON, H. (1947). Adaptation-level as a frame of reference for prediction of psychophysical data. *American Journal of Psychology*, **60**, 1-29.
- HENRY, F. M. (1948). Discrimination of the duration of a sound. *Journal of Experimental Psychology*, **38**, 734-743.
- HIRSH, I. J. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America*, **31**, 759-767.
- IVRY, R. I., & HAZELTINE, R. I. (1995). Perception and production of temporal intervals across a range of duration: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 3-18.
- JONES, M. R., & YEE, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 693-709.
- KEELE, S. W., NICOLETTI, R., IVRY, R. B., & POKORNY, R. A. (1989). Mechanisms of perceptual timing: Beat-based or interval-based judgments? *Psychological Research*, **50**, 251-256.
- LARGE, E. W., & JONES, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, **106**, 119-159.
- LICKLIDER, J. C. R. (1951). Basic correlates of the auditory stimulus. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 985-1039). Oxford: Wiley.
- MADISON, G. (2004). Detection of linear temporal drift in sound sequences: Empirical data and modeling principles. *Acta Psychologica*, **117**, 95-118.
- 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.
- MILLER, J. D., WIER, C. C., PASTORE, R. E., KELLY, W. J., & DOOLING, R. J. (1976). Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. *Journal of the Acoustical Society of America*, **60**, 410-417.
- PASHLER, H. (2001). Perception and production of brief durations: Beat-based versus interval-based timing. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 485-493.
- PASTORE, R. E., CRAWLEY, E. J., BERENS, M. S., & SKELLY, M. A. (2003). "Nonparametric" *A'* and other modern misconceptions about signal detection theory. *Psychonomic Bulletin & Review*, **10**, 556-569.
- PISONI, D. B. (1980). Adaptation of the relative onset time of two-component tones. *Perception & Psychophysics*, **28**, 337-346.
- POVEL, D.-J., & ESSENS, P. J. (1985). Perception of temporal patterns. *Music Perception*, **2**, 411-441.
- RAMMSAYER, T. H. (1994). Effects of practice and signal energy on duration discrimination of brief auditory intervals. *Perception & Psychophysics*, **55**, 454-464.
- RAMMSAYER, T. H., & LEUTNER, D. (1996). Temporal discrimination as a function of marker duration. *Perception & Psychophysics*, **58**, 1213-1223.
- ROUSSEAU, L., & ROUSSEAU, R. (1996). Stop-reaction time and the internal clock. *Perception & Psychophysics*, **58**, 434-448.
- SCHULZE, H. H. (1978). The detectability of local and global displacements in regular rhythmic patterns. *Psychological Research*, **40**, 173-181.
- SMALL, A. M., & CAMPBELL, R. A. (1962). Temporal differential sensitivity for auditory stimuli. *American Journal of Psychology*, **75**, 401-410.
- SNYDER, J., & KRUMHANS, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception*, **18**, 455-489.
- SORKIN, R. D. (1990). Perception of temporal patterns defined by tonal sequences. *Journal of the Acoustical Society of America*, **87**, 1695-1701.
- WESTHEIMER, G. (1999). Discrimination of short time intervals by the human observer. *Experimental Brain Research*, **129**, 121-126.
- WOODWORTH, R. S., & SCHLOSBERG, H. (1959). *Experimental psychology*. Oxford: Slovenska Akademia Vied.

NOTES

1. It is unclear from Schulze's (1978) description of the memory model whether the standard is an average of only the induction sequence intervals or of the induction sequence and the T intervals. Because the critical comparison between the C and the T intervals should not benefit from a simple comparison of the T interval with an average of ear-

lier intervals, we assume that the average will be based on *all* intervals preceding the C interval, thus including the T interval.

2. The distinction between mean alone and mean + variance can be evaluated by having an induction sequence consisting of intervals with a variable duration, but with an average duration equal to the T interval (Berens & Pastore, 2003).

3. Schulze (1978) and Keele et al. (1989) tested the beat-based mechanism by manipulating the number of intervals in a sequence that were synchronous versus asynchronous with the conjectured internal beat.

4. For example, when there is 50% period correction, the period after an interval will equal the average of that interval and the period of the internal oscillator when that interval was presented.

5. We define threshold as $d' = 1$, which is approximately equal to 69% correct.

6. McAuley and Jones's (2003) manipulation of the timing of the ending marking tone of the T interval and the beginning marking tone of the C interval was *not* designed to address the effects of ISI duration.

7. This threshold estimation procedure is equivalent to a probit analysis of $p(D|D)$ after using $p(D|S)$ to correct for bias.

8. To explain why the introduction of a pause between the induction sequence and the judgment intervals did not affect relative temporal duration judgment for their task, McAuley and Jones (2003) proposed a similar parameter for a timing model that averages the induction sequence intervals.

APPENDIX

Table A1
Threshold Means (in Milliseconds, With Standard Errors) for
Musically Trained and Musically Untrained Listeners from
Experiment 1, Listed According to T and Intervening Interonset
Interval (IOI) Duration (in Milliseconds)

| T | Intervening IOI | Musically Trained | | Musically Untrained | |
|-----|-----------------|-------------------|-----------|---------------------|-----------|
| | | <i>M</i> | <i>SE</i> | <i>M</i> | <i>SE</i> |
| 450 | 0 (adjacent) | 29.00 | 1.46 | 40.00 | 3.30 |
| | 350 | 51.74 | 6.90 | 59.01 | 11.35 |
| | 450 | 24.26 | 5.01 | 26.08 | 2.81 |
| | 900 | 39.75 | 5.42 | 41.53 | 5.45 |
| 650 | 0 (adjacent) | 44.52 | 5.81 | 44.19 | 5.59 |
| | 350 | 64.95 | 3.87 | 64.34 | 10.00 |
| | 650 | 30.63 | 3.27 | 30.63 | 4.14 |

Table A2
Threshold Means (in Milliseconds, With Standard Errors) for
Musically Trained and Musically Untrained Listeners from
Experiment 3, Listed According to the Induction Sequence Interval
Duration (in Milliseconds)

| Duration | Musically Trained | | Musically Untrained | |
|----------|-------------------|-----------|---------------------|-----------|
| | <i>M</i> | <i>SE</i> | <i>M</i> | <i>SE</i> |
| 250 | 34.95 | 3.59 | 36.00 | 5.11 |
| 450 | 20.14 | 4.29 | 26.75 | 3.86 |
| 650 | 47.75 | 7.24 | 48.20 | 4.39 |

Table A3
Threshold Means (in Milliseconds, With Standard Errors) for
Musically Trained and Musically Untrained Listeners from
Experiment 3, Listed According to Interonset Interval Extension
Duration (in Milliseconds)

| Duration | Musically Trained | | Musically Untrained | |
|----------|-------------------|-----------|---------------------|-----------|
| | <i>M</i> | <i>SE</i> | <i>M</i> | <i>SE</i> |
| 50 | 32.02 | 2.88 | 35.92 | 4.73 |
| 100 | 34.36 | 5.35 | 46.03 | 10.50 |
| 200 | 24.43 | 3.11 | 19.95 | 3.32 |
| 225 | 18.90 | 1.82 | 29.34 | 5.80 |
| 300 | 22.68 | 2.75 | 24.63 | 2.68 |
| 400 | 26.82 | 1.40 | 29.20 | 4.19 |
| 450 | 17.63 | 5.37 | 18.28 | 3.83 |
| 500 | 27.55 | 4.46 | 35.25 | 4.02 |
| 900 | 21.95 | 3.02 | 28.74 | 3.67 |
| 950 | 19.89 | 5.35 | 27.70 | 4.65 |
| 2,250 | 22.56 | 4.13 | 22.32 | 2.49 |