Journal of Motor Behavior, 2003, Vol. 35, No. 4, 355–370

1320

Rate Limits in Sensorimotor Synchronization With Auditory and Visual Sequences: The Synchronization Threshold and the Benefits and Costs of Interval Subdivision

Bruno H. Repp
Haskins Laboratories
New Haven, CT

ABSTRACT. Synchronization of finger taps with an isochronous event sequence becomes difficult when the event rate exceeds a certain limit. In Experiment 1, the synchronization threshold was reached at interonset intervals (IOIs) above 100 ms with auditory tone sequences (in a 1:4 tapping task) but at IOIs above 400 ms with visual flash sequences (1:1 tapping). Using IOIs above those limits, the author investigated in Experiment 2 the reduction in the variability of asynchronies that tends to occur when the intervals between target events are subdivided by additional identical events (1:1 vs. 1:*n* tapping). The subdivision benefit was found to decrease with IOI duration and to turn into a cost at IOIs of 200–250 ms in auditory sequences and at IOIs of 450–500 ms in visual sequences. The auditory results are relevant to the limits of metrical subdivision and beat rate in music. The visual results demonstrate the remarkably weak rhythmicity of (nonmoving) visual stimuli.

Key words: rate limits, subdivision, synchronization, tapping

It is generally recognized that vision is superior to audition in spatial resolution but is poorer in temporal resolution. For example, in studies of temporal discrimination, better performance has consistently been found in audition than in vision (e.g., Goodfellow, 1934; Grondin, 1993; Grondin, Meilleur-Wells, Ouellette, & Macar, 1998; Grondin, Ouellet, & Roussel, 2001; Grondin & Rousseau, 1991; Repp & Penel, 2002; Rousseau, Poirier, & Lemyre, 1983). Moreover, in studies of sensorimotor synchronization (finger tapping), the variability of asynchronies has been found to be greater in synchronization with isochronous visual sequences, typically composed of light flashes, than in synchronization with temporally matched auditory sequences, typically composed of clicks or tones (Bartlett & Bartlett, 1959; Dunlap, 1910; Fraisse, 1948; Klemmer, 1967; Kolers & Brewster, 1985; Repp & Penel, 2002).

One question of considerable theoretical and methodological interest is how fast an event sequence can be in either modality before synchronization of an action with (selected) sequence events becomes impossible. In the auditory modality, the rate limit, which is referred to herein as the *synchronization threshold,* reflects a perceptual or sensorimotor constraint that also governs the smallest metrical subdivision in music (London, 2002), which is sometimes called the "tatum" (Bilmes, 1993).[1] The tatum plays an important role in the perception and production of complex rhythms, particularly in African and Afro-American music (Iyer, 2002). Relying mainly on their musical intuitions and experience, but also on empirical studies, both London and Iyer estimated the limit to be in the vicinity of 100 ms. More generally, the synchronization threshold reveals rate limits of the internal timekeeping and coordinative motor processes that enable a person to track event sequences and carry out anticipatory actions (e.g., Large, 2000; Large & Jones, 1999). In comparing auditory and visual modalities, the synchronization threshold provides an interesting measure of perceptual temporal resolution in the service of action.

Given the elementary nature of the question, it is surprising how few relevant studies there are in the literature. Neither the auditory nor the visual synchronization threshold ever seems to have been assessed precisely. In one of the earliest studies on the subject, Dunlap (1910, p. 404) mentioned that "even a 1/3 [s] rate proved too fast for visual stimulation" in a synchronization task. At least 1 participant, however, successfully synchronized with auditory click sequences having interonset intervals (IOIs) as short as 250 ms. Almost half a century later, Bartlett and Bartlett (1959) reinvestigated the limits of synchronization. To

---

Correspondence address: Bruno H. Repp, Haskins Laboratories, 270 Crown Street, New Haven, CT 06511-6695, USA. E-mail address: repp@haskins.yale.edu

avoid running into motor limits, especially when tapping to auditory sequences, they asked participants to make only a single tap to each sequence. The tap closed a switch that terminated the sequence, and participants were instructed to make the tap coincide with any sequence event. Five highly trained participants were tested with two ranges of IOIs. With auditory click sequences, 3 participants performed at chance level with IOIs of 125 ms (i.e., they produced a random distribution of relative phases between taps and tones across a number of trials). The other 2 participants performed better than chance with IOIs of 167 ms, which apparently were the shortest IOIs they attempted. Only 1 of the 5 had difficulties with IOIs of 250 ms. With visual flash sequences, however, all participants performed at chance level with IOIs of 250 ms or less. One of 2 participants who were presented with 333-ms IOIs also failed. All did well with 500-ms IOIs, although variability was always higher than that found with auditory sequences of the same rate.

A serial reaction-time study by Klemmer (1967) is also pertinent. He instructed participants to depress a key quickly in response to each event in isochronous auditory or visual sequences. At short IOIs, that procedure automatically led to synchronization of responses with stimuli. Klemmer found, however, that participants were unable to maintain such synchrony with visual flash sequences having IOIs of 300 ms and were able to do so in only 50% of the trials when the IOIs were 500 ms. By contrast, with auditory tone sequences, performance was synchronous in 50% of the trials when the IOIs were 300 ms and in all of the trials with IOIs of 500 ms.

Although it might seem that the stationarity of the visual stimuli might have been the problem, comparable limits have been observed in tasks that required synchronization with moving visual stimuli. For example, Noble, Fitts, and Warren (1955) found that highly practiced participants' synchronization in a discrete visual-motor pursuit tracking task broke down at IOIs of 333 or 500 ms. More recently, Wimmers, Beek, and van Wieringen (1992, Experiment 3) reported that in-phase synchronization was lost at rates of 1.7 to 2.3 Hz (IOIs of 435 to 588 ms) in a task that required synchronization of arm movements with an oscillating visual signal. (For similar results, see Byblow, Chua, & Goodman, 1995; Peper & Beek, 1998; Repp & Penel, 2002.) To what extent certain forms of visual motion facilitate synchronization is an interesting question, but it was not addressed in the present study, in which the focus was on nonmoving visual flash stimuli of the kind used in the studies cited earlier.

The results of those studies suggested that the synchronization threshold for auditory sequences, in terms of IOI duration, might be as low as 125 ms and as high as 300 ms, perhaps depending on participants' rhythmic skills, whereas the threshold for visual sequences might be as low as 333 ms and as high as 500 ms. A clear modality difference is already evident, but one would need a more detailed investigation using a larger number of closely spaced IOIs to estimate synchronization thresholds more precisely. That was my purpose in Experiment 1.

In Experiment 2, a closely related but likewise little investigated phenomenon that occurs at IOIs above the synchronization threshold, at least with auditory sequences, was examined. When participants tap, in one condition, with every event of an isochronous auditory sequence (1:1 tapping) and, in another condition, with every other event of a sequence that is twice as fast (1:2 tapping), their asynchronies tend to be less variable in the 1:2 condition. The reduction in variability is referred to herein as a (physical) *subdivision benefit.*[2] The interval between the prescribed target events (every other event in the 1:2 example) is referred to as the *interbeat interval* (IBI). The IBI corresponds to the average intertap interval (ITI). It is well known that variability of asynchronies and ITIs increases with ITI duration, both in free and in synchronized tapping (Madison, 2001; Peters, 1989; Semjen, Schulze, & Vorberg, 2000). That increase has been attributed to the scalar variability of a central timekeeper (Gibbon, Church, & Meck, 1984). In the situation just described, however, the average ITI is the same in 1:1 and in 1:2 tapping; therefore, the subdivision benefit cannot derive from ITI duration. Its presumable cause is that in 1:2 tapping, the fast event rate entrains an internal timekeeper or oscillator that is twice as fast as the one entrained in 1:1 tapping at the same IBI. Rather than being launched in every cycle of the timekeeper, taps are alternately executed and withheld in 1:2 tapping. The subdivision benefit thus reflects the lower variability associated with a faster internal periodicity.

Semjen, Schulze, and Vorberg (1992) demonstrated the subdivision benefit with auditory sequences; they found it to be rate dependent. For an IBI of 600 ms, the variability of asynchronies and ITIs was lower in 1:2 tapping (IOI = 300 ms) than in 1:1 tapping (IOI = 600 ms), but for an IBI of 400 ms, there was no subdivision benefit. That finding suggests that in 1:2 tapping, the benefit disappeared somewhere between 300 and 200 ms of IOI duration, and it implies that there is a rate limit to the internal periodicity that is entrained by an auditory sequence. Like the synchronization threshold, the limit seems relevant to metrical subdivision in music as well as to modality differences in temporal processing. The results of Semjen and his colleagues suggest that the auditory subdivision benefit disappears at an IOI duration that is well above the synchronization threshold. That duration might be related to a second limit of metrical structure identified by London (2002), namely, the fastest possible beat rate. London hypothesized that a viable rhythmic beat implies the possibility of at least simple (i.e., two beats to a measure, or duple) subdivision, and he estimated the fastest possible beat rate to be around 250 ms. London's estimate agrees roughly not only with the findings of empirical studies of beat rate and beat induction (e.g., Parncutt, 1994; van Noorden & Moelants, 1999) but also with the results of Semjen et al. (1992). Indeed, one should expect the limit to be at about twice the synchronization threshold if the same basic rate limit prevents synchronization and a benefit of subdivision. That issue was investigated further in Experiment 2, in both the auditory and visual modalities.

EXPERIMENT 1

The Synchronization Threshold

My purpose in Experiment 1 was to obtain estimates of synchronization thresholds by varying IOI duration in small steps within an appropriate range in each modality. Because the auditory synchronization threshold was expected to be at an event rate faster than the maximal possible tapping rate (about six–eight beats/s), participants were required to tap with every fourth tone in auditory sequences (1:4 tapping). In that way, tones could be presented at fast rates while the tapping rate was kept within a comfortable range. It was assumed that the synchronization threshold reflects primarily perceptual processes required for the tracking of sequence events and not the specific response requirements (cf. Bartlett & Bartlett, 1959, who required only a single tap per sequence). In any case, 1:4 tapping was believed to be the optimal response requirement for synchronization with fast auditory sequences because it corresponds to a common (quadruple) musical meter and brings the beat close to its preferred rate of about two beats/s (Parncutt, 1994; van Noorden & Moelants, 1999). With visual sequences, the event rates were slow enough so that participants could tap comfortably with every light flash (1:1 tapping). In a third condition, slow auditory sequences matching the visual sequences were presented (1:1 tapping) so that variability could be compared between the modalities for identical tasks. The slow auditory sequences also served to familiarize novice participants with the tapping task.

Method

Participants

The participants included 12 paid volunteers (9 women, 3 men) and me. (I routinely run myself first in every experiment.) Five individuals (including me) had moderate to extensive experience in auditory synchronization tasks, and 4 of them had previously participated in Experiment 2, which preceded Experiment 1 chronologically. The others were novices who had responded to an advertisement that solicited individuals with "a good sense of rhythm" and a willingness to participate in finger-tapping experiments on a weekly basis.[3] Only I had some previous experience with visual synchronization. Musical experience ranged all the way from professional training to no training at all, although the majority of the participants had substantial musical training (see Table 1). Ages ranged from 20 to 27 years, except for one novice and me, who were 56 years old. All were right-handed.

Materials

There were three conditions: slow auditory, fast auditory, and visual. The IOIs (= IBIs) in the slow auditory and visual conditions ranged from 320 to 680 ms in steps of 40 ms; thus, there were 10 different isochronous sequences. The sequences in the fast auditory condition started with the same IOIs (indicating the beat rate) but, from the seventh tone on, continued at a rate 4 times as fast; the IOIs therefore ranged from 80 to 170 ms in 10-ms steps, with the IBIs being 4 times as long. Each sequence contained 56 target events (separated by IBIs) that were defined only by their temporal position. The 10 sequences in each condition were presented in 10 different random orders (trial blocks).

TABLE 1. Individual Synchronization Thresholds (IOI duration in ms) in the Fast Auditory and Visual Conditions of Experiment 1 and Participants' Tapping Experience and Musical Training

Participant	Auditory	Visual	Tapping	Musical training
A.M.	158	500	Novice	Amateur
A.S.	123	456	Experienced	Limited
A.C.	> 170	480	Novice	None
B.K.	110	460	Novice	Semi-professional
E.W.	137	540	Experienced	Amateur
B.S.	105	507	Some experience	Active amateur
B.R.	115	413	Experienced	Active amateur
H.R.	157	420	Novice	Amateur
J.S.	125	450	Novice	Amateur
K.W.	148	451	Novice	Amateur
R.F.	105	~432	Novice	Professional
S.V.	97	432	Some experience	Active amateur
V.T.	100	457	Novice	Amateur

Note. Individuals designated as amateur had 6 or more years of musical training.

All sequences were instantiated as files of musical instrument digital interface (MIDI) instructions and were played by a program written in MAX Version 3.0 running on a Macintosh Quadra 660AV computer, which also collected and saved the tapping data.[4] Auditory sequences were composed of high-pitched (4,186 Hz) piano tones ("pings") that had sharp attacks and decayed rapidly within about 100 ms. The tones came from a Roland RD-250s digital piano and were heard over Sennheiser HD540 II headphones. Visual sequences consisted of discrete blinks of a green light-emitting diode (LED) 3 mm in diameter. The LED was the message indicator of an Opcode MIDI Translator II box, and it lit up whenever a "note on" message was received from the computer. (There were no "note off" messages in the instructions.) Successive flashes were clearly visible as discrete events in a darkened room.

Procedure

Participants came for three 1-hr sessions in the following order: slow auditory, fast auditory, and visual. They sat in front of a computer monitor and tapped with the index finger (1 participant used the middle finger) of their preferred hand on a Roland SPD-6 percussion pad that they held on their lap. Some rested their wrist and other fingers on the surface of the pad and tapped by moving the finger only; others preferred to tap from above by moving either or both the wrist and the elbow joints of the free arm. The impact of the finger on the rubber pad provided some direct auditory feedback (a thud), whose intensity depended on the individual tapping force. The digital sound output from the percussion pad was disconnected.

During the visual session, participants did not wear headphones, and I darkened the room by extinguishing all lights, dimming the computer screen, and covering it with a cardboard flap, leaving only a small gap at the bottom from which a dim glow emanated. The box containing the LED was placed on the rear of the computer keyboard, about 50 cm from the participant's eyes. Between blocks, I uncovered and brightened the computer screen to save the data.

The sessions were self-paced: Participants started a block by clicking a button on the computer screen with the mouse, and each subsequent sequence started 3 s after the space bar on the computer keyboard was depressed. Participants were instructed to start tapping with the third event in each sequence and to continue tapping in synchrony with the sequence events. In the fast auditory condition, their task was to continue tapping at the initial tempo of the sequence while trying to stay in synchrony with the target events (i.e., every fourth tone).

Data Editing and Analysis

I inspected the data to identify occasional missing or extra taps and to make appropriate corrections to the asynchronies and ITIs. A true inability to synchronize was indicated by unidirectional phase drift (i.e., steadily increasing or decreasing asynchronies), which was common in the fast auditory and visual conditions but absent in the slow auditory condition. Unsuccessful trials in the fast auditory condition were defined as those in which the asynchronies had a standard deviation exceeding 10.0% of the IBI (40.0% of the IOI). In the visual condition, because the variability was much higher, the criterion was relaxed to 16.7% of the IOI (= IBI). The exact choice of criterion was not considered crucial because trials with phase drift generally had standard deviations that far exceeded any reasonable criterion. Only a few trials were ambiguous—for example, when phase drift started near the end of the sequence. Unsuccessful trials were excluded from analyses of asynchronies but not from analyses of ITIs. Asynchronies and ITIs were computed from the 6th tap onward and thus were based on 48 taps per trial.

Results

Synchronization Threshold Estimates

The average percentages of successful trials (as just defined) as a function of IOI duration in the fast auditory condition are shown in Figure 1A. (The corresponding IBI durations can be read off the abscissa label in Figure 1B.) The percentage decreased monotonically with IOI duration, as expected. (The graphs in this article should generally be read from right to left, going from slow to fast rates.) The 50% intercept of the function (linearly interpolated between the adjacent data points) provided an estimate of the average synchronization threshold; it corresponded to an IOI of 123 ms.

There were large individual differences in the auditory synchronization threshold. Threshold estimates for individual participants are shown in Table 1, together with some information about participants' tapping experience and musical training. Individual thresholds ranged from 97 ms to more than 170 ms (i.e., outside the range of IOIs used). Although I had not designed the experiment to investigate specifically the role of musical training, one can see that the 5 active musicians in the group all achieved low thresholds, whereas the only musically untrained participant had the highest threshold by far.

The average percentages of successful trials as a function of IOI (= IBI) duration in the visual condition are shown in Figure 1B. Again, the percentages decreased monotonically with IOI duration. The average synchronization threshold, estimated by the 50% intercept, was 459 ms—remarkably, almost 4 times as high as the threshold in the auditory condition.[5] The individual threshold estimates are listed in Table 1. They ranged from 413 to 540 ms. (One threshold was uncertain because of erratic data.) There appeared to be no relation to musical experience, and the correlation between the individual auditory and visual synchronization thresholds was only .24 (*df* = 11, *ns*).

There were no instances of phase drift in the slow auditory condition. That finding confirms that no tapping rate in this experiment challenged the motor system in any way.

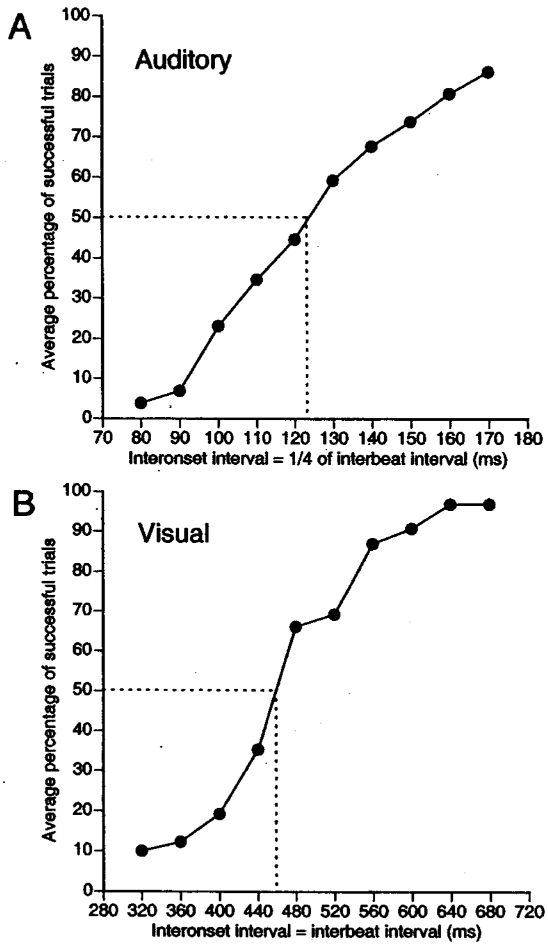

FIGURE 1. Average percentage of successful synchronization trials as a function of interonset interval duration (A) in the fast auditory condition and (B) in the visual condition of Experiment 1. Dotted lines show the estimation of the synchronization threshold.

Therefore, the synchronization thresholds can be safely attributed to limits of perception or sensorimotor coordination. Although the synchronization thresholds were the data of primary interest, I next briefly present analyses of asynchronies and ITIs to give a better picture of participants' performance in the tasks.

Asynchronies

The average asynchronies in successful trials as a function of IBI duration for all three conditions are presented in Figure 2A. The standard error bars reflect between-participant variability. It should be noted that the reliability and representativeness of the data decreased as IBI duration decreased in the fast auditory and visual conditions because of the decreasing number of successful trials (see Figure 1). Nevertheless, an impression of major trends and differences can be obtained from the graph. The significance of differences can be gauged (conservatively, in view of the within-participant design) from nonoverlapping standard error bars.

The asynchronies in the slow auditory condition were negative, indicating that the taps preceded the tone onsets. That anticipation tendency is commonly found in synchronization tasks, although it was rather small here, probably because of direct auditory feedback received from the taps (see Aschersleben & Prinz, 1995; Fraisse, Oléron, & Paillard, 1958). Some participants (all active musicians), however, did not show any anticipation at all. The asynchronies tended to decrease (i.e., to become less negative) as the IBI (= IOI, average ITI) decreased, at least from 600 ms onward, but there were large individual differences in that regard. Other investigators have also observed a decrease in the anticipation tendency with decreasing ITI duration (e.g., Mates, Radil, Müller, & Pöppel, 1994).

The average asynchronies in the fast auditory condition

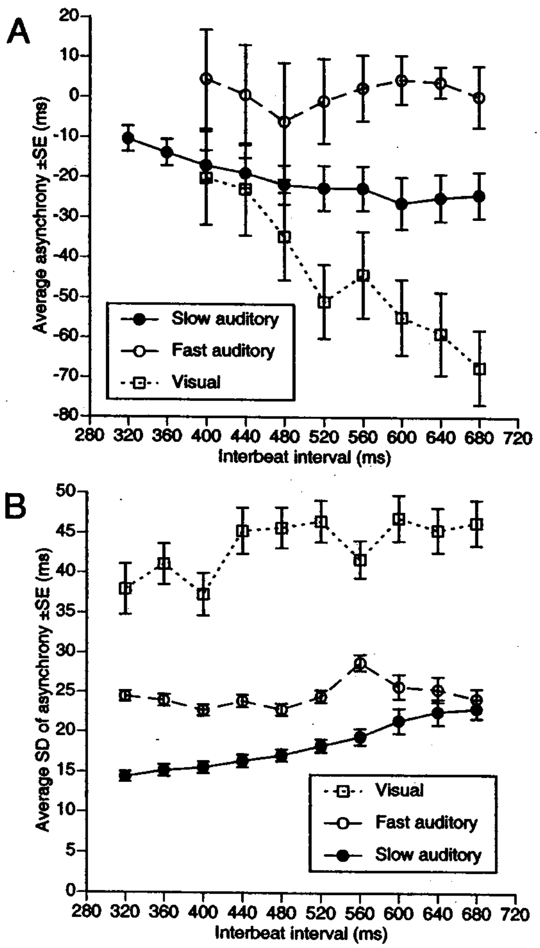

FIGURE 2. (A) Average asynchrony and (B) average within-trial standard deviation of asynchronies as a function of interbeat interval (IBI) duration in the three conditions of Experiment 1, with between-participant standard error bars.

hovered around zero and changed little with IBI duration.[6] That finding clearly reflects an effect of subdivision (i.e., of IOI duration), because the IBIs and the average ITIs were the same as those in the slow auditory condition. The effect is consistent with the gradual decrease in negative asynchrony as IBI (= IOI) duration decreased in the slow auditory condition, if one assumes that the decrease did not continue beyond asynchronies of zero. Wohlschläger and Koch (2000) have also observed that additional tones between target tones reduce or eliminate the anticipation effect.

At long IBIs, the asynchronies in the visual condition were much more negative than were those in the slow auditory condition.[7] However, they decreased rapidly with IBI duration and converged with the slow auditory asynchronies in the vicinity of the average visual synchronization threshold (459 ms). The finding of larger negative asynchronies in the visual condition is contrary to what was found in some previous studies (Fraisse, 1948; Kolers & Brewster, 1985; Repp & Penel, 2002).

In Figure 2B are the average within-trial standard deviations of the asynchronies as a function of IBI duration in the three conditions, with between-participant standard error bars.[8] The standard deviations in the slow auditory condition decreased monotonically as IBI duration decreased, as expected. The variability was clearly greater in the fast auditory condition than in the slow auditory condition, except at the longest IBIs. Thus, there was no (quadruple) subdivision benefit within that range of IBIs. On the contrary, there was an increasing cost of subdivision as the IBI decreased, because variability in the slow auditory condition decreased with IBI duration, whereas that in the fast auditory condition did not.[9] As expected, the average standard deviations of successful trials in the visual condition were substantially higher than were those in the slow auditory condition, at all IBI durations.

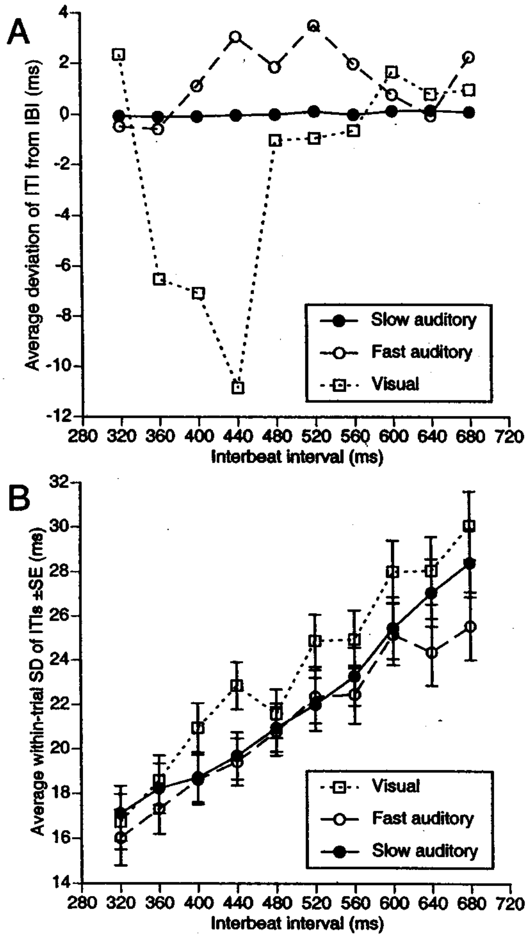

FIGURE 3. (A) Average deviation of intertap interval (ITI) duration from interbeat interval (IBI) duration and (B) average within-trial standard-deviation of ITIs as a function of IBI duration (with between-participant standard error bars) in the three conditions of Experiment 1.

Intertap Intervals

Because participants generally maintained a steady rate of tapping even when they failed to synchronize, one can compare the ITIs of all three conditions without eliminating unsuccessful trials. The deviations of the average ITIs from the IBIs are informative about the direction of drift in unsuccessful trials. The average deviations, which are a kind of overall bias or constant error, are shown in Figure 3A. Of course, they were near zero in the slow auditory condition. In the fast auditory condition, there was a tendency to tap too slowly at some intermediate IBIs, whereas in the visual condition there was a stronger tendency to tap too fast at some of the shorter IBIs. However, between-participant variability (not shown in the figure for graphic reasons) was very large. A small average deviation does not mean that participants generally tapped at the correct rate (except in the slow auditory condition); it implies only that positive and negative deviations from the IBI were about equally frequent. Individual participants tended to be consistent in the directions of their deviations. In the fast auditory condition, 7 participants generally tapped too slow and 6 tapped too fast. In the visual condition, 4 participants generally tapped too slow and 9 tapped too fast (except at the shortest IBI). These individual tendencies were similar in the fast auditory and visual conditions: At the two shortest IBIs (320 and 360 ms), where nearly all trials were unsuccessful, the individual deviations in the two conditions correlated .74 and .78, respectively ($df = 11, p < .01$).

The average within-trial standard deviations of the ITIs are shown in Figure 3B. Compared with the large differences among conditions in terms of the variability of the asynchronies (see Figure 2B), the differences among conditions were surprisingly small here, and the variability decreased as a strongly linear function of IBI duration in all three conditions. Subdivision of IBIs in the fast auditory condition clearly did not interfere with participants' ability to tap at a steady rate, even as synchronization became difficult or impossible. The average variability of ITIs in the visual condition was only slightly higher than that in the slow auditory condition. Thus, synchronization difficulties were generally not evident in the variability of the ITIs; they were restricted to the asynchronies. That finding is consistent with my impression as a participant that there was little awareness of asynchrony (or synchrony, for that matter) below the synchronization threshold. Thus, participants rarely attempted to correct their tapping but simply continued to tap "blindly" at a steady rate, hoping that the taps would coincide with the sequence events.

Discussion

Auditory Synchronization Threshold

London's (2002) and Iyer's (2002) suggestions of a rate limit of about 100 ms (IOI) for metrical subdivision are somewhat lower than the average auditory synchronization threshold of about 120 ms obtained in the present study. However, the lowest individual thresholds were indeed in the vicinity of 100 ms. That comparison seems more appropriate, because London and Iyer surely intended their estimates to apply to expert musicians. Individuals with special rhythmic expertise, such as professional percussionists, perhaps could do even better than the best of the present participants. It seems rather obvious that the auditory synchronization threshold is a measure of rhythmic ability and that musicians will generally have lower thresholds than will musically untrained individuals.

The present results are also in reasonable agreement with the findings of Bartlett and Bartlett (1959), even though a 1:4 tapping task was used here, whereas they required only a single tap per sequence. It seems likely that the response requirement made little difference, because the synchronization threshold clearly reflects a perceptual or sensorimotor limit, not one of motor behavior. The 1:4 tapping task requires a subjective organization of the tones into groups of four (i.e., a quadruple meter), as investigated long ago by Bolton (1894). On the basis of introspective reports, Bolton estimated that the subjective grouping of sounds becomes difficult at IOIs of 158 ms and impossible at IOIs of 115 ms. The latter estimate is also in good agreement with the average auditory synchronization threshold found in the present study.

Another estimate that falls in the same ballpark concerns the rate at which the number of sequentially presented auditory stimuli can be estimated accurately. Taubman (1950a) presented 1–10 tones at different rates and found that participants' number estimates were perfect when the IOI was 125 ms but that errors (generally underestimation of the number of events) began to arise at the next-shortest IOI in his design, which was 100 ms. Roughly comparable results were obtained in studies by Cheatham and White (1954), Garner (1951), Lechelt (1975), and Massaro (1976). According to Massaro, the maximal rate of verbal counting is about 6 words/s, or about 170 ms per utterance. It seems likely that the maximal rate at which auditory events (up to 10 or so) can be counted covertly (perhaps retrospectively, by retrieving the pattern from an auditory short-term memory) is related to the auditory synchronization threshold.

What is the underlying cause of that limit? The limit might be purely perceptual, or it might be sensorimotor. A perceptual limit could arise from persistence and temporal overlap of neural activity at some auditory processing level that result in successive stimuli cohering with each other and thus losing their individuality as distinct events. For example, the result of a study of the mismatch negativity (MMN) in event-related potentials elicited by missing tones in a sequence (Yabe, Tervaniemi, Reinikainen, & Näätänen, 1997) suggested an auditory temporal integration window of about 125-ms duration. Alternatively, a perceptual limit could reflect a rate limit of an internal periodicity that a sequence must entrain so that perceptual tracking can occur, or it might reflect the temporal resolution of a periodic attentional process (cf. Large & Jones, 1999). A sensorimotor explanation would be that there is a rate limit for processing the error feedback (asynchronies) that makes continuous phase correction and hence synchronization possible (Vorberg & Wing, 1996). If so, that limit presumably resides at the stage of registering or computing the asynchronies, which again points to an essentially perceptual problem resulting from the close spacing of the tones.

It should be emphasized that the auditory synchronization thresholds reported here apply to unstructured sequences only, that is, to isochronous sequences composed of identical events. Synchronization is likely to be facilitated if different pitches, event durations, or intensities recur regularly in an auditory sequence so that a metrical pattern emerges or perhaps even stream segregation (Bregman, 1990) occurs. In that case, synchronization with a higher level periodicity (1:*n* tapping) might be possible even when IOIs are shorter than 100 ms, although that remains to be demonstrated. Conversely, the synchronization threshold might be elevated if production of a more complex rhythm is required.

Visual Synchronization Threshold

The results of Experiment 1 revealed a remarkably large difference in people's ability to synchronize with visual and auditory sequences. The average visual synchronization threshold was about 460 ms of IOI, and no participant could synchronize reliably at IOIs shorter than 400 ms. The visual synchronization threshold thus was about 4 times as high as the auditory one. That finding is in general agreement with the results of other studies in which similar sequences composed of light flashes were used, and even with those of some studies in which moving stimuli were used (see introductory comments). Although better synchronization performance might be obtained with different visual displays, it seems highly unlikely that performance will ever be as good as with auditory sequences.

It is interesting to note that the visual synchronization threshold obtained here is also in rough agreement with results for visual numerosity estimation. Taubman (1950b) found perfect estimation of the number of light flashes when the IOI was 500 ms but significant underestimation at 333 ms, the next-shortest IOI in his design. Forsyth and Chapanis (1958) and Lechelt (1975) obtained comparable results. Thus, the large difference between the auditory and visual synchronization thresholds parallels a similarly large difference between the IOIs at which errors begin to appear in auditory and visual numerosity estimation.

There seems to be no relationship between the auditory and visual synchronization thresholds. The auditory threshold almost certainly depends on rhythmic ability and skill, but the visual threshold apparently does not. Although the present data are merely suggestive in that regard, it is likely that the two thresholds reflect different, modality-specific limitations.

What is it, then, that makes the threshold for synchronization with visual sequences so high? Again, one could think of purely perceptual or sensorimotor explanations. If the limitation were perceptual, however, that would amount to a serious handicap, which is not evident in other forms of visual perception. It seems more likely that the cause in this case is a sensorimotor one. The anatomical and functional connection between sensory processing areas and the motor system might be less close for vision than for audition (Fraisse, 1948). It is a plain fact that people readily tap their foot or move otherwise in synchrony with rhythmic auditory stimuli such as music, whereas they rarely do so in synchrony with rhythmic visual stimuli. One might argue that that difference is a consequence of the fact that rhythmic auditory sequences frequently occur in our environment, whereas rhythmic visual sequences rarely do. It seems more likely, however, that the paucity of rhythmic visual sequences in our environment is the result of the inherently unengaging nature of such stimuli. Although simple visual sequences can be regular and periodic, they are rhythmically inert, both in the sense that they do not readily elicit synchronous actions and also in that they cannot support any degree of rhythmic complexity or metricality (Gault & Goodfellow, 1938; Patel, Iversen, Chen, & Repp, 2002). Music presumably evolved in the auditory modality because only auditory rhythms readily entrain movements of the human body at multiple time scales.

Those general considerations do not really pinpoint the nature of the temporal limit in visual synchronization. Presumably, that limit reflects a difficulty in computing and processing the sensorimotor asynchronies that enable error correction in synchronization. That possibility could also account for the findings on covert counting of visual stimuli, because, at those slow rates, counting amounts to synchronization of inner speech with a sequence. No other visual tasks with such a severe temporal limit come to mind readily. Therefore, the visual synchronization threshold might reflect a limit that is specific to the rhythmic coordination of movement with visual stimuli. In future research, investigators will have to clarify how that hypothesis can be reconciled with the high temporal accuracy humans can achieve in nonrhythmic visual tasks such as ball catching or collision avoidance.

EXPERIMENT 2

The Benefits and Costs of Subdivision

In this experiment, I examined the extent to which physical subdivision of the IBIs in an isochronous sequence reduces the variability of the asynchronies between taps and target events, and how that subdivision benefit varies as a function of IOI duration in the auditory and visual modalities.[10] I expected that a benefit would occur when the IOIs are relatively long but that it would disappear and turn into a cost as IOI duration decreases. The critical IOI at which the transition occurs was expected to be well above the synchronization threshold, at least in the auditory modality. The auditory results of Experiment 1 provide some relevant evidence suggesting that the critical IOI is longer than 680/4 = 170 ms (see Figure 2B).

In the auditory condition, four IBI durations were used, each with no subdivision or with duple, triple, or quadruple subdivision (requiring 1:1, 1:2, 1:3, and 1:4 tapping, respectively). In the visual condition, eight longer IBIs were used, each with no or duple subdivision (requiring 1:1 and 1:2 tapping, respectively).[11] A second auditory condition with the same 4 × 4 design as in the first one, but with IBIs in the same range as those of the visual condition, was included for further comparison. The two auditory conditions are again referred to as *fast* and *slow*, respectively.

I used three subdivision conditions in the auditory conditions to compare the relative benefits of duple, triple, and quadruple subdivision. The results of several studies in the literature have indicated that triple subdivision is less easy or less natural than duple or quadruple subdivision. For example, Bolton (1894) found that subjective rhythmization of uniform auditory sequences occurs more commonly in groups of two or four than in groups of three (see also Parncutt, 1994). Drake (1993) observed that adults and children reproduce rhythms with a ternary meter less accurately than rhythms with a binary meter. Therefore, I considered it possible that the subdivision benefit would disappear sooner (i.e., at a longer IOI duration) with triple than with either duple or quadruple subdivision of IBIs.

Method

Participants

The participants included 8 paid volunteers (4 women, 4 men) and me. Six participants had moderate to extensive experience in auditory synchronization tasks; some of them also had limited experience with visual synchronization. Three participants were novices. Ages ranged from 20 to 27 years, except for one novice and me, who were 56 years old at the time. Musical training ranged from considerable (6 or more years) to none at all (see Table 3). All participants were right-handed.

Materials

Auditory and visual sequences were composed of the same tones or flashes as in Experiment 1. There were two sets of auditory sequences, each containing sequences with four different IBIs. In the fast set, the IBIs were 480, 600, 720, and 840 ms; in the slow set, they were 840, 960, 1,080, and 1,200 ms. The sequences with 840-ms IBIs were shared by the two sets so that any (unanticipated) effect of IBI range would be evident. Each sequence began with five tones at one of the four IBIs. From the fifth tone onward, the sequence either continued in the same way or became 2, 3, or 4 times as fast (i.e., the IBI was subdivided into two, three, or four IOIs). Thus, there were 16 different sequences in each set. Each sequence contained 44 target events, which were defined by temporal position only.

In the visual sequences, there were eight different IBIs, ranging from 840 to 1,260 ms in steps of 60 ms. After the fifth flash, each sequence continued either at the same tempo or became twice as fast (i.e., duple subdivision). Thus, there were also 16 different sequences in the visual set, each containing 44 target events.

The equipment for stimulus presentation was the same as in Experiment 1. The sequences in each set were presented five times in successive blocks, each containing a different random order of the same 16 sequences. The interval between sequences was 3 s (not self-paced).

Procedure

The experiment consisted of three sessions. The order of conditions was fixed: fast auditory, visual, and slow auditory. Participants were instructed to start tapping with the second sequence event and to continue tapping at the initial tempo (the IBI) until the end of the sequence, such that each tap coincided with a sequence event.

A different tapping device was used in this (earlier) experiment because the drum pad used in Experiment 1 had not yet been purchased. Participants sat in front of a computer monitor and tapped with the index finger of their preferred hand on a white key of a Fatar Studio 37 MIDI controller (a three-octave piano keyboard), which they held on their lap. The key moved about 10 mm and produced no sound unless it was struck rather hard, in which case some impact noise was audible. (The impact may have been audible for a few participants.) The key depression was sensed during the downward movement of the key, which added a negative constant (depending on the average key velocity) to the asynchronies.[12] The procedure was otherwise the same as that used in Experiment 1.

Data Editing and Analysis

Asynchronies were computed from the 5th tap onward and thus were based on 38 taps. I corrected misalignments caused by missing or extra taps before computing asynchronies. Trials exhibiting phase drift or other gross anomalies were identified by eye (usually without difficulty) and excluded. A special case of phase drift was synchronization at an incorrect IBI (e.g., with every fourth event instead of every third event in the sequence). Phase slips were another anomaly that occurred occasionally: Participants skipped an event in subdivided sequences but then continued to tap at the correct IBI. If the slip happened abruptly just once during a sequence, the trial was considered successful, and the asynchronies were recomputed relative to the effective target events following the slip. If phase slips happened more than once in a sequence or were gradual, the trial was considered unsuccessful and was excluded. The numbers of excluded trials are listed in Table 2.

Results

Slow Auditory Condition

This condition had been expected to be quite easy. Nevertheless, some participants experienced difficulties. One in particular, who had no musical training, seemed unable to tap with every third tone (triple subdivision) at any IBI. Usually, he tapped with every fourth tone instead (as did two other participants occasionally). To make it possible to include his data in the statistical analysis, I substituted the average of his data for the duple and quadruple subdivision conditions for his missing triple subdivision data at each IBI. Synchronization at an incorrect IBI never occurred in duple or quadruple subdivision conditions.

Participants' average asynchronies ranged from as much as –94 ms to as little as –6 ms. The grand average across participants was –40 ms.[13] The pattern of average asynchronies is shown in Figure 4A. The data are plotted as a function of IBI duration, with connected points representing the same subdivision condition. Standard error bars are not included because they would reflect only the large individual differences in the magnitude of the average asynchronies. A 4 × 4 repeated measures analysis of variance (ANOVA) confirmed that, as in Experiment 1, asynchronies tended to decrease (i.e., become less negative) as the IBI decreased, $F(3, 24) = 8.5$, $p < .006$, $\varepsilon = .58$.[14] Although subdivision also tended to reduce asynchronies, the main effect of subdivision did not reach significance, $F(3, 24) = 3.4$, $p < .08$, $\varepsilon = .48$. However, the interaction was significant, $F(9, 72) = 4.5$, $p < .02$, $\varepsilon = .34$: At the two shorter IBIs, any subdivision reduced the asynchronies, whereas at the two longer IBIs, only quadruple subdivision reduced the asynchronies. That interaction is difficult to interpret, and there were considerable individual differences in the pattern of asynchronies.

The results concerning the variability of the asynchronies were more consistent and straightforward. The average within-trial standard deviations as a function of IBI and subdivision are presented in Figure 4B. Again, standard errors are not included because they would mainly reflect individual differences in overall variability. (Individual average standard deviations ranged from 17 to 39

TABLE 2. Numbers of Unsuccessful Synchronization Trials (Out of 45 per Entry) in Experiment 2

Slow auditory		Fast auditory		Visual	
IBI/sub	No. trials	IBI/sub	No. trials	IBI/sub	No. trials
1,200/1	0	840/1	1	1,260/1	2
1,200/2	1	840/2	0	1,260/2	11
1,200/3	10	840/3	7	1,200/1	1
1,200/4	1	840/4	5	1,200/2	15
1,080/1	0	720/1	2	1,140/1	0
1,080/2	1	720/2	1	1,140/2	18
1,080/3	7	720/3	5	1,080/1	0
1,080/4	1	720/4	2	1,080/2	13
960/1	0	600/1	0	1,020/1	1
960/2	3	600/2	0	1,020/2	17
960/3	7	600/3	6	960/1	0
960/4	3	600/4	5	960/2	26
840/1	0	480/1	1	900/1	0
840/2	2	480/2	0	900/2	25
840/3	8	480/3	9	840/1	1
840/4	2	480/4	19	840/2	29

Note. IBI = interbeat interval (ms); sub = number of subdivisions.

ms.) Variability clearly decreased as IBI decreased, $F(3, 24) = 10.4$, $p < .001$, $\varepsilon = .84$, and there was also a striking decrease in variability as a consequence of subdivision (i.e., a subdivision benefit), $F(3, 24) = 21.9$, $p < .001$, $\varepsilon = .94$. The subdivision benefit was present at all IBIs; the interaction did not reach significance, $F(9, 72) = 2.1$, $p < .12$, $\varepsilon = .44$. Most interesting, there was little difference among duple, triple, and quadruple subdivision; they all conferred similar benefits. That finding was confirmed in a separate ANOVA on only those three subdivision conditions, in which both the main effect of subdivision and the interaction with IBI were nonsignificant. Clearly, the slow auditory condition did not include the critical IOI at which the subdivision benefit disappears, which means the critical IOI is shorter than 840/4 = 210 ms.

Fast Auditory Condition

This condition was expected to be more challenging than the slow auditory condition, although only one subcondition (the quadruple subdivision of IBIs of 480 ms, resulting in IOIs of 120 ms) extended into the region of the synchronization threshold. Indeed, in that subcondition, 19 out of 45 trials (42%) were unsuccessful, which suggests an average synchronization threshold somewhat below 120 ms, just slightly lower than that obtained in Experiment 1 (123 ms). Otherwise, the number of unsuccessful trials did not differ much from the slow auditory condition (see Table 2).[15] For the statistical analyses, a few missing data points (i.e., where all 5 trials were unsuccessful) were again replaced by the average values of adjacent subconditions for that individual.

The average asynchronies of the successful trials are shown in Figure 5A. They clearly decreased as the IBI decreased, $F(3, 24) = 25.7$, $p < .001$, $\varepsilon = .65$, at least from 720 ms on. They also decreased as subdivision level increased, $F(3, 24) = 21.6$, $p < .001$, $\varepsilon = .65$, but only up to triple subdivision; there was no significant difference between triple and quadruple subdivision. The interaction was not significant, $F(9, 72) = 1.8$, $p < .17$, $\varepsilon = .37$.

The average within-trial standard deviations of the asynchronies are shown in Figure 5B. They revealed an interesting pattern. Variability clearly decreased as the IBI decreased, $F(3, 24) = 9.6$, $p < .005$, $\varepsilon = .53$. Especially in baseline sequences with no subdivision (sub-1), the decrease was strikingly linear and steeper than in the slow auditory condition (Figure 4B). Subdivision also had a significant effect, $F(3, 24) = 6.6$, $p < .02$, $\varepsilon = .49$, but the effect was not monotonic and it interacted with IBI duration, $F(9, 72) = 3.5$, $p < .03$, $\varepsilon = .36$. Duple subdivision reduced variability, but the benefit decreased as the IBI decreased. Triple and quadruple subdivision were beneficial at the longest IBI only; from 720 ms on, they increased the variability, and as the IBI decreased, the subdivision cost tended to become greater for quadruple than for triple subdivision.

The Auditory Subdivision Benefit–Cost Transition

The average subdivision benefits for duple, triple, and quadruple subdivision as a function of IOI duration in the slow and fast auditory conditions are shown in Figure 6. The subdivision benefit was calculated as the difference between the standard deviations in baseline (sub-1) and subdivided sequences with the same IBI. The differences were calculated for individual participants before averaging, and the error bars represent double standard errors here, equivalent to 95% confidence intervals, so significant differences from

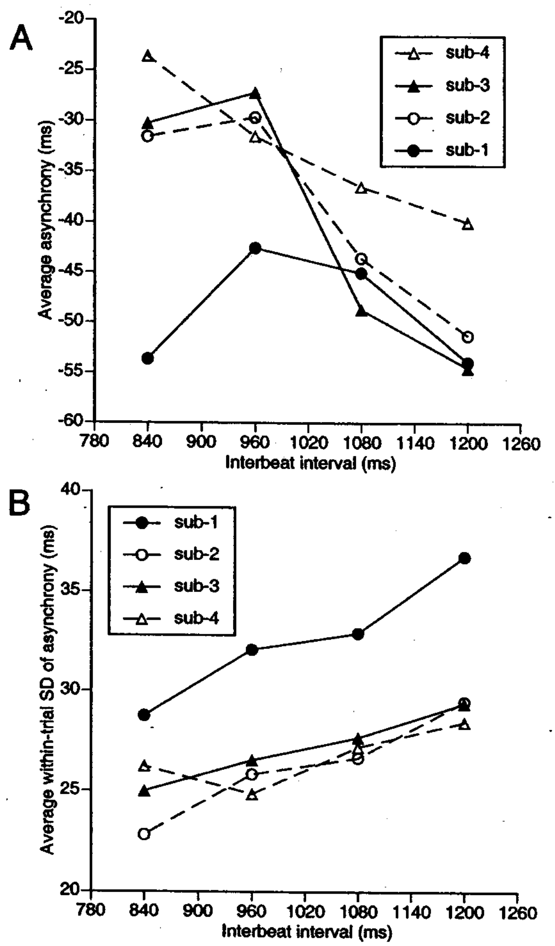

FIGURE 4. (A) Average asynchrony and (B) average within-trial standard deviation of asynchronies in the four auditory subdivision conditions (legend) as a function of interbeat interval (IBI) duration in the slow auditory condition of Experiment 2. Sub-1, -2, -3, and 4 = 1:1, 1:2, 1:3, and 1:4 tapping, respectively.

zero can be gauged by eye. The magnitude of the error bars betrays considerable individual differences.

The average subdivision benefit decreased with IOI duration in a roughly linear fashion within each condition, but the decrease was steeper in the fast than in the slow auditory condition. Even though similar benefits were obtained at the shared IBI of 840 ms in the two conditions, the different slopes suggest a range effect of some sort or a truly nonlinear decrease over a wide range of IOIs. Therefore, I used the regression line for only the fast auditory condition to estimate the average IOI at which the subdivision benefit vanished and turned into a cost. The zero crossing occurred at 197 ms (extrapolated) for duple, at 243 ms for triple, and at 189 ms for quadruple subdivision. The estimates derived from duple and quadruple subdivision thus were in close agreement (clearly, it was the IOI that limited the subdivision benefit, not the IBI), whereas that for triple subdivision suggested less benefit from subdivision in that condition, as predicted.

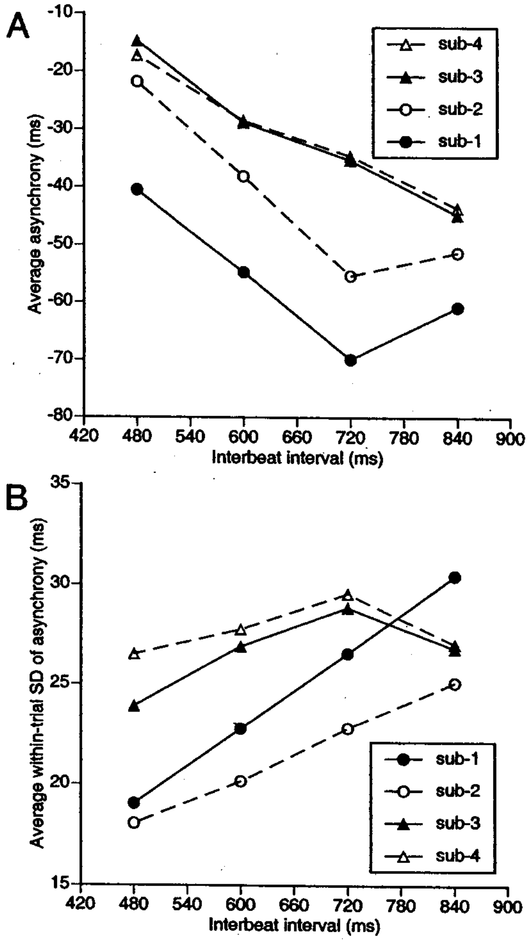

FIGURE 5. (A) Average asynchrony and (B) average within-trial standard deviation of asynchronies in the four auditory subdivision conditions (legend) as a function of interbeat interval (IBI) duration in the fast auditory condition of Experiment 2. Sub-1, -2, -3, and 4 = 1:1, 1:2, 1:3, and 1:4 tapping, respectively.

Benefit–cost transition estimates for individual participants are shown in Table 3. In six instances, no estimate could be obtained from the fast auditory condition because the regression line had a flat or even a negative slope.[16] In three of those cases, it was possible to obtain an estimate from the slow auditory condition. There was no systematic difference between the available individual estimates for duple and triple subdivision. However, all 8 participants for whom both pertinent estimates could be obtained showed a higher benefit–cost transition for triple than for quadruple subdivision, $t(7) = 12.3, p < .001$.

Visual Condition

The numbers of unsuccessful trials in that condition are listed in Table 2. Nobody had much difficulty with the

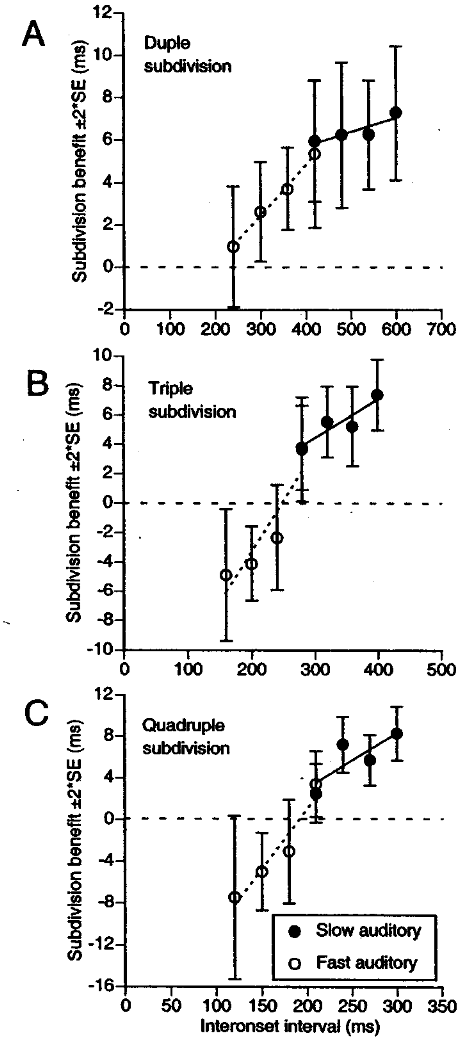

FIGURE 6. The auditory subdivision benefit (the difference between the average within-trial standard deviations of asynchronies in subdivided and unsubdivided sequences with the same interbeat interval, for (**A**) duple, (**B**) triple, and (**C**) quadruple subdivision conditions as a function of interonset interval duration in Experiment 2, with double standard error bars. Separate regression lines are shown for the fast and the slow auditory conditions.

baseline (sub-1) sequences, but there were many failures with subdivided (sub-2) sequences, even at rather long IBIs. Two participants were unable to synchronize with any subdivided sequences, and a third with nearly all. Three others had trouble with the faster subdivided sequences but did reasonably well with the slower ones. Only three participants were able to cope with all subdivision conditions. The number of unsuccessful trials exceeded 50% for the three shortest IOIs (480, 450, and 420 ms), which were in the vicinity of the average visual synchronization threshold obtained in Experiment 1 (459 ms).

In Figure 7A are shown the average asynchronies in baseline (sub-1) sequences as a function of IOI (= IBI) duration. The corresponding data from the slow auditory condition (Figure 4A) have been included for comparison. Standard errors have been omitted for reasons given earlier. It is evident that asynchronies were more negative in the visual than in the auditory task and that they also decreased more rapidly with IOI duration, as in Experiment 1 (Figure 2A). There were substantial individual differences, however, and neither the decrease with IOI duration nor the difference from the auditory condition was statistically reliable overall.

In Figure 7B, the average within-trial standard deviations of the asynchronies in baseline visual sequences are displayed as a function of IOI duration, with single standard error bars. The data from the slow auditory condition (Figure 4B) are shown for comparison. It is clear that the variability was much larger in the visual than in the auditory condition and that it decreased steadily as IOI duration decreased, $F(7, 56) = 6.5, p < .005, \varepsilon = .37$.

The effects of duple subdivision of visual sequences are shown in Figure 8. Because of the considerable data attrition (see Table 3), those data could not be subjected to an ANOVA. Instead, the average results for baseline sequences were subtracted from those for successful trials (if any) with subdivided sequences for each IBI and each participant, and the differences were then averaged across all participants for each IBI. The variable number of participants contributing data for each IBI is reflected indirectly in the double standard errors bars.

The positive differences between baseline and subdivision conditions in Figure 8A indicate that subdivision resulted in a substantial reduction of the negative asynchronies. For six of the eight IBIs, the reduction was significant on the basis of the confidence intervals, despite obviously large individual differences. That finding suggests that the magnitude of the negative asynchronies was primarily a function of sequence rate (i.e., IOI), not of tapping rate (IBI or average ITI). In Figure 8B, the subdivision benefit (the reduction in variability caused by subdivision) as a function of IOI duration is shown. Significant benefits were obtained at two long IOIs (570 and 630 ms), and benefits tended to decrease as IOI duration decreased, as expected. However, there was no clear evidence of a subdivision cost at short IOIs. The regression line intersects the abscissa at an IOI of 468 ms, which is close to both the average synchronization threshold obtained in Experiment 1 (459 ms) and the threshold suggested vaguely by the present data (Table 2).

Discussion

The Auditory Subdivision Benefit–Cost Transition

In addition to the smallest metrical subdivision of about 100 ms, London (2002) proposed another limit for metrical structures in music, which he estimated to be at IOIs of 200–250 ms. It is known from research on auditory pulse

TABLE 3. Individual Estimates of the Auditory Subdivision Benefit–Cost Transition (IOI Duration in ms) for Duple (Sub-2), Triple (Sub-3), and Quadruple (Sub-4) Subdivision

Participant	Sub-2	Sub-3	Sub-4	Tapping	Musical training
A.S.	184	261	193	Experienced	Limited
A.B.	292	196	158	Novice	Amateur
B.S	410*	237	170	Novice	Active amateur
B.R.	NA	302*	249*	Experienced	Active amateur
C.P.	267	249	206	Experienced	Amateur
D.P.	NA	219	NA	Novice	None
E.P.	241	235	181	Some experience	None
N.T.	289	219	175	Experienced	Limited
S.V.	158	183	143	Some experience	Active amateur

Note. Asterisk (*) = estimate derived from the slow auditory condition. NA = no estimate available. Individuals designated as amateur had 6 or more years of musical training.

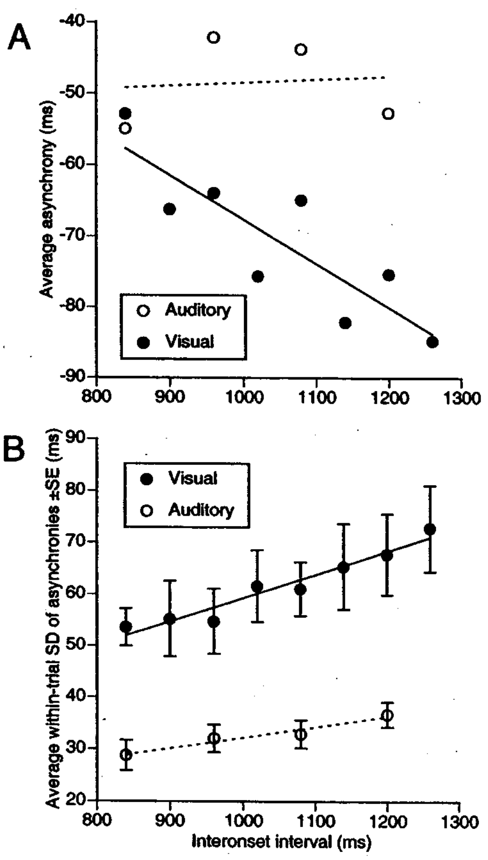

FIGURE 7. (A) Average asynchrony and (B) average within-trial standard deviation of asynchronies as a function of interonset interval duration in baseline (sub-1) visual sequences; auditory data (from Figure 4) are shown for comparison. Regression lines as well as standard error bars are shown in panel B.

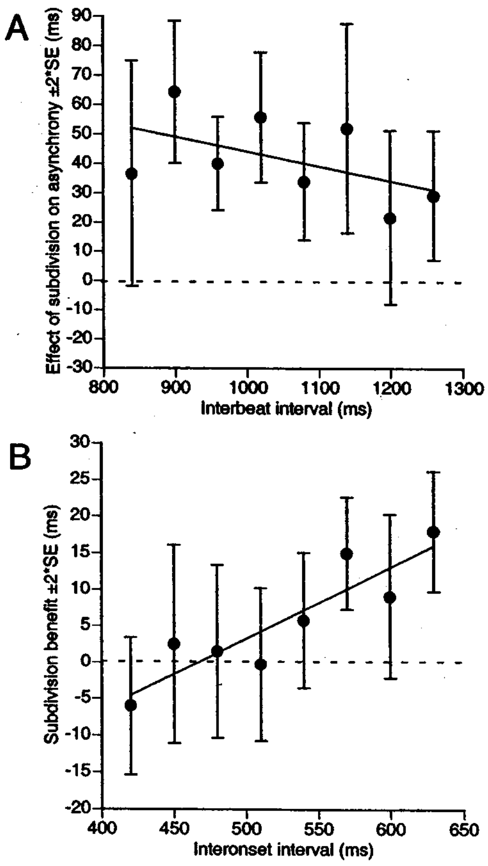

FIGURE 8. (A) Effect of subdivision on asynchronies and (B) subdivision benefit as a function of interonset interval duration in the visual sequences of Experiment 2. The regression lines shown, as well as the double standard error bars, were determined from limited available data.

salience (e.g., Handel & Lawson, 1983; Parncutt, 1994; van Noorden & Moelants, 1999) that that region corresponds to the fastest possible pulse or beat in music. London hypothesized that a viable beat must be potentially divisible. In other words, there must be another possible level of metrical structure below the beat; the beat itself cannot be the lowest level in a metrical hierarchy. If so, then the synchronization threshold, which estimates the rate limit of subdivisions, should also limit the maximal beat rate. The minimal IBI then should be about twice as long as the minimal IOI at the synchronization threshold, because that is the point where subdivision of the IBI into two IOIs is no longer possible.

The subdivision benefit–cost transitions estimated in Experiment 2 seems relevant to London's hypothesis. For duple, triple, and quadruple subdivision, the transition occurred at IOIs of about 180, 240, and 190 ms, respectively. Those values are not only similar to London's estimates but are also approximately twice as large as the average synchronization threshold obtained in Experiment 1, at least in the case of triple subdivision. One complication, however, is that those IOIs do not represent the IBIs of the sequences but rather the physical subdivisions of the IBIs. To interpret the results in the spirit of London's hypothesis, one might conclude that only potentially divisible subdivisions of a functional beat can confer a subdivision benefit.

There are additional reasons (see London, 2002) to expect some kind of discontinuity in the 200- to 250-ms region of the temporal continuum. The most important of those reasons is that Weber's law for duration discrimination breaks down around 250 ms; at shorter IOIs, the absolute discrimination threshold is constant rather than roughly proportional to duration, so the Weber fraction increases steeply as IOI decreases (see, e.g., Friberg & Sundberg, 1995). To obtain a subdivision benefit, participants in a synchronization task must derive temporal information from the IOIs that enables them to anticipate the next target tone more accurately. That information may no longer be precise enough to yield a benefit when the tones are separated by less than 200–250 ms.

In the present experiment, the rate of change of the (absolute) subdivision benefit or cost as a function of IOI duration was found to be smaller at long IOI durations. One possible reason for that finding is that the longer IOIs take over the role of functional IBIs because the prescribed IBIs are too long (cf. Parncutt, 1994). Thus, the observed subdivision benefit at the longer IOIs might be interpreted as being caused by the physical instantiation of functional beats versus their absence. That effect of beat instantiation may be less dependent on IOI duration than is the effect of beat subdivision.

In this experiment, three forms of subdivision (duple, triple, and quadruple) were compared. The average subdivision benefit or cost transition was similar for duple and quadruple subdivision. In other words, it was just as easy to tap with every fourth tone as with every other tone. Tapping with every third tone in a sequence, however, tended to be more difficult, and the subdivision benefit disappeared sooner in that condition. The greater ease of binary meters has been noted in several other studies in the literature (cited previously) and is supported by the relative prevalence of binary meters in the music of many different cultures. The relative difficulty of triple subdivision might disappear when the target events are physically accented (Drake, 1997).

The Visual Subdivision Benefit–Cost Transition

Like the visual synchronization threshold, the visual subdivision benefit–cost transition occurred at a much longer IOI (about 470 ms) than did the corresponding point for auditory sequences (200–250 ms). In contrast to audition, the visual subdivision benefit disappeared near the visual synchronization threshold, and there was no evidence of any subdivision cost preceding that threshold as IOI duration decreased. In other words, with visual sequences, there appeared to be only a single limit for both the ability to synchronize and the ability to derive a benefit from subdivision, whereas with auditory sequences there were two distinct limits. To the extent that those auditory limits reflect different aspects of metrical structure, with the upper limit relating to the maximal beat rate and the lower limit to minimal beat subdivision (London, 2002), the single visual limit can be interpreted as support for the claim that visual sequences cannot support a metrical hierarchy or a feeling of pulse (Patel et al., 2002). Or, if the auditory upper limit reflects a nonlinearity in temporal perception around 250 ms or so, then the visual results simply suggest that no such nonlinearity occurred within the range of longer IOIs (> 400 ms) that was used in the visual sequences. The visual subdivision benefit then merely reflects a decrease in variability as sequence IOI duration decreases, as long as synchronization is possible.

ACKNOWLEDGMENTS

National Institutes of Mental Health Grant MH-51230 supported this research. I am grateful to Yoko Hoshi and Helen Sayward for assistance; to Jack Vees for lending essential equipment; and to Simon Grondin, John Iversen, Peter Keller, Justin London, Aniruddh Patel, and three anonymous reviewers for helpful comments on earlier drafts.

NOTES

1. The term *tatum* is an abbreviation of "temporal atom," with the second vowel modified in honor of the legendary jazz pianist Art Tatum, whose recordings suggest an exceptional ability of rhythmic subdivision at fast tempi.

2. A subdivision benefit can also arise if participants create the subdivisions themselves by tapping between events (*n*:1 tapping; see, e.g., Pressing, 1998), counting (e.g., Grondin, Meilleur-Wells, & Lachance, 1999), or imagining subdivisions (e.g., Palmer & Krumhansl, 1990). However, those forms of self-generated subdivision were not investigated here.

3. The data of 1 additional novice were excluded because she was unable to synchronize at any tempo in both the fast auditory condition and in the visual condition.

4. All temporal intervals reported in this article have been men-

tioned as they appeared in the MAX environment. It is known from output measurements that real-time intervals are 2.4% shorter than those represented in MAX. If exactness is desired, the millisecond values reported should be multiplied by 0.976.

5. Although the two thresholds are more similar in terms of IBI duration, 492 versus 459 ms, that coincidence is almost certainly meaningless.

6. I edited those data to eliminate a few individual outliers beyond ±50 ms that indicated poor synchronization despite low variability and would have increased the error bars substantially. The data for the two shortest IOIs, which were based on very few trials and were very inconsistent, are not shown.

7. I edited the visual data to eliminate some individual values beyond ±100 ms as well as the data for the two shortest IOIs, for the same reasons as in the fast auditory condition.

8. Because the standard deviations were much less variable across trials than the asynchronies were, data are shown even for the shortest IBIs.

9. The relative constancy of the standard deviation for fast auditory sequences could be artifactual because the successful trials for short IBIs came primarily from participants with low synchronization thresholds, who also had the smallest standard deviations. Individually, however, those participants also showed fairly constant variability in their successful trials. Moreover, some of those participants actually showed small subdivision benefits, which in some cases extended very nearly to their synchronization threshold. Thus, the average cost of subdivision was mainly the result of participants with high synchronization thresholds.

10. No analysis of ITIs was conducted. Although a subdivision benefit would also be predicted for the variability of ITIs, the primary data were considered to be the asynchronies (of which the ITIs are the first difference, plus a constant) because only they are a measure of synchronization.

11. Triple or quadruple subdivision would have resulted in IOIs below the visual synchronization threshold for most IBIs.

12. Although velocities were not analyzed in this study, previous observations have suggested that variation in tap velocity is relatively small and unsystematic in isochronous tapping at moderate to slow rates. Therefore, velocity-dependent timing variability was probably negligible.

13. Because taps were registered during the downward movement, the asynchronies with respect to the bottom contact of the key were somewhat less negative (perhaps by about 10 ms).

14. All p values for F tests with multiple degrees of freedom include the Greenhouse-Geisser correction and are followed by the corresponding ε value.

15. However, there were no instances of synchronization at a wrong IBI in triple subdivision trials. Remarkably, the participant who had had difficulty with triple subdivision in the slow auditory condition, although he again had some trouble at the longest IBI, did fine at the shorter IBIs. That might have been because at the faster rates the IBI corresponded to the salient beat (tactus), whereas at slower rates the subdivisions constituted the beat, so that counting was required (see Parncutt, 1994). By contrast, another participant, who had done quite well in the slow auditory condition, had great difficulty with triple subdivision at faster rates. Those observations suggest that counting three beats requires somewhat different skills than does triple subdivision of a beat.

16. A novice without musical training (D. P.) and, most surprising, I (B. R.), contributed the most deviant results. The author generally showed no subdivision benefits in the fast auditory condition.

REFERENCES

Aschersleben, G., & Prinz, W. (1995). Synchronizing actions with events: The role of sensory information. *Perception & Psychophysics, 57,* 305–317.

Bartlett, N. R., & Bartlett, S. C. (1959). Synchronization of a motor response with an anticipated sensory event. *Psychological Review, 66,* 203–218.

Bilmes, J. A. (1993). *Timing is of the essence: Perceptual and computational techniques for representing, learning, and reproducing expressive timing in percussive rhythm.* Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, MA. <http://ssli.ee.washington.edu/~bilmes/mypapers/mit-thesis.pdf>

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology, 6,* 145–238.

Bregman, A. (1990). *Auditory scene analysis.* Cambridge, MA: MIT Press.

Byblow, W. D., Chua, R., & Goodman, D. (1995). Asymmetries in coupling dynamics of perception and action. *Journal of Motor Behavior, 27,* 123–137.

Cheatham, P. G., & White, C. T. (1954). Temporal numerosity: III. Auditory perception of number. *Journal of Experimental Psychology, 47,* 425–428.

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

Drake, C. (1997). Motor and perceptually preferred synchronization by children and adults: Binary and ternary ratios. *Polish Quarterly of Developmental Psychology, 3,* 43–61.

Dunlap, K. (1910). Reactions to rhythmic stimuli, with attempt to synchronize. *Psychological Review, 17,* 399–416.

Forsyth, D. M., & Chapanis, A. (1958). Counting repeated light flashes as a function of their number, their rate of presentation, and retinal location stimulated. *Journal of Experimental Psychology, 56,* 385–391.

Fraisse, P. (1948). Rythmes auditifs et rythmes visuels [Auditory and visual rhythms]. *L'Année Psychologique, 49,* 21–41.

Fraisse, P., Oléron, G., & Paillard, J. (1958). Sur les repères sensoriels qui permettent de controler les mouvements d'accompagnement de stimuli périodiques [On the sensory data that permit control of movements accompanying periodic stimuli]. *L'Année Psychologique, 58,* 322–338.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America, 98,* 2524–2531.

Garner, W. R. (1951). The accuracy of counting repeated short tones. *Journal of Experimental Psychology, 41,* 310–316.

Gault, R. H., & Goodfellow, L. D. (1938). An empirical comparison of audition, vision, and touch in the discrimination of temporal patterns and ability to reproduce them. *The Journal of General Psychology, 18,* 41–47.

Gibbon, J., Church, R., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (no. 423, pp. 52–77). New York: Annals of the New York Academy of Sciences.

Goodfellow, L. D. (1934). An empirical comparison of audition, vision, and touch in the discrimination of short intervals of time. *American Journal of Psychology, 46,* 243–258.

Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception & Psychophysics, 54,* 383–394.

Grondin, S., Meilleur-Wells, G., Ouellette, C., & Macar, F. (1998). Sensory effects on judgments of short time-intervals. *Psychological Research, 61,* 261–268.

Grondin, S., Ouellet, B., & Roussel, M.-E. (2001). About optimal timing and stability of Weber fraction for duration discrimination. *Acoustical Science and Technology, 22,* 370–372.

Grondin, S., & Rousseau, R. (1991). Judging the relative duration of multimodal short empty time intervals. *Perception & Psychophysics, 49,* 245–256.

Handel, S., & Lawson, G. R. (1983). The contextual nature of rhythmic interpretation. *Perception & Psychophysics, 34,* 103–120.

Iyer, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception, 19,* 387–414.

Klemmer, E. T. (1967). Sequences of responses to signals encoded in time only. *Acta Psychologica, 27,* 197–203.

Kolers, P. A., & Brewster, J. M. (1985). Rhythms and responses. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 150–167.

Large, E. W. (2000). On synchronizing movements to music. *Human Movement Science, 19,* 527–566.

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

Lechelt, E. C. (1975). Temporal numerosity discrimination: Intermodal comparisons revisited. *British Journal of Psychology, 66,* 101–108.

London, J. (2002). Cognitive constraints on metric systems: Some observations and hypotheses. *Music Perception, 19,* 529–550.

Madison, G. (2001). Variability in isochronous tapping: Higher order dependencies as a function of intertap interval. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 411–422.

Massaro, D. W. (1976). Perceiving and counting sounds. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 337–346.

Mates, J., Radil, T., Müller, U., & Pöppel, E. (1994). Temporal integration in sensorimotor synchronization. *Journal of Cognitive Neuroscience, 6,* 332–340.

Noble, M., Fitts, P. M., & Warren, C. E. (1955). The frequency response of skilled subjects in a pursuit tracking task. *Journal of Experimental Psychology, 49,* 249–256.

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

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception, 11,* 409–464.

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2002). Tapping to the beat of auditory and visual rhythmic sequences. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the Seventh International Conference on Music Perception and Cognition, Sydney, 2002.* Adelaide, Australia: Causal Productions (CD-ROM). (Abstract only.)

Peper, C. E., & Beek, P. J. (1998). Are frequency-induced transitions in rhythmic coordination mediated by a drop in amplitude? *Biological Cybernetics, 79,* 291–300.

Peper, C. E., Beek, P. J., & van Wieringen, P. C. W. (1995). Coupling strength in tapping a 2:3 polyrhythm. *Human Movement Science, 14,* 217–245.

Peters, M. (1989). The relationship between variability of intertap intervals and interval duration. *Psychological Research, 51,* 38–42.

Pressing, J. (1998). Error correction processes in temporal pattern production. *Journal of Mathematical Psychology, 42,* 63–101.

Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1085–1099.

Rousseau, R., Poirier, J., & Lemyre, L. (1983). Duration discrimination of empty time intervals marked by intermodal pulses. *Perception & Psychophysics, 34,* 541–548.

Semjen, A., Schulze, H.-H., & Vorberg, D. (1992). Temporal control in the coordination between repetitive tapping and periodic external stimuli. In C. Auxiette, C. Drake, & C. Gérard (Eds.), *Proceedings of the Fourth Rhythm Workshop: Rhythm Perception and Production* (pp. 73–78). Bourges, France: Imprimérie Municipale.

Semjen, A., Schulze, H.-H., & Vorberg, D. (2000). Timing precision in continuation and synchronization tapping. *Psychological Research, 63,* 137–147.

Taubman, R. E. (1950a). Studies in judged number: I. The judgment of auditory number. *The Journal of General Psychology, 43,* 167–194.

Taubman, R. E. (1950b). Studies in judged number: II. The judgment of visual number. *The Journal of General Psychology, 43,* 195–219.

van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research, 28,* 43–66.

Vorberg, D., & Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action* (Vol. 2, pp. 181–262). London: Academic Press.

Wimmers, R. H., Beek, P. J., & van Wieringen, P. C. W. (1992). Phase transitions in rhythmic tracking movements: A case of unilateral coupling. *Human Movement Science, 11,* 217–226.

Wohlschläger, A., & Koch, R. (2000). Synchronization error: An error in time perception. In P. Desain & L. Windsor (Eds.), *Rhythm perception and production* (pp. 115–127). Lisse, The Netherlands: Swets & Zeitlinger.

Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *NeuroReport, 8,* 1971–1974.

Submitted October 7, 2002
Revised January 8, 2003