On the Nature of Timing Mechanisms in Cognition

J. Devin McAuley
Department of Psychology
University of Queensland
Brisbane, Queensland 4072
Australia
`devin@psy.uq.edu.au`

Abstract

The ability to resolve timing differences within and between patterns is critical to the perception of music and speech; similarly, many motor skills such as music performance require fine temporal control of movements. Two important issues concern (1) the nature of the mechanism used for time measurement and (2) whether timing distinctions in perception and motor control are based on the same mechanism. In this paper, clock- and entrainment-based conceptions of time measurement are discussed; and predictions of both classes of model are then evaluated with respect to a tempo-discrimination experiment involving isochronous auditory sequences. The results from this experiment are shown to favor entrainment- over clock-based approaches to timing. The implications of these data are then discussed with respect to the hypothesized role of the cerebellum in timing.

Introduction

The dominant conception of time measurement for both perceptual and motor tasks has been based on a clock timer on a ms scale that records the duration of an event, such as the duration between the onset of two tones, as the number of millisecond "time slices" that occurred during that event. Braitenberg (1967) proposed that such a timer might be implemented in the cerebellum via a series of adjustable delay lines. In order to resolve timing differences that require duration comparisons, it has been commonly assumed that time estimates can be stored in memory, and then later retrieved (Keele et al., 1989). For motor tasks, recorded time measurements retrieved from memory have been assumed to serve as input to a motor program (Keele et al., 1985; Ivry and Keele, 1989).

An alternative view is that time measurement is phase based, involving the entrainment of an oscillatory timer. In contrast to the passive recording of time involved with clock models, entrainment is a dynamic timing process that adapts an oscillator's period to match the target duration; the oscillator's period then provides a direct estimate of duration, and the timing of events can be related to phase. In this paper, the clock and entrainment conceptions of time measurement are discussed in detail, and the predictions of both classes of model are then evaluated with respect to a tempo-discrimination experiment involving isochronous auditory sequences. The results from this experiment are shown to favor entrainment- over clock-based approaches to timing. The implications of these data are then discussed with respect to the hypothesized role of the cerebellum in timing (Ivry and Keele, 1989).

Clock Models

A wide range of clock models of timing have been proposed, although most share the same basic assumptions of a central timer (or clock), a perceptual store, a reference memory, and a comparator (Church and Broadbent, 1990); they differ mainly in the form of the clock and perceptual store. For many of the models, the clock is a fast neural pacemaker which generates discrete neural pulses at an average rate (λ), and the perceptual store maintains a "count" of the number of pulses that occur during the target interval (T) (Abel, 1972; Creelman, 1962; Treisman, 1963; Divenyi and Danner, 1977). Duration discrimination for two time intervals (T and $T + \Delta T$) is modeled by comparing the number of pulses that occured during the first interval ($\mu = \lambda T$) with the number of pulses that occured during the second interval ($\mu = \lambda[T + \Delta T]$). This approach requires a switch which starts the counting process at the beginning of the target time interval, and clears the counter when the estimate of the target interval is transferred from the perceptual store to the reference memory. Temporal resolving power is modeled by the variance of the pulse generating source, with small inter-pulse-variance corresponding to accurate estimates of duration and high discrimination sensitivity. In modeling duration discrimination across a range of T values, debate has centered on the precise form of the pulse-generating source.

There have been several connectionist approaches to time measurement based on the clock conception (Church and Broadbent, 1990; Miall, 1989). These differ from the "counter" variety in important ways. For Church and Broadbent (1990), the pulse-generating source is replaced by a set of oscillators with periods spanning a wide range of time intervals; the pulse counter is replaced by a binary vector representation of time, according to the +1/-1 phase of each oscillator; and, the reference memory storing the pulse count is replaced by a set of connection weights, permitting more than a single time interval to be stored in the reference memory at once. In connectionist clock models, duration discrimination is based on similarity in the representations of the to-be-compared interval retrieved from memory; if the measure of similarity is less than a prespecified threshold, the time difference is detected.

A weakness of both the connectionist- and counter-based clock models is that their predictions are usually limited to the perception of isolated intervals (i.e, one interval compared with another in isolation); whereas, important temporal distinctions in music and speech occur within the context of a pattern of intervals. However, in the direction of incorporating pattern context, clock models have recently been proposed to model the effect of isochronous contexts on tempo (rate) discrimination. These models are based on a "multiple-look" hypothesis, in which each interval in an isochronous sequence provides an independent statistical (clock-based) estimate (or "look") (Drake and Botte, 1993; Schulze, 1989). With multiple-observations of the same target interval, the perceiver improves the estimate of the target interval's duration by a process of averaging the multiple-looks, or in the case of tempo discrimination, improves the estimate of the the sequence's tempo, by the same process. An identical suggestion is that the stability of a target interval's memory trace improves with repetitions of the target interval. (Keele et al., 1989; Ivry and Hazeltine, 1995).

Entrainment Models

At the foundation of all entrainment models of timing is the assumption that the timing mechanisms of the nervous system are coupled to the environment. In the development of an entrainment theory, Jones (1976) proposed a central role for rhythm in cognitive processing, suggesting that the temporal organization of perception, attention, and memory is inherently rhythmic. As part of this theory, it is assumed that the rhythms of music and speech entrain periodic attentional "pulses", forming an attentional rhythm. Based on the concept of attentional entrainment, Jones and Boltz (1989) have proposed an expectancy/contrast model of timing. They assume that an isochronous series of tones, marking out identical time intervals (T_i) , will entrain an attentional oscillator with a period similar to the T_i's. The pulses of the oscillator provide dynamic "expectancies" for when the next tone (specifying T_{i+1}) will occur. In this way, the adapting period of the oscillator (Ω_i) provides a continuously updated estimate of the time intervals (T_i). Intervals (T_i), which violate the oscillator's period-based expectancies create a temporal contrast ($\Omega_i - T_i$). As the attentional oscillator is entrained by the sequence, temporal contrast is minimized (i.e., Ω_i approaches T_i).

Temporal resolving power in the Jones and Boltz model is based on the predicted magnitude of temporal contrasts. It is assumed that with small temporal contrasts, listeners will be more sensitive to a timing change than with larger temporal contrasts. Thus, similar to the "multiple-look" model, the expectancy/contrast model predicts that increasing the number of tones in an isochronous sequence should improve listeners' ability to detect a difference in the timing of an interval that continues the sequence, and should also improve listeners' ability to detect changes in the tempo of that sequence. The expectancy/contrast model also predicts that listeners' temporal resolution should be better with metrical sequences (of which isochronous sequences are an instance) than for irregularly-timed sequences, since attentional entrainment should occur more readily with regularly-timed sequences.

Adaptive Oscillators

A shortcoming of the expectancy/contrast model is that the hypothesized process of attentional entrainment is a descriptive component of the model, and not linked to a specific mathematical model of coupled oscillation. This makes the expectancy/contrast model, and other entrainment models (Schulze, 1978) a relatively easy target for criticism, since many of its predictions are underspecified, and open to multiple interpretations (Keele et al., 1989). As steps toward clarifying its predictions, the entrainment model has been formalized and a timing mechanism proposed that is based on an adaptive oscillator (McAuley, 1994; McAuley, 1995). This work parallels similar recent work by Large (1994).

The adaptive oscillator is a processing unit that has some resting rate at which it periodically "fires", but will adapt that resting rate when it is stimulated at different frequencies, combining both phase coupling and period coupling. In its simplest "phase-resetting" form, the adaptive oscillator resets its phase in response to an input pulse and will use the input's phase to adjusts its natural period (Ω) to be a little closer to the "perceived" periodicity of the input. This process permits adaptive oscillators to track periodic components of rhythmic patterns, despite intrinsic or expressive variability in their timing (see McAuley (1995) for a mathematical description of this process).

Time as Phase

For entrainment models of timing, the oscillator period (Ω) provides an implicit estimate of a duration (T). For those that are based on the phase-resetting adaptive oscillator, a time change (ΔT) in the duration (T) will trigger a phase change ($\Delta\phi$) in the reset-phase of the oscillator mechanism, where

$$\Delta\phi = \frac{T + \Delta T}{\Omega} \pmod 1. \tag{1}$$

In essence, this phase difference ($\Delta\phi$) registers the effect of a time difference (ΔT) on an oscillator tracking a series of equal intervals (T). In an entrainment model, the relationship between phase differences ($\Delta\phi$) and time differences (ΔT) varies as a function of the ratio between the base interval T and the estimate of duration Ω, expressed as the fraction $\frac{T}{\Omega}$. This fraction provides a measure of the amount of over- or underestimation of duration by the entrainment mechanism; for $\frac{T}{\Omega} > 1$, duration is underestimated, for $\frac{T}{\Omega} < 1$ duration is overestimated, and for $\frac{T}{\Omega} = 1.0$ estimated duration is identical to the actual duration.

To understand how the relationship between time differences and phase differences varies as a function of this ratio ($\frac{T}{\Omega}$), it is useful to rewrite Equation 1 as

$$\Delta\phi = [\frac{T}{\Omega}] + [\frac{\Delta T}{\Omega}] \pmod 1 \qquad (2)$$

and to represent phase on $[-0.5, 0.5]$ instead of on $[0, 1]$; in this representation, positive and negative phase differences indicate positive and negative time differences, respectively. For this reason, the phase differences associated with $+\Delta T$ and $-\Delta T$ will be distinguished as $\Delta\phi_+$ and $\Delta\phi_-$. There are then three cases to consider.

Case 1: $\frac{T}{\Omega} = 1.0$. For Case 1, illustrated in Figure 1, estimated duration is equal to actual duration, Equation 2 reduces to

$$\Delta\phi = \pm\frac{\Delta T}{T}. \qquad (3)$$

In this case, lengthening or shortening T by $X\%$ maintains the magnitude of phase difference regardless of whether the time change ΔT is positive or negative (i.e., $|\Delta\phi_-| = |\Delta\phi_+|$). To provide a concrete example, suppose T is lengthened by 10%, then $\Delta\phi_+ = 0.1$. On the other hand, if T is shortening by 10%, then $\Delta\phi_- = -0.1$. For either an increase or decrease in duration of 10%, the magnitudes of the triggered phase differences are equal ($|\Delta\phi_+| = |\Delta\phi_-| = 0.1$).

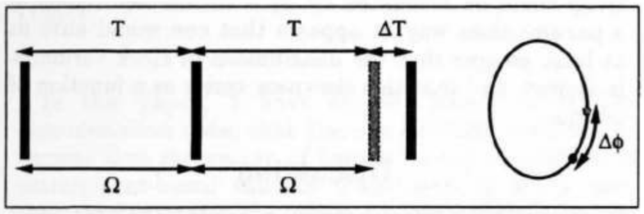

Figure 1: Illustration of Case 1. For an oscillator with a period Ω that is a perfect estimate of T (as shown on the left), lengthening T by ΔT triggers a phase difference $\Delta\phi$ relative to zero phase of the oscillator (as shown on the right).

Case 2: $\frac{T}{\Omega} > 1.0$. For Case 2, illustrated in Figure 2, actual duration is underestimated ($\Omega < T$). As a result, lengthening or shortening T by $X\%$ does not preserve the magnitude of the resulting phase differences. Instead, lengthening T triggers a phase difference that is larger than that for shortening T: $|\Delta\phi_+| > |\Delta\phi_-|$. Another effect of underestimation is that it stretches the mapping between ΔT and $\Delta\phi$; notice that in terms of Equation 2, $\frac{\Delta T}{\Omega}$ is a larger fraction of the base interval T than $\frac{\Delta T}{T}$ is.

Case 3: $\frac{T}{\Omega} < 1.0$. For Case 3, illustrated in Figure 3, actual duration is overestimated ($\Omega > T$). As in case 2, lengthening or shortening T by $X\%$ does not preserve the magnitude of the resulting phase differences. Symmetric with the effect in Case 2, shortening T triggers a phase difference that is larger than lengthening: $|\Delta\phi_-| > |\Delta\phi_+|$. And the additional effect of overestimation is that it compresses the mapping between ΔT and $\Delta\phi$; notice that in terms of in Equation 2, $\frac{\Delta T}{\Omega}$ is a smaller fraction of the base interval T than $\frac{\Delta T}{T}$ is.

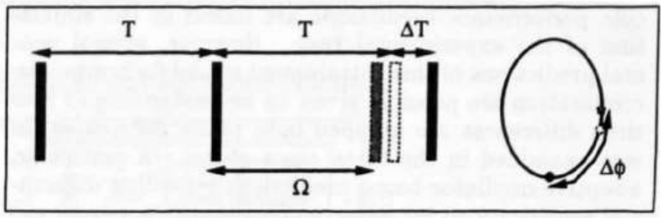

Figure 2: Illustration of Case 2. For an oscillator with a period Ω that is an underestimate of T, lengthening T by ΔT triggers a phase difference $\Delta\phi$ that is larger than that for Case 1.

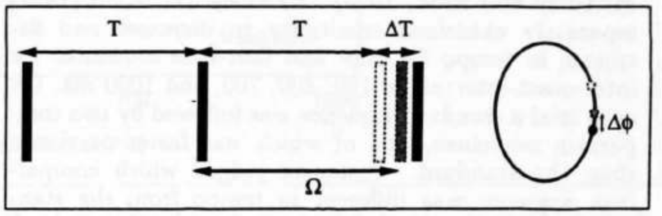

Figure 3: Illustration of Case 3. For an oscillator with a period Ω that is an overestimate of T, lengthening T by ΔT triggers a phase difference $\Delta\phi$ that is smaller than that for Case 1.

The Just-Noticeable Phase Difference

In order to specify the predictions of an entrainment model for tempo discrimination, it was assumed that the detection of tempo differences is linked to the detection of phase differences, triggered by "expectancy violations" in the entrainment process (McAuley, 1995). The suggestion that phase differences are used to detect timing differences pins the predictions of an entrainment model to the dynamics of the underlying oscillatory timer. For the adaptive oscillator mechanism, a just-noticeable phase difference (JND_ϕ) is assumed that specifies the threshold time difference (ΔT) that is detectable in a interval T. Thus, if the magnitude of the phase-difference ($\Delta\phi$) triggered by a time difference (ΔT) is greater than JND_ϕ then the time difference (ΔT) is detected, otherwise it is not. In accordance with the entrainment hypothesis, the just-noticeable phase difference decreases (sensitivity improves) as the tracking adaptive oscillator is entrained by the input (see McAuley (1995) for details).

In evaluating the predictions of the adaptive-oscillator-based model for tempo discrimination, the main focus is the relationship between listeners' sensitivity, measured as the ΔT necessary for unbiased 70% correct performance (the just-noticeable difference or JND) and the model's sensitivity measured as a just-noticeable phase difference. The relationship between JND and JND_ϕ is a dynamic one, depending on the amount of under- or overestimation of the tracking adaptive oscillator and its degree of entrainment by the input. Specific performance predictions are linked to the simulation of the experimental task. However, several general predictions of this entrainment model for tempo discrimination are possible given an understanding of how time differences are mapped onto phase differences, as was examined in the three cases above. In particular, adaptive-oscillator-based predictions regarding differential sensitivity to increases and decreases in tempo will be discussed below, and compared with those derived from clock models.

Tempo Discrimination

In several recent studies, listeners' ability to detect differences in the tempo of isochronous tone sequences has been systematically investigated (Drake and Botte, 1993; Drake and Botte, 1994; McAuley and Kidd, 1994; McAuley and Kidd, 1995). McAuley and Kidd (1994) separately examined sensitivity to increases and decreases in tempo for two- and four-tone sequences for inter-onset-intervals of 100, 400, 700, and 1000 ms. On each trial a standard sequence was followed by two comparison sequences, one of which was faster or slower than the standard. Listeners judged which comparison sequence was different in tempo from the standard. Separate thresholds were obtained for "faster" and "slower" trials using an adaptive tracking procedure (Levitt, 1971). Consistent with Drake and Botte (1993), thresholds were found to be lower with four-tone sequences than with two-tone sequences, especially at the faster tempos. However, at the fastest tempos, listeners showed greater sensitivity to increases in tempo than to decreases in tempo, while the reverse was true at the slower tempos. These data are illustrated in Figure 4.

The results from this experiment confirm the predictions of the entrainment model. If an intrinsic "preferred" period of the system of around 600 ms is assumed, long IOIs (slow tempos) are underestimated and short IOIs (fast tempos) are overestimated, consistent with empirical data (Fraisse, 1982). For entrainment models, under- or overestimation indicates that the phase difference ($\Delta\phi$) corresponding to *no* time difference ($\Delta T = 0.0$) is skewed from $\Delta\phi = 0$ to a positive or negative value, as described in Cases 2 and 3. With underestimation, a tempo increase of $X\%$ triggers a phase difference that is smaller than that triggered by the same tempo decrease; thus, with underestimation, the model predicts greater sensitivity for tempo decreases (slowing down) than for tempo increases (speeding up). On the other hand, with overestimation, a tempo increase of $X\%$ triggers a phase difference that is larger than that triggered by the same tempo decrease; thus, with overestimation, the model predicts greater sensitivity for tempo increases than for tempo decreases. As described above, this pattern of differential sensitivity is found with listeners (McAuley and Kidd, 1994).

If it is assumed that in the limit (e.g., an isochronous standard sequence with a large number of tones), the adaptive oscillator is perfectly entrained by the tempo of the standard, in which Case 1 applies and $\Omega = T$, then comparison sequences that are faster or slower by the same percentage will trigger equal phase differences, and the detection thresholds for increases and decreases in tempo will be the same. Thus, in the limit, differential sensitivity should disappear. For the reported data, differential sensitivity did disappear for the four-tone sequences, but only for the fastest tempos. In a simulation of the entrainment model (McAuley, 1995), this same interaction was produced by the dynamic interaction between period-coupling and a period-decay process. At fast rates, the input pulses driving the entrainment process occurred at a fast enough rate to enable the model to achieve perfect entrainment, in spite of the counteracting effects of period-decay. In contrast, at slower rates, the input pulses did not drive entrainment quickly enough to completely counteract the effects of period-decay, and the model did not achieve perfect entrainment; hence, differential sensitivity did not disappear for the slower tempos, as was also found with the listeners.

In contrast, the clock models discussed in this paper do not predict the pattern of differential sensitivity found with listeners, since temporal resolution is based on the variability of the clock process, independent of the stimulus. Consequently, no performance distinctions between faster/slower comparisons are made. In addition, it is not clear how the observed pattern of differential sensitivity could be accounted for by a clock-based model in a parsimonious way; it appears that one would have to at least assume that the distribution of clock variances is skewed, and that this skewness varies as a function of tempo.

Discussion

Wing and Kristofferson (1973) extended the concept of a clock timer to the production of regularly timed intervals (such as finger tapping). They assume that tapping variability arises from two independent processes: the clock component and a motor delay component. By making the independence assumption, they provide an elegant method based on analysis of auto-covariance of inter-tap-intervals, to decompose the tapping variability into the variances of the component processes (see Wing and Kristofferson (1973) for details). They suggest that the negative lag-one covariance often observed in tapping tasks may not be due to a compensatory timing mechanism, but can instead be explained as an implementation delay introduced by an independent motor component.

Based on this decomposition of tapping variability into clock and motor components, Ivry and Keele (1989) propose that the cerebellum acts as a central clock timer for both the perception of duration and the control of movements. They report that for both the perception and production of temporal intervals, larger clock-variability estimates are obtained with neurological patients with cerebellar deficits than are obtained with controls. However, this support for a cerebellar role in timing is linked to the assumptions of the Wing and Kristofferson model, requiring a firm commitment to the nature of the cerebellar timing mechanism, one that is clock-based (see also (Keele et al., 1985)).

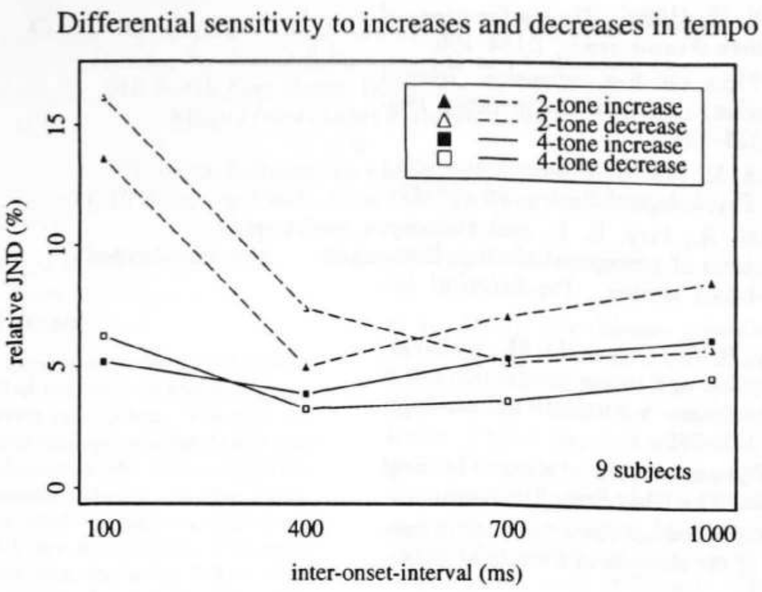

Figure 4: Mean discrimination thresholds of listeners for the detection of increases and decreases in the tempo of two- and four-tone isochronous sequences for base IOIs of 100, 400, 700, and 1000 ms (reproduced from McAuley (1995)).

In this paper, I have argued, based on tempo-discrimination data, that there is sufficient evidence to suggest that the nature of human timing capabilities is entrainment-based and not clock-based, in which case, the assumptions of the Wing and Kristofferson model (1973) are wrong. This calls into question the decomposition of tapping variability into clock and motor components. I do not intend here to argue against a cerebellar role in timing, but rather to suggest that data used to support the cerebellar timing hypothesis should at least be reanalyzed, taking into account *the possibility* that timing is based on entrainment. Braitenberg (1967) proposed that a clock-based timer might be implemented in the cerebellum via a series of adjustable delay lines; it has since been argued that the maximum neural delays in the cerebellum are too short to provide a reasonable clock implementation (Fahle and Braitenberg, 1984). The adaptive-oscillator mechanism discussed in this paper provides an alternative suggestion for a cerebellar timer, one that is entrainment-based.

Acknowledgments

The author thanks Simon Dennis, Gary Kidd and Catherine Rogers for their constructive comments on this research. This research is supported by a University of Queensland Postdoctoral Fellowship.

References

Abel, S. M. (1972). Discrimination of temporal gaps. *Journal of the Acoustical Society of America*, 52(2):519–524.

Braitenberg, V. (1967). Is the cerebellar cortex a biological clock in the millisecond range? *Progress in Brain Research*, 25:334–346.

Church, R. M. and Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition*, 37:55–81.

Creelman, C. D. (1962). Human discrimination of auditory duration. *Journal of the Acoustical Society of America*, 34(5):582–593.

Divenyi, P. L. and Danner, W. F. (1977). Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra. *Perception and Psychophysics*, 21(2):125–142.

Drake, C. and Botte, M. C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception and Psychophysics*, 54(3):277–286.

Drake, C. and Botte, M. C. (1994). The measure and modelization of tempo sensitivity in auditory sequences. *Journal of the Acoustical Society of America*, 95(5):2966.

Fahle, M. and Braitenberg, V. (1984). Some quantitative aspects of cerebellar anatomy as a guide to speculation on cerebellar functions. In *Cerebellar functions*. Spinger-Verlag, Berlin.

Fraisse, P. (1982). Rhythm and tempo. In Deutsch, D., editor, *The Psychology of Music*, pages 149–180. Academic Press, New York.

Ivry, R. I. and Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of duration: Evidence for a common timing mechanism. *Jour-*

nal of Experimental Psychology: Human Perception and Performance, 21(1):3–18.

Ivry, R. I. and Keele, S. W. (1989). Timing functions of the cerebellum. *Cognitive Neuroscience*, 1:134–150.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83:323–355.

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

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

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

Large, E. W. (1994). Dynamic representation of musical structure. Ph.D. Thesis, The Ohio State University.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49:467–477.

McAuley, J. D. (1994). Time as phase: A dynamic model of time perception. In *Proceedings of the Sixteenth Annual Meeting of the Cognitive Science Society*, pages 607–612, Hillsdale, New Jersey. Lawrence Erlbaum Associates.

McAuley, J. D. (1995). Perception of Time as Phase: Toward an Adaptive-Oscillator Model of Rhythmic pattern processing. Ph.D. Thesis, Indiana University.

McAuley, J. D. and Kidd, G. R. (1994). Differential sensitivity to increases and decreases in tempo: Evidence for an entrainment model. *Journal of the Acoustical Society of America*, 96(5):3257.

McAuley, J. D. and Kidd, G. R. (1995). Temporally directed attending in the discrimination of tempo: Further evidence for an entrainment model. *Journal of the Acoustical Society of America*, 97(5):3278.

Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Computation*, 1(3):359–371.

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

Schulze, H. H. (1989). The perception of temporal deviations in isochronic patterns. *Perception and Psychophysics*, 45(4):291–296.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the internal clock. *Psychological Monographs: General and Applied*, 77(13).

Wing, A. M. and Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception and Psychophysics*, 14(1):5–12.