Perception & Psychophysics
2002, 64 (6), 909-918

Detection of metric structure in auditory figural patterns

SYLVIE HÉBERT and LOLA L. CUDDY
Queen's University, Kingston, Ontario, Canada

This series of experiments dealt with discrimination between two temporal patterns differing only by the insertion of an additional silent gap. In Experiment 1, patterns varied in metric and figural structure. Metric structure is described as the sense of temporal regularity that may occur between subjectively accented tones. Figural structure is described as the grouping of temporally adjacent tones separated by silences. Standard patterns were either strongly or weakly metric; comparison patterns differed from the standards by the insertion of a silence that disrupted either the metric structure alone or both the metric and the figural structures. Experiment 1 provided support for the roles of both metric and figural structures and provided support for the clock-induction model of Povel and Essens (1985) as an account of metric processing. In Experiments 2–4, discrimination of patterns with differing metric structures but identical figural structures was examined more closely. Rate of presentation of the patterns was varied. Multiple regression indicated that, independent of rate variations, discrimination improved as the absolute (not relative) duration of the silent gap increased. We argue that an additional timing mechanism, independent of pattern structure, is operative in temporal pattern discrimination. All the results were replicated across levels of music training of the listeners.

There are several main approaches in rhythm research. One is concerned with interactions between rhythm and other dimensions encountered in music. In some studies, for example, the influence of rhythm on the perception of melodic structure (Boltz, 1989a, 1989b; Jones & Ralston, 1991; Jones, Summerell, & Marshburn, 1987; Smith & Cuddy, 1989) or the influence of harmonic structure on the perception of rhythm (Dawe, Platt, & Racine, 1993, 1994) has been examined. In another approach, studies have uncovered instances in which the processing of rhythm and the processing of melody are co-existent but independent (e.g., Palmer & Krumhansl, 1987a, 1987b) or are functionally dissociated following brain insult (e.g., Peretz, 1990; Peretz & Kolinsky, 1993). Still another approach has been concerned with rhythmic processing per se. Stimulus patterns vary temporally, unaccompanied by variations in pitch, timbre, or dynamics. Studies have focused on both production (Drake, 1993, 1998; Drake & Gérard, 1989; Essens, 1986, 1995; Essens & Povel, 1985; Povel & Essens, 1985; Smith, Cuddy, & Upitis, 1994) and perception (Bharucha & Pryor, 1986; Handel, 1992, 1998; Hirsh, Monahan, Grant, & Singh, 1990; Monahan & Hirsh, 1990; Palmer & Krumhansl, 1990; Povel & Essens, 1985; Ross & Houtsma, 1994). In the present paper, the tradition of the last approach is followed. We examine discrimination between two auditory tone patterns differing only in their temporal properties.

The patterns may be described as figural structures. A figure consists of events (here, tones) similar in spectral composition and contiguous in time. An event bounded by silence—an isolated event in the pattern—is also considered a figure. A simple perceptual code for figural structures has been proposed, a code that preserves information about the number of tones in each successive figure and the number of figures (Bamberger, 1982; Handel, 1992, 1998; Povel & Essens, 1985; Ross & Houtsma, 1994). For example, a sequence of four figures might be coded 3212, indicating three tones in the first figure, two in the second figure, and so on. The duration between tones within a figure and the duration of the time intervals between figures are not preserved in the figural code. To establish the code, it is only necessary to detect a difference between a short (between tones of a figure) and a long (between figures) duration.

A number of experiments have been used to assess listeners' ability to discriminate between two short auditory patterns differing in figural structure. Discrimination is generally very high for both infant listeners (Demany, McKenzie, & Vurpillot, 1977; Thorpe & Trehub, 1989) and adult listeners (Handel, 1998; Ross & Houtsma, 1994, Experiment 1). The perception of auditory figures may represent an application of general Gestalt grouping principles of similarity and proximity (Koffka, 1935; Köhler, 1947).

As well as figural organization, however, another source of information may be available to the listener. The music

---

This paper was supported by a postdoctoral fellowship from Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (Government of Quebec) to S.H. and by a research grant from the Natural Sciences and Engineering Research Council of Canada, awarded to L.L.C. Some of the results were presented at the meeting of the Society for Music Perception and Cognition held in Boston in 1997. We are grateful to Carol Krumhansl, Jaan Ross, and Dirk J. Povel for their comments on an earlier version of the paper. S.H. is now at the School of Speech Pathology and Audiology, University of Montreal. Correspondence concerning this article should be sent to L. L. Cuddy, Department of Psychology, Queen's University, Kingston, ON, K7L 3N6 Canada (e-mail: cuddyl@psyc.queensu.ca).

literature has identified another rhythmic structure, known as metric structure (e.g., Lerdahl & Jackendoff, 1983). Metric structure is based on the coding of precise timing information between successive tones both within and between figures and also between nonsuccessive tones. Moreover, metric structure calls for a concept of temporal hierarchies in which the timing of successive events is subservient to the timing of nonsuccessive events separated by longer temporal durations. For the present paper, this notion will be elaborated with reference to the clock-induction model proposed by Povel and Essens (1985). The clock-induction model invokes two hierarchic levels. One level corresponds to the *ticks* of the clock and defines the perceptual unit; the other corresponds to subdivisions of the unit. The unit also corresponds to the musical concept of beat or tactus, regular temporal intervals at which a listener might clap along with a rhythmic sequence.

According to Povel and Essens (1985), two critical factors determine the instantiation of the internal clock and the strength of the clock. The first is the perception of subjective accents, the perceptual salience of certain tones within a sequence of physically identical tones. Isolated tones, the second tone in a pair of two tones, and the first and last tones in a group of three or more tones will be subjectively accented, even in the absence of physical accents (Povel & Okkerman, 1981). The second factor is the distribution in time of the subjective accents. Subjective accents regularly distributed in time will induce a strong clock whose ticks are synchronized with the accented tones. If the accented tones are irregularly distributed in time, ticks of a clock and the accented tones cannot be (or will be less) synchronized. A weak clock or no clock at all will be induced. The more frequently the accented tones coincide with regular ticks of the clock, the stronger the induced clock. Povel and Essens expressed the total amount of counterevidence against a clock by the formula $C = 4E0 + E1$. $E0$ represents the number of silences coinciding with a tick, and $E1$ represents the number of unaccented tones coinciding with a tick. The strongest clocks are those with the lowest C value; conversely, the weakest clocks are those with the highest C value. Production studies showed that the reproduction of patterns assumed to induce strong clocks was more accurate than the reproduction of those inducing weak clocks (Essens, 1986, 1995; Essens & Povel, 1985; Povel & Essens, 1985).

Ross and Houtsma (1994, Experiment 2) further examined the role of metric structures in a discrimination experiment. Patterns consisted of randomly ordered sequences of clicks and silences. Each trial required discrimination between a standard and a comparison pattern, where the comparison pattern was either the same as the standard or included a disruption. The disruption was an additional silent gap inserted between figures; thus, the figural structure of the standard pattern was always preserved in the comparison. Patterns differed in metric strength, however, according to the measure proposed by Povel and Essens (1985). Discrimination was found to be more accurate for patterns assumed to induce a strong clock than for patterns assumed to induce a weak clock. Ross and Houtsma concluded that strong metric patterns generated a more precise timing grid than did nonmetric or weakly metric patterns and, hence, yielded better discrimination.

Ross and Houtsma's (1994) conclusion was challenged by more recent work by Handel (1998). In Handel's (1998) study, comparison patterns differed from standard patterns by the permutation of a click and a silence. This change either preserved or modified the figural structures and also altered the metric strength. As might be expected, Handel (1998) found figural structure to be a powerful cue. When the figural structures of the standard and the comparison patterns differed, discrimination was very high. When figural structure was preserved, however, discrimination between patterns was at chance. This performance failure when figural structure was preserved is an indication that the timing between figures was not encoded. Also, whether or not figural structures differed, there was little evidence of metric strength. Metric strength affected discrimination under only two conditions, and then only weakly so—when an external pulse was added to enhance metric structure and when the standard pattern contained a strong meter and the comparison a weaker meter. This severely limited role of metric structure led Handel (1998) to question the application of the musical concept of meter to discrimination tasks.

Thus, Handel's (1998) conclusion, that timing between figures is generally unavailable, contrasts with the results of Ross and Houtsma (1994) and with the view of Povel and Essens (1985). Our series of experiments, therefore, returns to the assessment of the role of metric structure in auditory pattern discrimination. Empirical support in favor of the role of metric, as well as figural, structures is gathered in Experiment 1. We provide evidence that listeners do process timing between figures (i.e., between groups of tones) and that a strong clock facilitates processing. Accordingly, Experiments 2–4 pursue the findings further. We ask, first, a procedural question: whether the evidence of timing between figures is restricted to the use of the same tempo (rate of presentation) throughout the experiment. We report that it is not so restricted. Second, we ask whether the timing between figures is measured in subdivisions of the clock unit. We report that it is not. Rather, we propose that an absolute context-free timing mechanism is engaged along with the context-dependent mechanism that determines metric strength.

EXPERIMENT 1

In Experiment 1, listeners were asked to discriminate between a standard and a comparison pattern. Each standard pattern was expected to induce either a strong or a weak clock (Povel & Essens, 1985). Each comparison pattern was either the same as or different from the standard. If different, a silence was added so that the figural structure was changed (hereafter, called *disruption within*, because the disruption was inserted within figures) or left unchanged (hereafter, called *disruption between*, because the disruption was inserted between two figures).

If metric structure facilitates the processing of timing information, we would expect a definite and overall influence of clock strength on discrimination. In particular, given the results of Ross and Houtsma (1994), the effect of clock strength should hold where the patterns to be compared have identical figural structures. However, if only figural structure is relevant for discrimination, as Handel's findings (1998) would suggest, only patterns with different figural structures will be discriminated. Comparison patterns with a disruption between figures should be undistinguishable from standard patterns—that is, performance should be at or near chance in the disruption-between condition.

Method

Listeners. The listeners were 12 women and 4 men from the university community, with a mean age of 26.6 years (range, 18–47). They either received participation credit for an introductory psychology course or were paid for their services. They were recruited for "an experiment in rhythm perception," without mention of music training or other prerequisites. The music background of each listener, however, was obtained after the experimental session and was scored as follows: One point was given for each Royal Conservatory (or equivalent) grade achieved, and half a point for each year of any other type of musical training. Music training points for this experiment averaged 3.3 (range, 0–14). All the listeners self-reported normal hearing.

Apparatus. FM-synthesized sounds were generated by a Yamaha TX81Z synthesizer. The "hand-drum" timbre was selected from the preset factory voices. The perceived pitch of the hand-drum corresponded approximately to middle C ($F0 = 262$ Hz). The amplitude envelope of each sound had a steep rise time of 10 msec, followed by a gradual negative exponential decay. All the sounds had the same intensity with no physical accents. The sounds were delivered through Monitor Audio TR-159 speakers located in a sound-isolated testing booth. Presentation and data collection were controlled by a IBM-compatible computer running MIDILAB (Todd, Boltz, & Jones, 1989).

Patterns. The standard patterns were the 10 patterns shown in Figure 1. Column 1 provides the category numbers and the pattern numbers from Povel and Essens (1985, Table 2, p. 423). Column 2 provides the intervals contained in each pattern. An interval is defined as the onset-to-onset duration between successive events. Intervals are numbered 1, 2, 3, and 4. Interval 1 was 200 msec, as in Povel and Essens; Intervals 2, 3, and 4 were 400, 600, and 800 msec, respectively. Each of the 10 patterns has a different permutation of the same set of four different intervals—111112234. In column 3, patterns are represented on a time axis, with each vertical line representing an event onset. Each dot represents a silence of 200 msec. In this representation, successive vertical lines that are spatially proximate are the figural groups. Ellipses in the first pattern of each category indicate the groups.

Above the event onsets in column 3 of Figure 1, the symbol > indicates the predicted locations of subjective accents. The horizontal line underneath each temporal pattern represents the "best" internal clock and its ticks. Each of these clocks was selected over other possible clocks because it minimizes the *C* value. The top five patterns in the figure are patterns that fit best with a clock that ticks at regular intervals of 800 msec. Ticks of the clocks coincide exactly with subjective accents. These five patterns should thus induce a strong clock, the strongest clock among Povel and Essens's (1985) patterns, with a *C* value of zero. The bottom five patterns do not fit a regular clock. Ticks of the clock sometimes coincide with silences or with unaccented events. These patterns are thus expected to induce a weak (or no) clock and are the weakest clocks among Povel and Essens's patterns, with a *C* value of 6.[1] The classification of strong- and weak-clock patterns was verified by an alternative analysis (see the Appendix).

Procedure. A trial consisted of the presentation of 1 of the 10 patterns repeated once without interruption, followed by a pause, followed by a comparison pattern, also repeated once. The pause between the standard and the comparison patterns was equal to the duration of 2.5 intervals. A block consisted of 30 trials (10 standard patterns × 3 comparison patterns) randomly ordered, with the constraint that no 2 trials with the identical standard pattern could occur in succession. Two blocks of trials, out of a possibility of eight different blocks, were selected randomly for each listener, with the restriction that each block would be selected twice across the 16 participants.

In the *same* condition, the comparison pattern was identical to the standard. The *different* conditions involved an added silence that was placed either within a group of tones (disruption-within condition) or between two groups of tones (disruption-between condition). In the disruption-within condition, an interval of 200-msec duration, randomly selected, was lengthened to 400 msec. In the disruption-between condition, an interval of 600 msec was lengthened to 800 msec. Different comparison patterns were always weak-clock patterns.

Practice trials preceding the experimental trials followed the same procedure as the experimental trials. Standard patterns for these trials were patterns in Povel and Essens (1985, Table 2) that represented moderately strong clocks.

Trials were presented at a comfortable listening level. The listeners were instructed to judge whether the two patterns in each trial were *same* or *different* and to input their response by pressing one of two keys on the computer keyboard. After each response was entered, the message "correct" or "incorrect" appeared on the monitor screen for 2 sec. After a 1-sec pause, the next trial began.

Results and Discussion

There were six experimental conditions: three levels of type of comparison (same, disruption within, and disruption between) crossed with two levels of clock strength of the standard pattern (strong and weak). The proportions of *different* responses for each listener and each condition were used to derive d' scores from the tables of signal detection theory (Green & Swets, 1966). Hits were defined as a correct *different* response to a different comparison pattern. False alarms were defined as an incorrect *different* response to a same comparison pattern. Proportions of 0 and 1 were adjusted by converting 0 proportions to $1/(2N)$ and 1 proportions to $1.5N$ (Macmillan & Kaplan, 1985). Each listener yielded four d' scores, one for each combination of different comparison conditions and clock strength. The mean d' scores are given in Table 1A.

Table 1A shows that the disruption-within condition yielded higher scores than the disruption-between condition [$F(1,15) = 40.65$, $MS_e = 0.40$, $p < .001$]. As well, scores were higher overall for strong-clock than for weak-clock patterns [$F(1,15) = 17.84$, $MS_e = 0.49$, $p < .002$]. Scores for the weak-clock, disruption-between condition were significantly higher than chance [$d' = 0.0$; $t(15) = 2.94$, $p < .02$]. The interaction between comparison con-

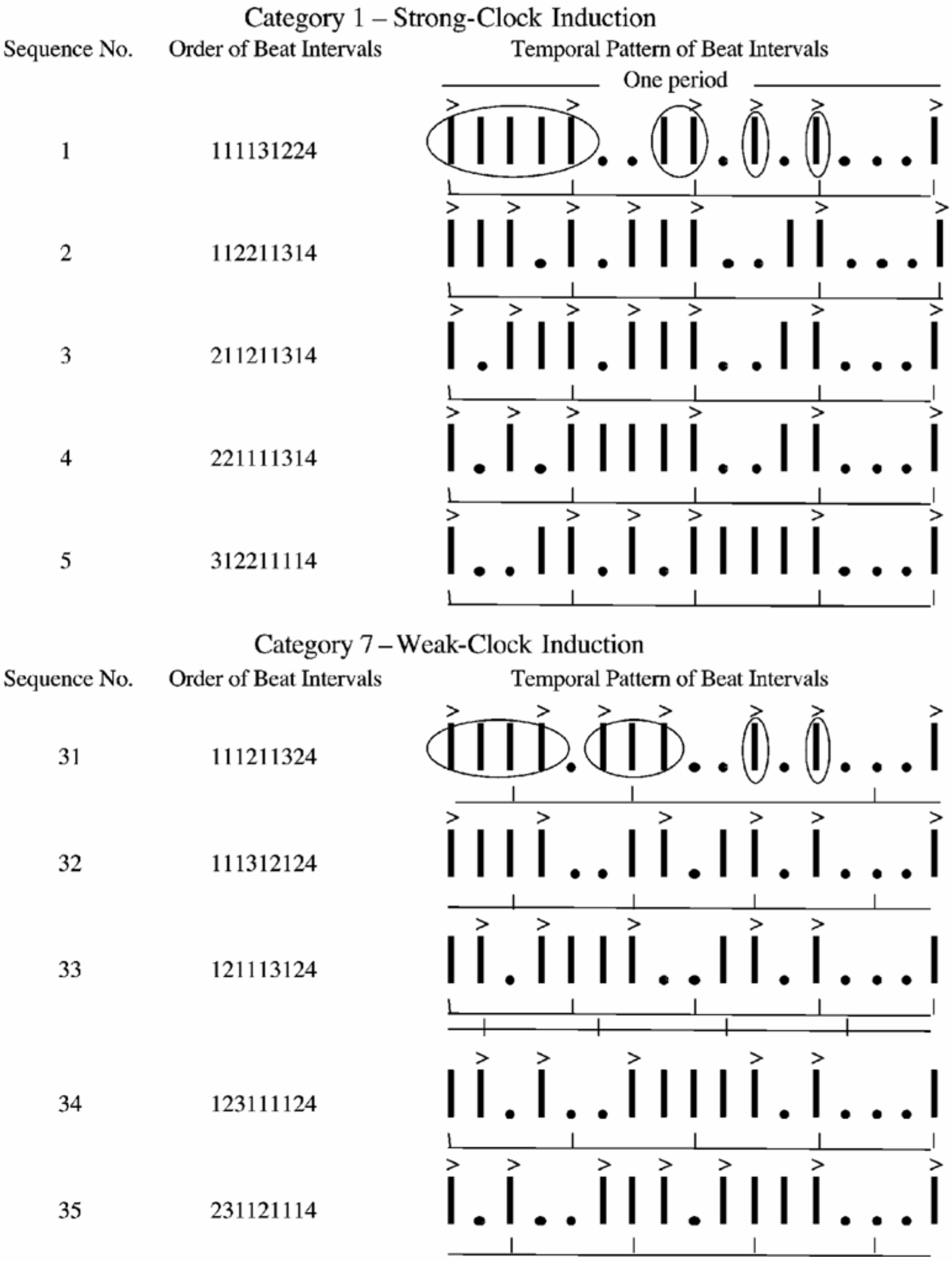

Figure 1. The 10 standard patterns used in the experiments (after Povel & Essens, 1985).

dition and clock strength was not significant [$F(1,15) = 1.43, MS_e = 0.18, p > .25$].

The general clock effect obtained here extends previous work with performance (Povel & Essens, 1985) to a different paradigm—discrimination. It also extends the findings of Ross and Houtsma (1994, Experiment 2) to patterns for which different figural structures are compared. Thus, it may be concluded that metric structure influenced discrimination whether or not a figural strategy could also be employed. An account involving figural coding of the groups of events exclusively (Handel, 1998) is not supported by the present data.

The above conclusion did not appear to be restricted to the musically trained listeners. Correlations between music training points and d' scores were not significant for either disruption-within trials [$r(14) = .35, p > .05$] or disruption-between trials [$r(14) =. 44, p > .05$].

The effect of metric structure and the sensitivity of listeners, in general, to the timing between figures are the most important findings of the experiment. Experiments 2–4

Table 1A
Mean *d′* Scores for Experiment 1

Comparison Condition	Clock Strength Strong	 Weak	 Mean
Disruption within	2.54	1.94	2.24
Disruption between	1.66	0.80	1.23
Mean	2.10	1.37	

pursue further questions regarding the findings for the condition in which figural strategies cannot be employed—the disruption-between condition.

Replication

Before turning to the next experiment, we briefly report the results of a replication of Experiment 1, conducted without feedback. Handel (1998) did not provide feedback, and he suggested that the absence of feedback might be responsible for the lack of discrimination he found between different rhythms with the same figural structures. Because figural structure is a rather prominent feature of temporal patterns, listeners without feedback could have interpreted Handel's (1998) task as one requiring only figural coding. Moreover, they could have inferred (incorrectly) that their strategy was adequate.

Sixteen participants were tested in our replication without feedback. Mean *d′* scores are given in Table 1B, which shows trends similar to those of Table 1A. The disruption-within condition yielded higher scores than did the disruption-between condition [$F(1,15) = 32.40, MS_e = 0.86, p < .001$]. However, although scores were higher overall for strong-clock than for weak-clock patterns, the difference did not reach significance [$F(1,15) = 3.16, MS_e = 0.79, p < .10$]. Scores for the weak-clock, disruption-between condition were not significantly higher than chance [$t(15) = 1.38, p = .19$]. The interaction between comparison condition and clock strength was not significant [$F(1,15) = 1.34, MS_e = 0.18, p > .25$].

Inspection of the replication data and the error terms indicated that listener variability in the replication was greater than that in the main experiment. It may also be noted that scores in the replication experiment were lower than those in the main experiment. An analysis combining both experiments revealed that the difference between the two groups of listeners was not significant [$F(1,30) = 2.29, MS_e = 2.56, p = .14$], nor were any interactions involving experiments significant.

We argue that, not surprisingly, the presence or absence of feedback influences response strategy. With no feedback, response strategies may be less consistent. Thus, for the replication, individual differences were greater, statistical power less, and evidence of a clock effect less reliable. The overall picture across the main experiment and the replication, however, is that the data trends are similar. Thus, absence of feedback alone does not appear to provide a complete account of Handel's (1998) data. We return to this issue at a later point.

EXPERIMENT 2

The primary purpose of Experiment 2 was methodological. In Experiment 1, it appeared that listeners were sensitive to metric structure. However, it may be questioned whether the pick-up of timing information for weak-clock patterns was dependent on or facilitated by a constant presentation rate for all patterns. Possibly, the metric structure of strong-clock patterns instantiated a regular beat that carried over to weak-clock patterns. In Experiment 2, each trial was presented at one of three different presentation rates.

For the three presentation rates, the smallest interval between tones was 150, 200, and 267 msec, respectively. Thus, the time between clock ticks was 600, 800, and 1,068 msec, respectively—values that lie within the estimated range of 200 msec to 1.4 sec for spontaneous or preferred tempos for rhythmic patterns (Handel, 1989, p. 385). The total duration of the standard pattern was 2.4, 3.2, and 4.3 sec, respectively—values that fall within the estimated range of 2–5 sec duration for auditory sensory memory and rhythmic patterns typically found in music (Krumhansl, 2000).

Method

Listeners. The listeners were 12 women and 6 men with a mean age of 18.9 years (range, 17–20). Music training points averaged 4.0 (range, 0–9).

Procedure. The procedure was similar to that in Experiment 1, but the design was altered as follows. Each trial (standard and comparison patterns) was presented at one of three different presentation rates. Interval 1 was set at 150, 200, or 267 msec, respectively, for the three rates, and the remaining intervals at whole-number multiples of Interval 1. The duration of disruption for the disruption-between condition was equal to Interval 1 for each of the three different rates. Thus, the disruption was 150, 200, and 267 msec for the 150-, 200-, and 267-msec presentation rates, respectively.

Only two, rather than three, comparison conditions were included in the design—same and disruption between. There were 60 experimental trials for each listener, resulting from the factorial combination of 10 patterns × 3 presentation rates × 2 comparison conditions. Twelve disruption-within trials (4 disruption-within trials at each presentation rate, 2 for each clock strength) were included to encourage listeners to attend to the entire pattern.[2]

The 72 trials were split into two blocks of 36 trials each. Six different random orders were created for each block, with the constraint that no more than 2 trials with the same standard patterns and no more than 3 trials with the patterns having the same presentation rate could occur in succession.

Results and Discussion

The data were scored and analyzed in the same manner as in the previous experiment. Hits and false alarms were calculated separately for the three presentation rates.

Table 1B
Mean *d′* Scores for Replication: Experiment 1

Comparison Condition	Clock Strength Strong	 Weak	 Mean
Disruption within	2.10	1.83	1.97
Disruption between	0.91	0.39	0.65
Mean	1.51	1.11	

Table 2
Mean d' Scores for Experiment 2

Comparison Condition	Clock Strength		
	Strong	Weak	Mean
150-msec rate	0.46	0.35	0.41
200-msec rate	1.58	0.62	1.10
267-msec rate	1.98	1.07	1.53
Mean	1.34	0.68	

Table 2 displays mean d' scores for the two levels of clock strength and the three presentation rates. Discrimination scores were higher for strong- than for weak-clock patterns [$F(1,17) = 10.29$, $MS_e = 1.14$, $p < .006$]. Discrimination improved as presentation rate slowed [$F(2,34) = 19.81$, $MS_e = 0.58$, $p < .001$]. The interaction was also significant [$F(2,34) = 4.14$, $MS_e = 0.49$, $p < .03$]. Post hoc analyses revealed that the clock effect was significantly smaller at the fastest presentation rate.

The mean d' scores for the weak-clock patterns at the 200- and 267-msec presentation rates, but not at the 150-msec presentation rate, were significantly greater than chance [for the 200-msec rate, $t(17) = 2.14$, $p < .05$; for the 267-msec rate, $t(17) = 4.53$, $p < .001$]. The correlation between listeners' music training points and d' scores was not significant [$r(16) = .04$, n.s.].

The results of this experiment replicated the clock effect found in the first experiment. Moreover, they showed that evidence for the availability of timing information in weak-clock patterns was not restricted to the use of a single presentation rate across trials. An unexpected finding was that discrimination improved as presentation rate slowed. However, the slowing of the presentation rate was accompanied by an increase in the duration of the disruption. Assessing the role of each of the two factors was our motive in the last two experiments.

EXPERIMENT 3

In Experiment 3, we again examined the effect of presentation rate but kept the duration of the disruption constant. If the improvement that occurred in Experiment 2 as the rate slowed was due only to the manipulation of presentation rate, the improvement should be replicated in Experiment 3. If, however, the improvement in Experiment 2 was due to the increase in the absolute duration of the disruption, there should be no effect of presentation rate in Experiment 3.

Method

Listeners. Listeners were 15 women and 3 men with a mean age of 19.2 years (range, 18–21). Music training points averaged 3.3 (range, 0–10).

Procedure. The procedure was similar to that in Experiment 2. The only difference was that the duration of the disruption was 200 msec, constant across presentation rates.

Results and Discussion

The data were scored and analyzed in the same manner as in the previous experiment.

Table 3 displays the mean d' scores for each of the two levels of clock strength and each of the three presentation rates. The clock effect was replicated [$F(1,17) = 21.06$, $MS_e = 1.14$, $p < .001$]. Although they showed a slight increase over 150, 200, and 267 msec, the mean scores for the three different rates were not significantly different [$F(2,34) = 2.03$, $MS_e = 0.79$, $p > .15$]. The interaction between clock and presentation rate did not reach conventional levels of significance [$F(2,34) = 2.91$, $MS_e = 0.41$, $p < .07$].

The mean d' scores for the weak-clock patterns at the 200- and 267-msec presentation rates, but not at the 150-msec presentation rate, were significantly greater than chance [for the 200-msec rate, $t(17) = 2.89$, $p < .01$; for the 267-msec rate, $t(17) = 4.41$, $p < .005$]. Finally, the correlation between listeners' music training points and d' score was not significant [$r(16) = .29$, $p > .05$].

The results show that for a constant duration of the disruption, the effect of presentation rate was greatly weakened. Weak-clock patterns tended to benefit more than strong-clock patterns from the slowing of presentation rate, but not to a significant degree. Thus, the absolute duration of the disruption, not presentation rate, appeared to be the main factor governing discrimination accuracy. In other words, the relative duration of the disruption—that is, the percentage of the duration of the disruption, relative to the duration of Interval 1—cannot be invoked to account for the data in both Experiments 2 and 3. In Experiment 2, the percentage for the three presentation rates was constant at 100%; in Experiment 3, it was 125%, 100%, and 75%, for the 150-, 200-, and 267-msec rates, respectively.

EXPERIMENT 4

In Experiment 4, the duration of the disruption was set in opposition to the presentation rate. The duration of the disruption was reversed for the fast and slow presentation rates so that the disruption was shorter as the rate slowed. If absolute duration of the disruption governs discrimination accuracy, performance should be best for the fast presentation rate. In other words, the direction of the effect of presentation rate found in Experiment 2 should be reversed.

Method

Listeners. The listeners were 8 women and 10 men with a mean age of 20.6 years (range, 18–35). Music training points averaged 2.7 (range, 0–8).

Table 3
Mean d' Scores for Experiment 3

Comparison Condition	Clock Strength		
	Strong	Weak	Mean
150-msec rate	1.69	0.34	1.02
200-msec rate	1.64	0.98	1.31
267-msec rate	1.83	1.01	1.42
Mean	1.72	0.78	

Table 4
Mean *d′* Scores for Experiment 4

	Clock Strength		
Comparison Condition	Strong	Weak	Mean
150-msec rate	2.15	1.46	1.81
200-msec rate	2.13	1.18	1.66
267-msec rate	1.32	0.38	0.85
Mean	1.87	1.01	

Procedure. The procedure was similar to that in Experiment 2, except that the duration of the disruption for the disruption-between condition was reversed for the fast and slow presentation rates. Thus, the disruption was 267 and 150 msec for the 150- and 267-msec rates, respectively. The disruption was left unchanged at 200 msec for the 200-msec rate.

Results and Discussion

Table 4 displays mean *d′* scores for the two levels of clock strength and each of the three presentation rates. The effect of presentation rate was significant [$F(2,34) = 13.93$, $MS_e = 0.51$, $p < .001$], but in this experiment, performance worsened as the rate slowed (and the duration of the disruption decreased). The analysis again yielded a significant effect of clock [$F(1,17) = 17.14$, $MS_e = 1.18$, $p < .001$]. The interaction between clock and presentation rate was not significant ($F < 1$).

The mean *d′* scores for the weak-clock patterns at the 150- and 200-msec presentation rates, but not at the 267-msec presentation rate, were significantly greater than chance [for the 150-msec rate, $t(17) = 4.78$, $p < .001$; for the 200-msec rate, $t(17) = 4.22$, $p < .001$]. Finally, the correlation between listeners' music training points and *d′* scores was not significant [$r(16) = -.16$, $p > .05$].

Further analyses were conducted to verify the reliability of the findings regarding presentation rate and absolute duration of the disruption across experiments. First, a standardized value was obtained for the one condition that was constant across Experiments 1–4, the 200-msec rate with a 200-msec disruption. No difference in the mean values across the four experiments was found ($F < 1$, $MS_e = 2.27$). The overall mean *d′* score for this condition, across levels of clock strength, was 1.33. For purposes of visual display (Figure 2) all the results were adjusted with respect to the overall mean, thus eliminating (nonsignificant) level differences among the experiments. Figure 2 displays this adjusted *d′* for each one of the three disruption durations for each of the three presentation rates averaged across clocks. There is a monotonic relationship between discrimination and disruption duration, showing that discrimination improved as the absolute duration of the disruption increased.

Also, Figure 2 shows how the data obtained by Ross and Houtsma (1994) and Handel (1998) fit neatly with the projected relationship. Ross and Houtsma employed a disruption of 250 msec. The obtained *d′* of 1.55 (averaged across clock strength) lies between our estimates for 200

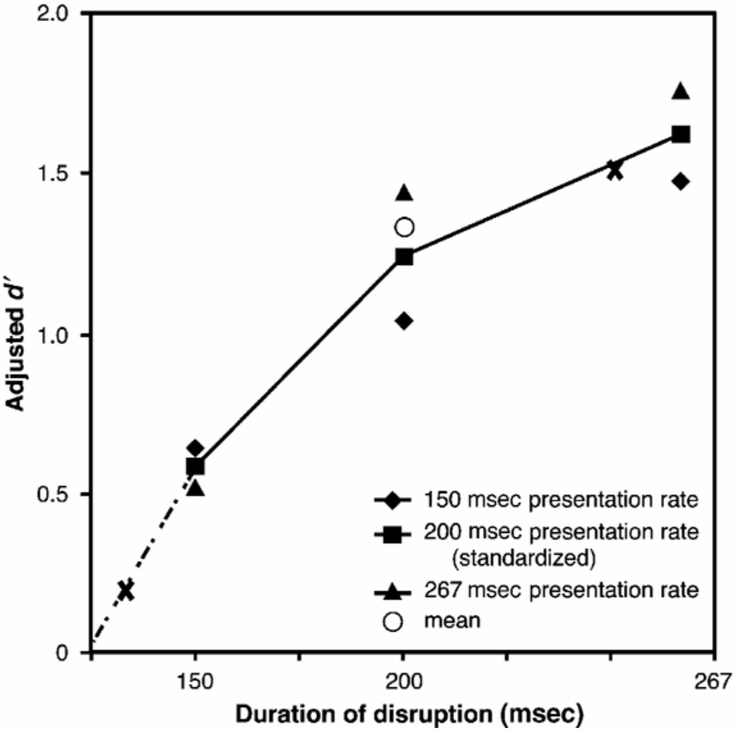

Figure 2. Adjusted *d′* values for each of the three disruption durations for each of the three presentation rates, averaged across clocks. The X mark at 250 msec represents Ross and Houtsma's (1994) actual *d′* value. The X mark at 133 msec represents Handel's (1998) projected *d′* value (at chance).

Table 5
A Summary of the Multiple Regression Analysis With Clock, Presentation Rate, and Absolute Duration of Disruption as Predictors

Predictor	β	t	p
Clock	.702	6.20	.001
Presentation rate	−.091	−0.80	.433
Duration of disruption	.548	4.82	.001

Note—Overall model, $R = .89$; $F(3,16) = 20.63, p < .001$.

and 267 msec. Handel's (1998) procedure used intervals of 133 msec; disruption involved the shifting of one tone by one interval. The chance performance that Handel (1998) obtained for disruption between (estimated in Figure 2 as $d' = 0.0$) is quite reasonably in line with the relationship.

A multiple regression was conducted to test three predictor variables (clock strength, presentation rate, and duration of disruption) against discrimination performance across experiments. There were 20 values for each variable, 2 from Experiment 1 (disruption between, strong and weak clock), and 6 from each of Experiments 2, 3, and 4. Experiment (rather than subjects) thus becomes the new random factor. The regression was conducted on both the original and the adjusted d' (Figure 2) with similar outcomes. Table 5 presents the results of the regression analysis on the original. Two predictors, clock strength and duration of disruption, yielded a significant contribution; presentation rate did not.

Experiments 2–4, therefore, confirmed that the effect of clock strength holds across variations in presentation rate. In addition, timing information is preserved between figures even if clock strength is weak. An unexpected finding was that discrimination appears to reference absolute time, not relative time, such as the beat rate established by the clock.

Finally, in order to increase the power of the correlation between music training points and performance, we ran an overall correlation for all the experiments. The correlations were non significant, with $r(68) = .06, p = .62$ and $r(68) = .15, p = .23$ for strong-clock and weak-clock patterns, respectively.

GENERAL DISCUSSION

This series of experiments dealt with how listeners discriminate between two temporal patterns differing by the insertion of an additional silent gap. We will summarize three major issues. First, we addressed the question of whether discrimination involved only one perceptual structure—figural grouping—or a second structure based on metric timing. Consistent with the existing literature, patterns with different figural structures were more easily discriminated than patterns with identical figural structures (Handel, 1998; Ross & Houtsma, 1994, Experiment 1). In addition, patterns assumed to induce a strong meter were easier to discriminate than nonmetric or weakly metric patterns. This latter finding is consistent with previous results involving pattern discrimination (Ross & Houtsma, 1994, Experiment 2) and pattern reproduction (Essens, 1986; Povel & Essens, 1985). It is not consistent, however, with findings by Handel (1998), who reported little or no role of metric structure in temporal discrimination.

Next, we addressed the discrepancy between our results and those of Handel (1998). One difference between our procedure and that of Handel (1998) was that we provided feedback to the listeners, whereas he did not. In a replication study, in which feedback was withdrawn, we found greater variability in listener performance but found the same trends in the data implicating metric structure. In the light of our findings in Experiments 2–4, we suggest that a more likely account of the discrepancy between our data and Handel's (1998) involves differences in the absolute size of the silent gap to be detected. The size of the gap in Handel's (1998) task may have been subthreshold for precise timing of the gap to be processed (see Figure 2). By *subthreshold*, we do not refer to threshold values for discrimination between two durations (values that Handel, 1998, clearly exceeded). Rather, the threshold referred to here is timing within a temporal pattern.

Third, Experiments 2–4 reveal the operation of an absolute timing mechanism in addition to the structure-dependent mechanism of meter. The clock effect, although present in all the experiments, could not account for how pattern discrimination varied with the duration of the additional silent gap under different presentation rates. According to the definition of the internal clock, clock ticks speed up or slow down with an increase or decrease in presentation rate. Therefore, a disruption equal in duration to the presentation rate occupies a constant proportion of the time between clock ticks. In other words, independent of presentation rate, the relative disruption is constant. Discrimination under such conditions, however, was not constant (Experiment 2). Discrimination varied with the absolute duration of the disruption and approached a constant value when the change was constant in absolute time (Experiment 3).

This new finding fits nicely with recent findings that memory for musical features has absolute, as well as relative, components. Music has always been regarded as a highly relational domain with respect to pitch, time, and timbre information. Listeners were thought to encode only the relative pitch (intervals) and time (relative duration values) information of a musical pattern, and this long-term mental code would not even retain timbre (instrument on which the musical pattern was played). Even though it is undoubtedly the case that relative information about musical tunes can be retrieved, recent work has shown that long-term auditory memory for musical recordings also contains absolute information about pitch (Levitin, 1994), tempo (Levitin & Cook, 1996), and timbre (Schellenberg, Iverson, & McKinnon, 1999). The results of our study, showing the presence of an absolute code for timing information, suggest a perceptual substrate for the absolute coding of musical tempo (Levitin & Cook, 1996).

The notion that there is an absolute timekeeper for musical patterns has been supported by Clarke and Krumhansl

(1990). Their study obtained indirect and direct estimates of the duration of musical segments varying in structural qualities, such as musical complexity and completeness. Listeners (performing musicians) produced veridical time estimates unaffected by the structural quality of the excerpt. Clarke and Krumhansl speculated that this sense of absolute time might be specific to the highly trained population sampled in their experiments. They noted that performing musicians have learned to maintain a steady tempo despite variations in the structural quality of the music. Our data suggest that the sense of absolute time may be invoked for pattern discrimination by listeners without specialized training.

Nonmusical tasks also provide support in favor of the encoding of absolute time by untrained listeners. Ivry and Hazeltine (1995) found evidence for a common, and absolute, timing mechanism operative in both the perception and the production (i.e., tapping) of temporal intervals ranging from 325 to 500 msec. The tasks involved either two events or a series of events. Therefore, the absolute timing mechanism may be invoked in a variety of temporal tasks.

The absolute clock can also be considered from a biological perspective. Two subcortical structures have been suggested to play a critical role in the timing of both movement and perception—that is, the cerebellum and the basal ganglia (Ivry, 1996). We recently documented the case of a dyslexic adult who had a deficit in using the absolute clock in a task involving discrimination of simple temporal intervals shorter than 1 sec (Rousseau, Hébert, & Cuddy, 2001). Other recent evidence has also shown that a deficit in (absolute) time estimation is observable in dyslexia, for tasks highly sensitive to cerebellar involvement (Nicolson, Fawcett, & Dean, 1995). Failure to engage absolute clock timing could, therefore, be associated with developmental dyslexia. The role of the absolute clock in such functional deficits could be examined more closely in future studies.

In summary, discrimination of temporal patterns put into play several types of mechanisms. One such mechanism relies on how close in temporal proximity the tones are. Sensitivity to figural structure is immediate and easy and is present in 3-month-old infants (Demany et al., 1977). Another mechanism involves timing that is dependent on the instantiation of an internal clock, or a sense of metric structure. It is likely that sensitivity to meter is acquired implicitly through exposure to music (e.g., Drake, 1998), as is argued to be the case for sensitivity to tonality (Krumhansl, 1990; Tillmann, Bharucha, & Bigand, 2000). Formal musical training is not required. Furthermore, sensitivity to the hierarchical structure of meter may involve different brain mechanisms from those involved with figural grouping (Liégeois-Chauvel, Peretz, Babaï, Laguitton, & Chauvel, 1998; Peretz, 1990) or nonhierarchical structures (Sakai et al., 1999). Finally, a separate context-independent mechanism involves a real-time clock, measuring absolute time. Both metric and absolute mechanisms can co-occur, such as with strongly metric patterns. Where context does not favor a strong internal meter, the context-independent timing mechanism remains an available resource for discrimination.

REFERENCES

Bamberger, J. (1982). Growing up prodigies: The midlife crisis. *New Directions for Child Development*, **17**, 61-77.

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

Boltz, M. (1989a). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 749-761.

Boltz, M. (1989b). Rhythm and "good endings": Effects of temporal structure on tonality judgments. *Perception & Psychophysics*, **46**, 9-17.

Clarke, E. F., & Krumhansl, C. L. (1990). Perceiving musical time. *Music Perception*, **7**, 213-252.

Dawe, L. A., Platt, J. R., & Racine, R. J. (1993). Harmonic accents in inference of metrical structure and perception of rhythm patterns. *Perception & Psychophysics*, **54**, 794-807.

Dawe, L. A., Platt, J. R., & Racine, R. J. (1994). Inference of metrical structure from perception of iterative pulses within time spans defined by chord change. *Music Perception*, **12**, 57-76.

Demany, L., McKenzie, B. A., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature*, **266**, 718-719.

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Perception & Psychophysics*, **53**, 25-33.

Drake, C. (1998). Psychological processes involved in the temporal organization of complex auditory sequences: Universal and acquired processes. *Music Perception*, **16**, 11-26.

Drake, C., & Gérard, C. (1989). A psychological pulse train: How young children use this cognitive framework to structure simple rhythms. *Psychological Research*, **51**, 16-22.

Essens, P. J. (1986). Hierarchical organization of temporal patterns. *Perception & Psychophysics*, **40**, 69-73.

Essens, P. [J]. (1995). Structuring temporal sequences: Comparison of models and factors of complexity. *Perception & Psychophysics*, **57**, 519-532.

Essens, P. J., & Povel, D. J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception & Psychophysics*, **37**, 1-7.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.

Handel, S. (1992). The differentiation of rhythmic structure. *Perception & Psychophysics*, **52**, 497-507.

Handel, S. (1998). The interplay between metric and figural rhythmic organization. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1546-1561.

Hirsh, I. J., Monahan, C. B., Grant, K. W., & Singh, P. G. (1990). Studies in auditory timing: 1. Simple patterns. *Perception & Psychophysics*, **47**, 215-226.

Ivry, R. B. (1996). The representation of temporal information in perception and in motor control. *Current Opinion in Neurobiology*, **6**, 851-857.

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

Jones, M. R., & Ralston, J. T. (1991). Some influences of accent structure on melody recognition. *Memory & Cognition*, **19**, 8-20.

Jones, M. R., Summerell, L., & Marshburn, E. (1987). Recognizing melodies: A dynamic interpretation. *Quarterly Journal of Experimental Psychology*, **39A**, 89-121.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace.

Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology* (rev. ed.). New York: Liveright.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* Oxford: Oxford University Press.

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, **126**, 159-179.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, **56**, 414-423.

Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, **58**, 927-935.

Liégeois-Chauvel, C., Peretz, I., Babaï, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, **121**, 1853-1867.

MacMillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, **98**, 185-199.

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

Nicolson, R. I., Fawcett, A. J., & Dean, P. (1995). Time estimation deficits in developmental dyslexia: Evidence of cerebellar involvement. *Proceedings of the Royal Society of London: Series B*, **259**, 43-47.

Palmer, C., & Krumhansl, C. L. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 116-126.

Palmer, C., & Krumhansl, C. L. (1987b). Pitch and temporal contributions to musical phase perception: Effects of harmony, performance timing, and familiarity. *Perception & Psychophysics*, **41**, 505-518.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 728-741.

Peretz, I. (1990). Processing of local and global musical information in unilateral brain-damaged patients. *Brain*, **113**, 1185-1205.

Peretz, I., & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: A neuropsychological perspective. *Quarterly Journal of Experimental Psychology*, **46A**, 301-325.

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

Povel, D. J., & Okkerman, H. (1981). Accents in equitone sequences. *Perception & Psychophysics*, **30**, 565-572.

Ross, J., & Houtsma, A. J. M. (1994). Discrimination of auditory temporal patterns. *Perception & Psychophysics*, **56**, 19-26.

Rousseau, L., Hébert, S., & Cuddy, L. L. (2001). Impaired short temporal interval discrimination in a dyslexic adult. *Brain & Cognition*, **46**, 249-254.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada, T., Iwata, N. K., & Nielsen, M. (1999). Neural representation of a rhythm depends on its interval ratio. *Journal of Neuroscience*, **19**, 10074-10081.

Schellenberg, E. G., Iverson, P., & McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review*, **6**, 641-646.

Smith, K. C., & Cuddy, L. L. (1989). Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 457-471.

Smith, K. C., Cuddy, L. L., & Upitis, R. (1994). Figural and metric understanding of rhythm. *Psychology of Music*, **22**, 117-135.

Thorpe, L. A., & Trehub, S. E. (1989). Duration illusion and auditory grouping in infancy. *Developmental Psychology*, **25**, 122-127.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, **107**, 885-913.

Todd, R. E., Boltz, M., & Jones, M. R. (1989). The MIDILAB research system. *Psychomusicology*, **8**, 83-96.

NOTES

1. Patterns were repeated once—that is, the onset from the first tone to the first tone of the repetition was 16 intervals. The total interval duration of each pattern was divisible by 4, not by 3, thus facilitating the four-beat clock (Povel & Essens, 1985).

2. Detection of disruptions within figures, in this and subsequent experiments, was near ceiling. Thus, the insertion of disruption-within trials fulfilled its purpose—to verify that listeners were attending to the entire pattern.

APPENDIX

The alternative analysis was motivated by a study by Palmer and Krumhansl (1990), who showed, for a sample of Western-harmonic compositional styles, that the notated meter (i.e., 2/4, 3/4, 4/4, and 6/8) was supported by the temporal distribution of events in the music. The authors argued that statistical regularity—here, the frequency of occurrence of events at different points of time—is a cue to perceptual structure. As well, statistical regularity may become internalized as a mental representation, guiding future expectation of events.

Palmer and Krumhansl (1990) divided musical bars into temporal locations representing the smallest subdivision of durations. Next, the frequency of events at each location was tallied. The frequency distribution across locations showed a regular periodicity that depended on the meter. These periodic components corresponded to a music-theoretic analysis of the relative metric strength of successive temporal events within a bar (Lerdahl & Jackendoff, 1983).

The patterns in the present experiment (Figure 1) may be treated as four beats of a 4/4 bar with 16 subdivisions. According to the frequency distributions of Palmer and Krumhansl (1990), an event may be strongly expected to occur on the first position, followed by the fifth, the ninth, and then the thirteenth positions—the musical beats within the bar. Events between the beats are expected less, although the half-beat, represented by positions three, seven, eleven, and fifteen, are expected more than the remaining subdivisions of the beat. According to Lerdahl and Jackendoff (1983), the weights for metric accent strength of the 16 positions are 5–1–2–1–3–1–2–1–4–1–2–1–3–1–2–1.

It can be shown that each pattern categorized by Povel and Essens (1985) as a strong clock has an event (tone) on each of the four musical beats, as defined above, and consistently fewer events on each of the subdivisions of the beats. Tallied across patterns, the frequency distribution is 5–2–4–3–5–1–4–4–5–1–2–4–5–0–0–0. The correlation between the frequency distribution and the music-theoretic prediction is significant [$r(14) = .63, p < .01$]. Each pattern classified as a weak clock does not consistently have an event on all the musical beats, nor are there consistently fewer events on each of the subdivisions of the beats. Tallied across patterns, the frequency distribution is 5–4–3–4–1–3–5–4–1–4–5–1–5–0–0–0. The correlation between the frequency distribution and the music-theoretic prediction is not significant [$r(14) = .16$]. Thus, the distribution of events supports either a strong or a weak clock in a manner consistent with Povel and Essens.

It is beyond the focus of this paper to explore distinctions between various proposals for the extraction of meter from a sound sequence. It is sufficient to note that the coding model of Povel and Essens (1985), the representational account of Palmer and Krumhansl (1990), and the music-theoretic analysis of Lerdahl and Jackendoff (1983) converge on the validity of the concept of clock strength.

(Manuscript received March 27, 1998;
revision accepted for publication October 30, 2001.)