TRACKING AN IMPOSED BEAT WITHIN A METRICAL GRID

BRUNO H. REPP
Haskins Laboratories

JOHN R. IVERSEN AND ANIRUDDH D. PATEL
The Neurosciences Institute

RHYTHMIC STRUCTURE OFTEN FAVORS a particular beat that is marked by frequent tone onsets and grouping accents. Using rhythms similar to those of Povel and Essens (1985), we asked musically trained participants to tap on physically or mentally imposed beats that either coincided with the favored beat or were phase-shifted relative to it. Surprisingly, tapping was equally stable. Actually, variability tended to be lowest when the imposed beat was in anti-phase with the favored beat; however, this tendency was reversed when participants were instructed to tap in anti-phase with the beat. These results demonstrate that precise on-beat synchronization with different imposed beats can be achieved by locking into the metrical grid defined by a rhythm's basic pulse. The favored beat provides the most stable reference for off-beat tapping but not necessarily for on-beat tapping, which relies to a greater extent on intervening rhythm tones as temporal references.

Received February 22, 2007, accepted March 12, 2008.

Key words: metrical interpretation, beat induction, rhythm, synchronization, tapping

THE INDUCTION OF BEAT perception by rhythmic patterns has been investigated in many studies (e.g., Desain & Honing, 1999; Hannon, Snyder, Eerola, & Krumhansl, 2004; Large, 2000, 2001; Large & Kolen, 1994; Large & Palmer, 2002; Longuet-Higgins & Lee, 1982; Snyder & Krumhansl, 2001; Toiviainen & Snyder, 2003). Various structural properties of a rhythm contribute to perception of a beat, but temporal structure tends to be most important (Hannon et al., 2004). In a now classic study, Povel and Essens (1985) constructed a set of rhythmic sequences by permuting a fixed set of eight intervals defined by the onsets of nine identical tones. Each rhythm contained five intervals of 200 ms duration, two intervals of 400 ms, and one of 600 ms, adding up to 2400 ms. Each rhythm was repeated cyclically, with an 800-ms interval separating the repetitions. Because of this particular temporal structure, all sequences had a basic pulse of 200 ms and were likely to be perceived as being in a duple (2/4) meter, with the most salient metrical level being an 800-ms beat starting with the first tone. However, the rhythms differed in how readily that beat could be perceived because they were more or less syncopated with respect to it. (See also Fitch & Rosenfeld, 2007.) When the task was to reproduce the rhythms, some were easier to reproduce than others, even though they all contained the same intervals and the same number of tones, and this was taken to reflect differences in the strength of the beat (or internal clock, as Povel and Essens called it) induced in listeners.

Povel and Essens (1985) were able to explain these differences by means of a model of beat induction that takes into account the distribution of accents due to temporal grouping of tones. According to their perceptual research, group-initial and group-final tones as well as isolated tones tend to be perceived as more prominent than other tones, although in groups of two tones the second tone is heard as more accented than the first (Povel & Okkerman, 1981). If these *grouping accents* are regularly spaced, they quickly induce perception of a beat; if not, the feeling of a regular beat does not emerge as readily. Here we call rhythms of the former kind *strongly beat-inducing* (SBI), and those of the latter kind, *weakly beat-inducing* (WBI).[1]

Figure 1 shows one example of each. Because of the empty 800-ms interval at the end of each rhythm cycle, both types of rhythm (when repeated cyclically) favor a beat with an 800-ms (or perhaps 400-ms) period, with the first beat coinciding with the initial event of the rhythm. However, SBI sequences imply that beat more

---

[1] In a previous study (Patel et al., 2005), we referred to these sequences as *strongly metrical* and *weakly metrical,* respectively, which we now find inappropriate because all the sequences of Povel and Essens (1985) are strongly metrical in the sense that they define a metrical grid, based on the basic pulse frequency of 200 ms, which can support a variety of different beats, as the present study shows.

Music Perception VOLUME 26, ISSUE 1, PP. 1–18, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 DOI:10.1525/MP.2008.26.1.1

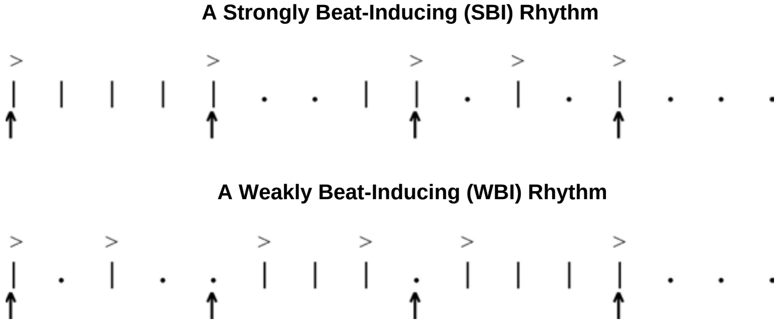

FIGURE 1. Examples of single cycles of strongly beat-inducing (SBI) and weakly beat-inducing (WBI) rhythms. The basic pulse (shortest interval) is 200 ms. Vertical bars indicate tone onsets, dots indicate metrical grid points without tone onsets, wedges indicate temporal grouping accents, and arrows indicate the favored beat. Note that in the SBI rhythm each favored beat coincides with a tone, whereas in the WBI rhythm only two of the four favored beats coincide with a tone. Also, in the SBI rhythm four out of five accented tones coincide with the favored beat, whereas in the WBI rhythm only two out of six do.

strongly than do WBI sequences because their grouping accent structure favors it as well. (See Figure 1 caption for further explanation.)

The less strongly a rhythm favors a particular beat (and the metrical hierarchy that comes with it), the more open it is to the mental construction of alternative beats and metrical interpretations. This constructive aspect of metrical perception has been neglected in research on rhythm and meter. In studies of beat induction, the listener is considered as a passive resonator or integrator of information who tries to discover the optimal metrical interpretation of the rhythmic structure presented to the ear. Although it is not uncommon to find that different listeners (or even the same listeners at different times) arrive at different metrical interpretations of the same rhythm, for any given listener at any given time his or her interpretation is considered to constitute a single, momentarily optimal solution to the informational jigsaw puzzle, with other possible solutions being discarded along the way or never even being considered. This approach contrasts with the complementary one taken here, which focuses on listeners' ability to construct and willfully impose different metrical interpretations on the same rhythmic substrate (see also Repp, 2005a, 2007).

Admittedly, endogenous construction and imposition of a beat is far less common than exogenous induction, and is usually found only in musically trained individuals. Nevertheless, it is a capability that is of considerable theoretical interest and important in musical contexts, most obviously so in music performance where a metrical framework for musical action must be constructed from memory or from musical notation. Metrical perception, too, can be endogenously determined when a rhythm is not strongly beat inducing. A metrical interpretation can be imposed on such a rhythm according to self-generated intentions, verbal instructions (Repp, 2005a), musical notation (Repp, 2007), or prior metrical context (i.e., via mental continuation of a previously induced beat). The present study used this last method to vary the assignment of different beats to the same rhythmic structure. One purpose of the research was simply to demonstrate that musically trained individuals are able to maintain an arbitrary induced beat when presented with relatively complex but metrically malleable (London, 2004) rhythms. To obtain an observable indicator of endogenously controlled metrical perception, participants were asked to tap on (i.e., in phase with) the imposed beat. A second, more specific, purpose of the research was to investigate whether the stability of on-beat tapping would be enhanced when the imposed beat coincides with the beat favored by the structure of the rhythm. In other words, the question was whether tapping with the favored beat of a SBI rhythm (the beat that would normally be induced by just listening to the rhythm) would be less variable than tapping with any other imposed beat.

In a previous study, we (Patel, Iversen, Chen, & Repp, 2005) investigated whether tapping on the strongly favored beat of SBI sequences would be more stable (less variable) than tapping on the less strongly favored beat of WBI sequences. From 35 sequences used by Povel and Essens (1985), which had been rank-ordered according to rhythm reproduction accuracy, we culled the 15 highest-ranked (SBI) and the 15 lowest-ranked (WBI) sequences. To prevent memorization of a particular rhythmic pattern, the rhythms were not repeated cyclically, as in the Povel and Essens study, but were concatenated in different random orders to yield unpredictable SBI and WBI sequences. However, each rhythm cycle still ended with an empty 800-ms interval, so that an 800-ms beat based on that interval was favored by both types of sequence, though more or less strongly. To eliminate any uncertainty about the beat, each sequence was preceded by an isochronous induction sequence that indicated the period and phase of the favored beat. Participants started synchronizing their taps with the induction sequence and then continued tapping on the beat as the rhythmic sequence unfolded. The participants were musically trained and had little difficulty with this task.

The results supported the hypothesis that tapping with the favored beat would be more stable in SBI than in WBI sequences. Asynchronies (computed relative to the theoretical beat location when the beat was not marked by a tone, as was often the case in WBI sequences) and inter-tap intervals were significantly less variable for SBI than for WBI sequences. The variability of tapping with the beat of SBI sequences was similar to that of tapping with a simple isochronous sequence having an 800-ms period.[2]

As illustrated in Figure 1, WBI rhythms often had no tone in the second and/or third beat locations of a cycle (the first and fourth beat locations were always marked by tones), whereas in SBI rhythms (and, of course, in isochronous sequences) all beat locations were marked by tones. This fact suggests an alternative, perhaps more fundamental explanation of the more variable synchronization with WBI sequences: Whenever a tap coincides with silence, there is no information about the synchronization error (i.e., there is no tap-tone asynchrony), and hence there can be no phase error correction or phase resetting on the next tap (cf. Repp, 2002, 2005b), which leads to increased variability. The results from two further conditions in our earlier study (Patel et al., 2005) are consistent with this hypothesis. In one condition, the missing beat tones of WBI sequences were filled in. Tapping variability in those sequences was similar to that in SBI sequences. In the other condition, the corresponding tones were removed from an isochronous sequence having 800-ms intervals. Tapping variability for the resulting sequence was greater than with an isochronous sequence, and similar to that for WBI sequences. However, an explanation in terms of intermittent error correction is difficult to distinguish from one based on reduced beat strength because beat strength is likely to be closely related to how often the beat is marked by a tone.

In the present study, we wanted to test further the hypothesis that sensorimotor coupling strength (reflected inversely in the variability of asynchronies) depends on beat strength and on the number of tones that mark the beat. In Experiment 1, in addition to comparing tapping on the favored beat of SBI and WBI sequences, as in Patel et al. (2005), we compared tapping on the favored beat with tapping on other beats that were physically or mentally imposed on the same sequence, the hypothesis being that tapping on imposed beats would be more variable and error-prone. To facilitate the imposition of different beats, to which the original SBI and WBI sequences might have been somewhat resistant, we modified these sequences by eliminating the recurrent empty 800-ms interval. That interval essentially constituted an inter-stimulus interval between rhythms encompassing three "measures" defined by four beats (see Figure 1). As such, it clearly constituted an important cue to the favored beat, and was probably the only such cue in WBI sequences. In the current sequences, the final tone of each component rhythm was simultaneously the first tone of the next (different) component rhythm, with the result that the sequences were no longer composed of identifiable rhythmic segments or cycles. We assumed that the consistent marking by tones and the regular patterning of grouping accents would still favor the same beat as previously in SBI sequences, though perhaps not as strongly, whereas the accent patterns of WBI sequences probably would not favor any particular beat phase (although it would still favor an 800-ms or perhaps 400-ms beat period). We examined the correctness of these assumptions in Experiment 2. In Experiment 3, we investigated both on-beat and off-beat tapping with respect to an imposed beat, in order to follow up on an unexpected finding in the two preceding experiments.

Experiment 1

To help participants impose different beats mentally, which was the task of primary interest, we employed a variant of the synchronization-continuation paradigm

---

[2] This result was considered surprising at the time. However, two control conditions suggested a possible reason. When participants tapped with every other tone of an isochronous sequence whose tone inter-onset intervals were 400 ms (duple subdivision of the 800-ms beat), variability was lower, but when they tapped with every fourth tone of an isochronous sequence having 200-ms intervals (quadruple subdivision of the 800-ms beat), variability increased again. These results are consistent with earlier findings showing a "subdivision benefit" in synchronization as long as the subdivision intervals are longer than 200-250 ms (Repp, 2003). The presence of five 200-ms intervals in the sequences may have canceled any subdivision benefit deriving from the two 400-ms intervals, and the 600-ms interval did not constitute a simple fraction of the beat.
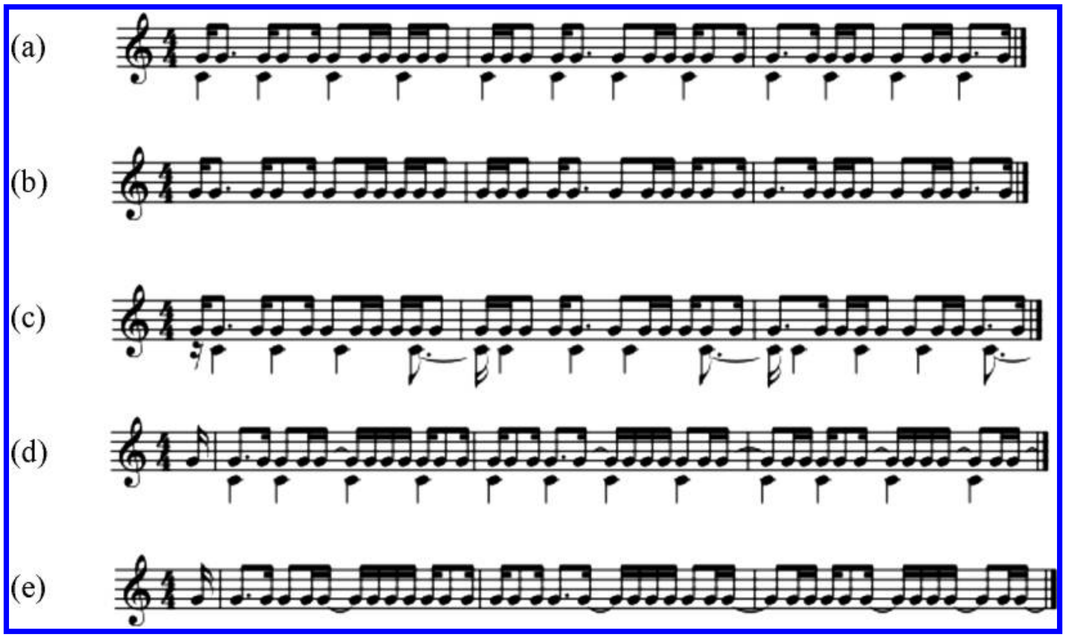
FIGURE 2. Schematic illustration of physical and mental beat imposition. (The pitches of the notes do not reflect those of the actual stimulus tones.) (a) Beginning of a SBI sequence, with an added beat marker (quarter notes) that is in phase with the favored beat (phase B1). (b) Metrical interpretation of the sequence after cessation of the beat marker (mental continuation of beat B1). (c) The same sequence with a beat marker that is one sixteenth-note out of phase with the favored beat (phase B2). (d) Actual metrical interpretation of this sequence forced by the beat marker. (e) Metrical interpretation after cessation of the beat marker (mental continuation of beat B2). Note that the rhythms in (b) and (e) are physically identical but radically different from each other at the subjective level.

(Stevens, 1886; Wing & Kristofferson, 1973). After an initial isochronous induction sequence consisting of low-pitched tones, the rhythm sequence started but was still accompanied during its first half by low-pitched tones that marked the designated beat. Thus, the beat was imposed physically at first, which is a form of beat induction (in this case by an extraneous stimulus rather than by the rhythm itself). Participants synchronized their taps with this explicit beat marker until it disappeared, and from that point on they continued tapping in the same beat phase until the end of the rhythm sequence. Participants thus had to maintain the imposed beat mentally during continuation tapping. We expected that the predicted differences in variability among the different beat phase conditions might already be evident during synchronization, but that they would become larger during continuation. (Note that the task was synchronization-continuation only with respect to the external beat signal; with respect to the rhythmic sequence and the beat imposed upon it, it was synchronization throughout.) Also included in the materials was an isochronous sequence consisting of 800-ms beat markers only, in order to replicate our earlier finding of no difference between tapping with an isochronous sequence and tapping with the favored beat of a SBI sequence.

It is important to understand that the tasks in Experiment 1 involving an imposed beat other than the favored beat (in SBI sequences) required on-beat tapping with respect to the imposed beat, not off-beat tapping with respect to the favored beat. Although this may seem like a spurious distinction from an observer's perspective, for a participant these two situations are very different. A beat cannot be perceived in two different phases at the same time: If a beat other than the favored beat is imposed successfully, the favored beat is not perceived. The perceptual organization of a rhythm is radically changed when different, mutually exclusive metrical interpretations are adopted (cf. Sloboda, 1985). When a beat other than the favored one is imposed on a SBI rhythm, the rhythm is perceived as highly syncopated, just like a WBI rhythm, and does not resemble at all the rhythm that is heard when the favored beat is adopted, even though the sequence remains physically invariant (if beat imposition is purely mental).[3] This subjective change is illustrated schematically in Figure 2. Because metrical

---

[3] This statement is based on the first author's impressions as a participant in the experiment. Of course, we do not have any direct knowledge of other participants' subjective experience, but it seems fair to assume that it was similar.

interpretation is a subjective phenomenon, we had to trust participants to follow instructions to the best of their ability and not to engage in off-beat tapping with respect to the favored beat when required to tap on a different beat. In fact, the danger of this happening was small because the favored beat was not likely to be perceived in the presence of a phase-shifted beat marker during the synchronization phase of a trial, where the beat marker naturally (and according to instructions) served as the beat.

Method

PARTICIPANTS

The participants were seven young paid volunteers (six women) and one of the authors (BHR). All were regular participants in synchronization experiments at Haskins Laboratories and had extensive music training. Three were graduate students or postgraduates of the Yale School of Music (viola, cello, bassoon); three were undergraduates, current or former members of the Yale Symphony Orchestra (cello, clarinet, percussion); one was an undergraduate who had had 7 years of flute instruction but did not play any more; and BHR (60 years old at the time of the study) has been an active amateur pianist all his life.

MATERIALS AND EQUIPMENT

The SBI and WBI sequences were created from those used by Patel et al. (2005) by closing up the recurring 800-ms intervals in text files containing MIDI instructions. The original sequences (see Patel et al. for a more detailed description) each consisted of 15 different, randomly concatenated rhythm cycles of the kind shown in Figure 1 ($15 \times 3.2 = 48$ s total duration), with 10 different random orders for each sequence type. Deletion of the 800-ms intervals reduced these sequences to linked concatenations of these rhythms ($15 \times 2.4 = 36$ s) in which the individual three-measure cycles were no longer recognizable as such. The rhythms were realized as sequences of identical high-pitched digital piano tones (A7, MIDI pitch 105, 3520 Hz). Each sequence was preceded by an induction sequence consisting of 9 low-pitched tones (A2, MIDI pitch 33, 55 Hz) with inter-onset intervals of 800 ms. The low tones continued as beat markers throughout the subsequent rhythmic sequence, which started with a delay (d) of 200, 400, 600, or 800 ms after the last induction tone. (The delay determined the beat phase.) The entire rhythm sequence was then repeated without interruption, but without the low tones. The total duration of a sequence (one trial) thus was $6.4 + d + 36 + 36 = d + 78.4$ s. Isochronous 800-ms beat sequences consisted of low tones only and lasted 44 s.

The sequences were played back on a Roland RD-250s digital piano under control of a program written in MAX 4.0.9. The software ran on an iMac G4 computer that was connected to the digital piano via a MOTU Fastlane USB MIDI translator. Low tones had a nominal duration of 50 ms; high tones had no specified duration and decayed freely within about 100 ms. Low tones (MIDI key velocity of 30) were softer than high tones (MIDI key velocity of 60) but nevertheless quite salient perceptually.

DESIGN AND PROCEDURE

Participants sat in front of a computer monitor on which the current trial number was displayed, listened to the sequences over Sennheiser HD540 II earphones at a comfortable loudness level, and tapped with the index finger of their preferred hand (the right hand for all but one) on a Roland SPD-6 percussion pad held on their lap. Most participants rested the wrist and other fingers of their hand on the surface of the pad and tapped by moving the index finger only; some, however, tapped by moving the wrist and elbow. The impact of the finger on the rubber pad was audible as a thud whose loudness depended on the tapping force.

Participants received one short practice block consisting of 5 trials and then 8 test blocks, each containing 10 trials (4 SBI, 4 WBI, 2 isochronous) in different random orders. In constructing the test blocks, 8 different versions (random concatenations of different component rhythms) of each rhythmic sequence type were used, each of which appeared twice with each beat phase (delay). The versions and beat phase conditions were assigned to the blocks in a balanced fashion, such that each beat phase condition occurred in each block, but with different versions of the SBI or WBI sequence. The experiment required two sessions of about 1 hour duration, typically one week apart. Participants were instructed to start tapping with the third induction tone, to keep synchronizing their taps with the low tones while they were present, and to continue tapping in the same beat phase after the low tones ended, in time with the rhythm sequence. The importance of maintaining the imposed beat mentally and tapping on that beat was emphasized. Participants were told that they should repeat a trial if they noticed that their mental beat had shifted its phase during the continuation part of a trial.

Results

No participant had any serious difficulties with the tasks. However, trials were occasionally repeated, and there were others in which the taps slipped into an

TABLE 1. Distribution of Repeats and Errors Across Conditions in Experiment 1.

Correct phase	Number of repeats		Number of errors (actual tapping phase)									
			SBI sequences					WBI sequences				
	SBI	WBI	B1	B2	B3	B4	Total	B1	B2	B3	B4	Total
B1	1	3	—	3	0	1	4	—	1	2	0	3
B2	2	2	0	—	1	0	2*	0	—	1	1	2
B3	3	3	1	1	—	0	2	0	0	—	1	1
B4	6	5	6	0	1	—	7	2	2	0	—	4

*One trial showed continuous phase drift.

incorrect phase during the continuation part without the participant noticing it, or without bothering to repeat the trial. Altogether 25 trials (3.9%) were repeated and then executed correctly; only the repetition was analyzed. Another 25 trials contained phase slips or phase drift (called "errors" in the following), and 8 trials were inadvertently skipped or contained a large number of missing taps. These 33 trials (5.2%) were excluded from analysis.

The distribution of repeats and errors across the various conditions, shown in Table 1, is of some interest. The four phases of the imposed beat (the intended "correct" phases of the taps) are referred to in the following as B1 (coinciding with the beginning of the sequence, 800 ms delay), B2 (600 ms delay), B3 (400 ms delay), and B4 (200 ms delay). Although the data in Table 1 are too sparse for statistical analysis, they do suggest that the B4 condition was slightly more difficult than the others in SBI sequences. Moreover, in that condition, six out of seven errors represented slips into the B1 phase, which suggests that the favored beat of SBI sequences functioned as an attractor in the rare cases when the mentally imposed beat became unstable (cf. Fitch & Rosenfeld, 2007). The other phase conditions for SBI trials and all four phase conditions for WBI trials did not differ much in terms of repeats and errors.

The principal performance measure for correct trials was the standard deviation of the asynchronies between taps and beats, computed with respect to the theoretical position of the imposed beat when it was not marked by a tone. These data are shown in Figure 3.

A repeated-measures ANOVA was conducted, with the variables of sequence type (SBI, WBI), tapping condition (synchronization, continuation), and beat phase (B1, B2, B3, B4). The only significant effect was

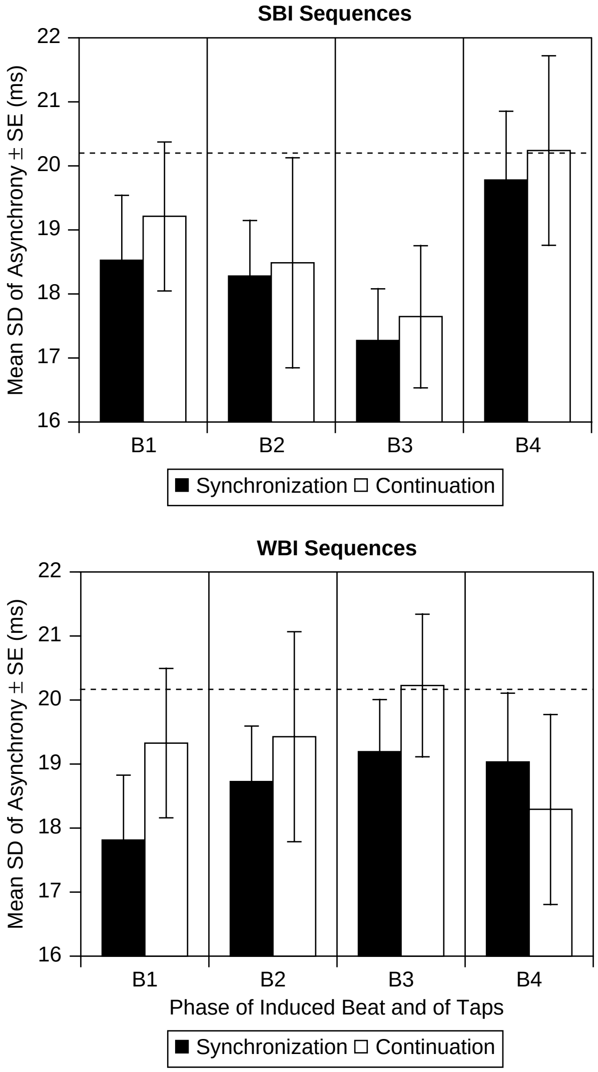

FIGURE 3. Mean standard deviation of asynchronies in the various conditions of Experiment 1, with standard error bars. The dotted horizontal line is the mean standard deviation of asynchronies for tapping with an isochronous sequence of beats.

the Sequence Type × Beat Phase interaction, $F(3, 21) = 4.41, p < .03$.[4] Separate ANOVAs were subsequently conducted on SBI and WBI sequences. For SBI sequences, there was a significant main effect of beat phase, $F(3, 21) = 5.35, p < .02$, whereas for WBI sequences there was no significant effect. This was just as predicted, but the pattern of differences among the beat phase conditions for SBI sequences was not the expected one: Variability was not lowest in the B1 condition, which represented the favored beat, but rather in the B3 condition, where the imposed beat and the taps were in anti-phase with the (presumably unperceived) favored beat. Variability was highest in the B4 condition, which agrees with the increased frequencies of repeats and errors in that condition (Table 1).[5] Furthermore, contrary to our hypothesis and the findings of Patel et al. (2005), tapping with the favored beat of SBI sequences (B1) was not less variable than tapping with the analogous beat (B1) of WBI sequences.

Another surprising but gratifying finding was that the main effect of tapping condition was not significant. Participants' taps were not significantly more variable during continuation tapping, where they synchronized with a mentally imposed beat, than during synchronization with an explicit beat marker (although there was a tendency in that direction), nor did differences among beat phase conditions increase from synchronization to continuation. This indicates that participants were quite successful in maintaining the imposed beat mentally, making it almost equivalent to a physically marked beat. Yet another unexpected result was that the tapping variability for isochronous beat sequences (where the taps were always in phase B1) tended to be *greater* than for rhythmic sequences, regardless of their beat phase (see dotted horizontal lines in Figure 3). However, there were substantial individual differences in that respect, and two participants (most notably author BHR) showed an opposite difference, so the overall difference was not significant.

There were no statistically reliable differences in mean asynchrony across conditions. The grand mean asynchrony was close to −16 ms in all conditions, but there were considerable individual differences in mean asynchrony.[6]

TABLE 2. Tones and Grouping Accents in Different Beat Phases.

	B1	B2	B3	B4
(A) Percentage of beats coinciding with tones				
SBI	100	50	50	63
WBI	59	63	74	67
(B) Percentage of beats coinciding with accented tones				
	B1	B2	B3	B4
SBI	63	30	40	14
WBI	33	30	46	39
(C) Percentage of accented tones				
	B1	B2	B3	B4
SBI	63	61	81	22
WBI	56	48	62	58

In order to test more precisely the hypothesis that the variability of tapping in a particular beat phase depends on the number of tones that coincide with taps, we analyzed the structure of SBI and WBI sequences by computing the percentage of beat locations in each phase that was marked by tones. These percentages are shown in Table 2A. It can be seen that the favored beat of SBI sequences (B1) was always marked by tones, as we have noted earlier, whereas the corresponding beat in WBI sequences (B1) was marked only 59% of the time. Yet these two tapping conditions did not differ in the variability of asynchronies. The percentages for the other beat phase conditions in Table 2 likewise bear no relationship to the pattern of variability results obtained (Figure 3). Indeed, the beat was marked more often by a tone in the condition with the highest variability (B4 in SBI sequences) than in the condition with the lowest variability (B3 in SBI sequences). These results flatly contradict the hypothesis that the number of tones coinciding with a beat is important for the stability of tapping, at least as long as 50% or more of the beat positions are marked.

Perhaps a more important factor is whether the tones coinciding with a beat carry a rhythmic grouping

---

[4] The Greenhouse-Geisser correction was applied to all F values with more than one degree of freedom in the numerator.

[5] We also analyzed the data in terms of the standard deviations of inter-tap intervals. The pattern of results was similar to the one for the standard deviations of asynchronies.

[6] Recent measurements using the current laboratory setup have revealed processing delays in the MIDI output and registration of taps that add about 15 ms to the measured asynchronies. This has been taken into account here, but asynchronies reported in previous studies by the first author using the same setup are about 15 ms too long. The statement (e.g., in Repp, 2005b) that musically trained participants sometimes show no negative mean asynchrony needs to be revised.

accent. As Povel and Essens (1985) have shown, it is the regularity of accented events that induces perception of a beat. We analyzed SBI and WBI sequences accordingly by calculating the percentage of beat locations in each phase that coincided with accented tones, using the accenting rules mentioned in the Introduction (Povel & Essens, 1985; Povel & Okkerman, 1981). For that purpose, only tones separated by 200 ms were considered to form a group; tones separated by longer intervals were considered isolated and accented by default. The percentages varied slightly across different sequences because different random concatenations of the component rhythms created different groupings. The mean percentages are shown in Table 2B. It can be seen that beat phase B4 in SBI sequences had a particularly low percentage of accented tones, which may account for its relative difficulty. Apart from that, however, these percentages still bear little relation to the variability results.

Finally, we calculated the percentages of tones in each beat phase that were accented (i.e., the entries in Table 2B expressed as a percentage of those in Table 2A), which are shown in Table 2C. These results are perhaps most revealing because they show that phase B3 in SBI sequences actually had a higher percentage than phase B1, the favored beat, whereas phase B4 had the lowest percentage. These values thus correspond to the observed differences in tapping variability across beat phases in SBI sequences. For WBI sequences there is still no relation with the variability results, but the differences in variability were not significant for WBI sequences. The results thus suggest that synchronization with an imposed beat is most stable when most of the tones coinciding with it—even though they may be relatively few in number—are accented. This may still not be the correct explanation, however. We will consider an alternative, preferred explanation in connection with Experiment 3.

Discussion

The results of Experiment 1 are surprising in several respects. First, they fail to replicate the finding of Patel et al. (2005) that tapping with the favored beat of SBI sequences is less variable than tapping with the analogous beat (B1) of WBI sequences. The earlier finding thus seemed to depend on the presence of the recurring empty 800-ms interval between component rhythms, not on the fact that the favored beat of SBI sequences was always marked by a tone. (See the General Discussion for consideration of other factors that might account for the different results.) The 800-ms interval alone undoubtedly helped induce a beat in phase B1 in both SBI and WBI rhythms, but in SBI rhythms that favored beat was supported by the rhythmic structure, whereas in WBI rhythms it was not. The present findings suggest that consistent marking of a beat by tones is not necessary for achieving high tapping stability; musically trained participants can tap in arbitrary beat phases, where only 50 percent of the beat locations are marked by tones, with essentially the same accuracy and stability. Put differently, tapping with the beat of a syncopated rhythm (three of the SBI beat phases and all of the WBI beat phases entailed frequent syncopation) is just as precise and almost as easy for musicians (apart from a few repetitions and phase slips) as tapping with the beat of a nonsyncopated rhythm (B1 in SBI rhythms). By failing to support the hypothesis that tapping stability rests on the simple percentage of beat locations that are marked by tones (at least for percentages above 50%), the results also suggest that tapping stability does not depend crucially on the frequent registration and correction of asynchronies. This is consistent with an interpretation of phase correction in synchronization as being based on phase resetting with respect to temporal references, rather than being strictly asynchrony-based (see Repp, 2005b, 2008). We will return to this issue in Experiment 3.

A second surprising result is that at least some participants' taps were more variable when they accompanied an isochronous beat sequence (in phase B1) than when they accompanied a rhythmic sequence. Given that Patel et al. (2005) had found equal variability in tapping with the favored beat of SBI sequences and with an isochronous sequence having an 800-ms period, the present results suggest paradoxically that the favored beat gained an advantage over an isochronous sequence after it was weakened by removal of the recurrent 800-ms empty interval in the SBI sequences. One possible explanation is that the recurrent empty interval, when present, introduced systematic tapping variability that increased the overall variability of tapping with the favored beat of SBI sequences, and Patel et al. indeed found some evidence of such systematic variability. Another possibility is that it is more difficult to tap consistently with the soft low-pitched tones used as beat markers here than with the high-pitched tones used in the isochronous sequences of Patel et al. In the present experiment, synchronization with low tones (and with a mentally imposed beat) may have been aided by rhythmic context, whereas synchronization with the high-pitched beat tones of Patel et al. did not show any such facilitation because they were simply part of the rhythm.

A third surprising result is that there was little difference in tapping variability between the synchronization and continuation phases of trials. In other words, participants were basically as accurate when tapping with a mentally imposed beat as when tapping with an explicit beat marker. This suggests strongly that the rhythmic context, not the physical instantiation of the beat itself, is responsible for the stability of sensorimotor synchronization with a beat. (Note that synchronization could not be maintained by relying on a mental beat alone; there must be some physical reference for phase correction or resetting.) This conclusion is also consistent with a phase resetting interpretation of phase correction in synchronization.

In Experiment 1, we assumed that the rhythmic structure of SBI sequences would still favor a B1 beat, despite removal of the recurrent empty 800-ms interval from the original sequences. We also assumed that WBI sequences would no longer favor the B1 (or any other) beat phase. We tested these assumptions in Experiment 2, in which participants could choose the beat themselves. This gave us also the opportunity to collect additional data on tapping variability in different beat phases.

Experiment 2

In Experiment 2, participants listened to the rhythm sequences, without hearing any preceding induction sequence or explicit beat markers, and chose the 800-ms beat they preferred. Once they had decided on a beat, they tapped along with it. We expected that participants would show a preference for B1 in SBI sequences, but no clear preference in WBI sequences. The question of interest was whether tapping variability in different beat phases would follow the same pattern as in Experiment 1, now that the beats were self-chosen.

Preference for a particular beat phase needs to be separated from a tendency to choose a beat that starts with the first tone of a sequence, which naturally would favor B1, albeit in both SBI and WBI sequences. Toiviainen and Snyder (2003) found such a tendency in their study of beat finding in excerpts from music by J. S. Bach, in which they varied the starting point of the excerpts. Therefore, we varied the starting point of the rhythm sequences in Experiment 2. The question of whether the starting point would influence the choice of beat in the present materials was of some interest in itself.

Method

PARTICIPANTS

Eight young paid volunteers (five women) and author BHR participated. Four (a cellist, a bassoonist, a clarinetist, and BHR) had participated in Experiment 1, but more than one year had elapsed and many other experiments had intervened. The other five participants included four pianists and one cellist, all undergraduate or graduate students with extensive music training but without previous experience in tapping experiments.

MATERIALS AND EQUIPMENT

The sequences corresponded to the continuation part (36 s) of the rhythmic sequences of Experiment 1. That is, each sequence was presented once without any preceding induction sequence or explicit beat markers. Obviously, there was also no need for an isochronous sequence in this experiment. The sequences were presented in 8 blocks of 8 trials each. Each block contained one SBI sequence and one WBI sequence, each presented with four different starting points, in random order. The nominal starting points were 0, 200, 400, and 600 ms from the beginning, corresponding to beat phases B1-B4. However, the later starting points were often not marked by tones, and therefore the actual starting point was taken to be the first tone following the nominal starting point. For example, for the WBI rhythm shown in Figure 1, the actual starting point for the 200 and 400 ms nominal starting points (corresponding to beat phases B2 and B3, respectively) would be the second tone (B3), and the actual starting point for the nominal 600 ms starting point (B4) would be the third tone (B2). This had to be taken into account in the analysis. Although the nominal starting phases were equally frequent, each occurring 8 times for each sequence type, the actual starting phases were not: Their frequencies for SBI sequences were 18 (B1), 7 (B2), 3 (B3), and 4 (B4), and those for WBI sequences were 8 (B1), 7 (B2), 11 (B3), and 6 (B4).[7]

PROCEDURE

Participants were told that some sequences would have a clear beat whereas others would not. They were asked not to start tapping immediately but to listen to each

---

[7] In hindsight, it might have been better to construct new sequences in which the nominal onset was always marked by a tone; this would have resulted in equal frequencies of the different actual starting phases. However, those sequences would no longer have matched exactly the sequences used in Experiment 1.

sequence until they had decided upon a beat and then to tap with that beat until the end of the sequence. An example of an 800-ms beat and a few practice trials were given. All participants readily chose an 800-ms beat during practice and never chose any other beat period during the experiment. (Although other beat periods are possible in principle, they were not of any interest in the present context.)

Results

Only trials in which participants' taps maintained the same beat phase throughout the sequence were included in the analysis. The data of two novice participants had to be excluded entirely. One of them exhibited phase slips or phase drift in almost all trials; the other one always started tapping almost immediately, contrary to instructions, and also had phase slips in 17 out of 64 trials. This demonstrates that even musically trained participants can have considerable difficulty with the beat tapping task. Of the remaining seven participants, only three committed any phase slips, amounting to a total of 16 trials (3.6% of all trials) that were excluded.

Table 3 shows the percentage of trials in which participants tapped in a particular beat phase (relative to the original starting point of the sequence, B1) for each of the *actual* starting phases of SBI and WBI sequences. It can be seen that there was a strong tendency to tap in the phase that the sequence started with (boldface percentages in the diagonal). On average, participants tapped in the starting phase of SBI and WBI sequences in 56.1% and 61.5% of the trials, respectively (with 25% being chance). However, this tendency varied greatly across participants, with some almost always tapping in the starting phase and others showing little influence of starting phase. Furthermore, it can be seen in Table 3A that there was a clear preference for B1, the favored beat, in SBI sequences. Participants were much more likely to tap in the starting phase when a SBI sequence started in phase B1 (80.2%) than when it started in another phase (about 48%). When participants did *not* tap in the starting phase, they were much more likely to tap in phase B1 (54.5%) than in any other phase (about 15%; see the normalized means of off-diagonal percentages at the bottom of Table 3A). There was no strong preference among the three other beat phases. The corresponding percentages for WBI (Table 3B) sequences reveal an unexpected preference for B4.

TABLE 3. Percentage of Trials in which Participants Tapped in a Particular Beat Phase (Relative to the Original Starting Point of the Sequence), for each *Actual* Starting Beat Phase.

(A) Strongly beat-inducing (SBI) sequences				
Starting phase of sequence	Tapping phase			
	B1	B2	B3	B4
B1	**80.2**	8.4	2.3	9.2
B2	37.5	**50.0**	12.5	0.0
B3	33.3	9.5	**47.6**	9.5
B4	25.0	10.7	17.9	**46.4**
Off-diagonal mean	54.5	16.3	18.6	10.6
(B) Weakly beat-inducing (WBI) sequences				
	B1	B2	B3	B4
B1	**53.7**	16.7	13.0	16.7
B2	12.8	**51.1**	10.6	25.5
B3	11.0	6.8	**61.6**	20.5
B4	0.0	7.7	12.8	**79.5**
Off-diagonal mean	15.4	20.3	23.6	40.7

Note: Column means of off-diagonal percentages (off-diagonal mean) are normalized to add up to 100%.

We also measured the time it took participants to start tapping. The mean starting time was 7.0 s (i.e., after about 8 beats had elapsed), with individual means ranging from 4.1 to 10.8 s.[8] A two-way repeated-measures ANOVA with the variables of sequence type (SBI, WBI) and tapping phase (B1, B2, B3, B4) revealed a mere tendency toward a two-way interaction, $F(3, 18) = 2.84$, $p < .10$: The mean starting time was shortest (6.0 s) for tapping phase B1 with SBI sequences, the favored beat.

Figure 4 shows the standard deviations of the asynchronies between taps and tones (or theoretical beat locations where tones were absent) as a function of tapping phase. In an ANOVA on these data, the two-way interaction approached significance, $F(3, 18) = 3.28$, $p < .07$, but separate ANOVAs on SBI and WBI sequences did not reveal significant main effects of tapping phase. Nevertheless, the similarity of the data for SBI sequences to those of Experiment 1 should be noted: Again, lowest variability was obtained for tapping in phase B3, not in phase B1. Thus there was no advantage for tapping with the favored beat of SBI sequences, even when that beat was self-chosen.

Discussion

Experiment 2 confirmed our assumption in Experiment 1 that the rhythmic structure of SBI sequences still favored the B1 beat, despite the absence

---

[8] One participant started tapping right away in four trials; these starting times were excluded. Two participants had empty cells (no trials for a particular tapping phase) that were filled in with their mean starting time in the ANOVA.

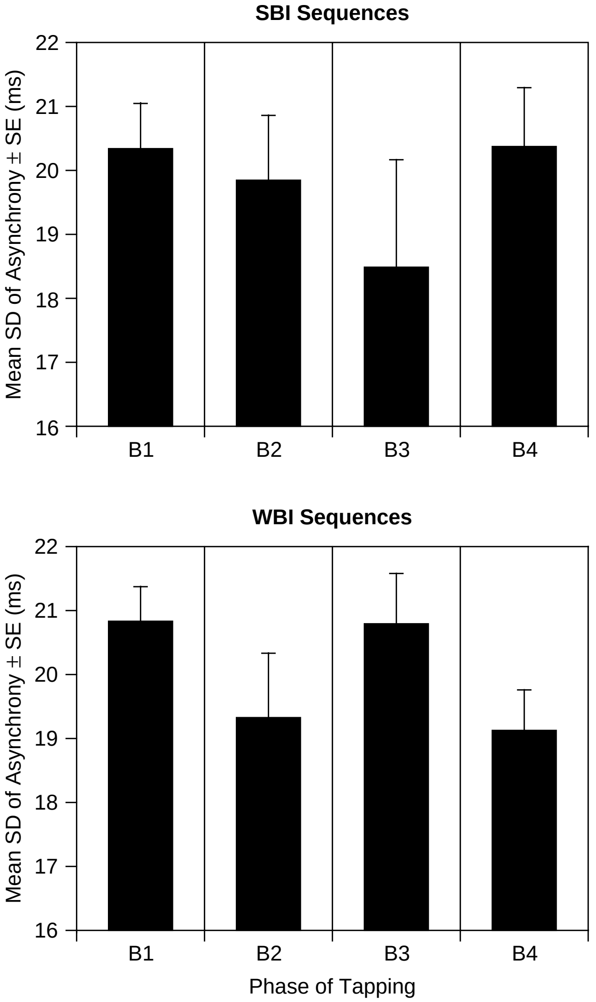

FIGURE 4. Mean standard deviations of asynchronies as a function of tapping phase in Experiment 2, with standard error bars.

of the recurring empty 800-ms interval that was present in the original materials of Povel and Essens (1985) and Patel et al. (2005). Clearly, when participants could choose the beat to tap on, they preferred to tap on the favored beat of SBI sequences, though not always. Sometimes participants tapped in a different phase, which means they did not detect the favored beat. Because it can be safely assumed that participants, if given a choice, would always have chosen the favored beat phase in the original SBI sequences of Patel et al., the results do demonstrate that deletion of the empty 800-ms interval reduced the salience of the beat favored by the rhythmic grouping structure.

Demany and Semal (2002) found that a regular beat is very difficult to detect when it alternates with randomly timed but otherwise identical tones in a sequence. Their result implies that such sequences are perceived as nonmetrical. In the present SBI sequences, a regular sequence of tones (the favored beat, B1) was embedded in a metrical context where other tones occurred not randomly but at integral subdivisions of the beat period (multiples of the basic pulse period). Under these conditions, the presence of a regular sequence of tones evidently could be detected more easily, though not always. It is possible that participants would have detected the favored beat even more often if they had waited longer before committing themselves to a particular beat. (Remember that adopting a particular beat entails not being able to perceive alternative beats having different phases.)

The favored beat of SBI sequences also tended to be chosen more quickly than other possible beat phases. However, there was still no tendency for tapping variability to be lower for tapping on the favored beat. Rather, variability again tended to be lowest for tapping in phase B3 with SBI sequences, although this did not reach significance and thus was true only for some participants.

McAuley and Semple (1999) considered three alternative models of beat induction: the original Povel and Essens (1985) model, which is based on negative evidence (silences and unaccented tones coinciding with the beat); a model based on positive evidence (accented and unaccented tones coinciding with the beat); and a hybrid model based on both kinds of evidence. The first two models correspond roughly to the statistics in Tables 2B and 2A, respectively, although the models allow for different weighting of different types of events. McAuley and Semple found that, at a tempo corresponding to that of the present sequences (a 200-ms basic pulse), the hybrid model fitted musicians' beat finding results best, whereas at slower tempi the positive evidence model fitted best. The present beat-finding results for SBI sequences seem to reflect mainly the percentage of beats coinciding with tones and thus support a positive evidence model. (The apparent preference for B4 in WBI sequences remains unexplained.) However, the variability results for SBI sequences, which match those of Experiment 1, suggest a different positive evidence model in which the percentage of tones that is accented matters (Table 2C).

Experiment 2 also demonstrated a strong tendency in most participants to choose the starting phase of a sequence as the beat, even after having listened to the sequence for a number of seconds. This is in agreement with the findings of Toiviainen and Snyder (2003) and suggests that the first tone heard tends to be perceived as a downbeat. Two participants, however (author BHR being one), were barely affected by the starting phase, which suggests that this is not an obligatory effect.

Experiment 3

Experiment 3 was conducted to examine further the curious tendency for tapping on B3 to be more stable than tapping on B1 (the favored beat, which was consistently marked by tones) in SBI sequences, in both previous experiments. Although this tendency did not reach significance in either experiment and thus was exhibited only by some participants, the fact that it was found twice suggests that it is reliable for those participants.

There are two ways of interpreting this tendency. One is suggested by the analysis in Table 2C: B3 may support a stronger perceived beat than B1 because 81% of the tones coinciding with B3 carry a grouping accent, whereas only 63% of the tones coinciding with B1 do. If B3 really provides a stronger beat than B1, then off-beat tapping relative to B3 should also be easier and more accurate than off-beat tapping relative to B1. This prediction was tested in Experiment 3. The hypothesis is not very plausible, however. A stronger beat can be conceptualized as a stronger internal resonance or oscillation (Large, 2000, 2001; Large & Kolen, 1994). Oscillator models, however, predict that internal resonance should increase in proportion with the percentage of beat locations that are marked by tones, regardless of grouping accents, and this favors B1 over B3. Therefore, the tendency of on-beat tapping to be less variable for B3 than for B1 is probably not a reflection of relative beat strength.

The other interpretation goes as follows. The stability of on-beat tapping may be determined not only by the strength of the beat itself but also by its rhythmic context—the various other tones that subdivide the intervals between beats (Large, Fink, & Kelso, 2002; Repp, 2008). The greater stability of tapping on B3 than tapping on B1 in SBI sequences could then be due to the consistent presence of tones in anti-phase with the imposed beat (viz., in phase B1). In other words, even though the B1 tones are not perceived *as the beat* when B3 is the physically or mentally imposed beat, they may nevertheless provide important temporal references for tapping on the B3 beat. The resulting stability of tapping can be viewed as a kind of *subdivision benefit* (Repp, 2003): Physical subdivision of a beat period increases the stability of on-beat tapping. If this is the correct interpretation, then different predictions follow for off-beat tapping. In off-beat tapping the beat tones themselves are the primary temporal references; otherwise, the tapping would not be conceptualized correctly as off-beat by the participant. Therefore, if B3 is not really a stronger beat than B1, off-beat tapping relative to B3 should be more difficult and more variable than off-beat tapping relative to B1.

Consider especially tapping in anti-phase with the imposed beat. When tapping is in anti-phase with B1, the taps are in phase B3, but B1 is the temporal and cognitive reference. When tapping is in anti-phase with B3, the taps are in phase B1 (which is always marked by tones), but B3 nevertheless is the reference, or else it would not be anti-phase tapping. Thus, an interaction is predicted: Although in-phase tapping tends to be less variable with B3 than with B1 (as suggested by Experiments 1 and 2), anti-phase tapping should be easier and less variable with B1 than with B3.

Although it would have been sufficient to consider in-phase and anti-phase tapping tasks to test these predictions, Experiment 3 also included two other off-beat tapping tasks: tapping upbeats or afterbeats relative to the imposed beat. We expected that these fairly challenging tasks, too, would be easier when B1 was the imposed beat than when it was B3.

Method

PARTICIPANTS

The participants were five young paid volunteers (three women) and two of the authors (BHR, JRI). Only BHR had participated in the previous experiments; all others were newly recruited. One was a highly trained graduate student percussionist; two others (JRI and one undergraduate) also had extensive percussion training. The remaining participants were two violinists and one cellist who played in the undergraduate Yale Symphony Orchestra.

MATERIALS

The synchronization-continuation paradigm of Experiment 1 was used. The sequences were a subset of the ones used in Experiment 1, namely SBI sequences with an imposed beat in phase B1 or B3, as well as an isochronous sequence. Eight blocks of 12 sequences each were formed, such that each type of sequence appeared four times, in random order. A nonrandom practice block of 12 sequences contained four isochronous sequences followed by four SBI sequences with an imposed B1 beat and finally four SBI sequences with an imposed B3 beat.

The prescribed tapping phase relative to the imposed beat in each trial was indicated in a changing visual display showing four quarter notes in a measure with a 4/4 time signature, with an arrow below one of the quarter notes. (Although the display was initially confusing to participants—it would have been better to show a single note followed by three rests, and perhaps a smaller note value—they soon understood that the notes were meant to indicate subdivisions of the beat period.) In each block, each of the four tapping phases (referred to in the following as T1, T2, T3, and T4) occurred once with an isochronous sequence, once with an SBI sequence with an imposed B1 beat, and once with an SBI sequence with an imposed B3 beat. The order of the tapping phase conditions was random, except that successive trials always had a different tapping phase. In the practice block, the order of tapping phases was T1, T3, T4, T2 for each sequence type.

PROCEDURE

Participants first completed the practice block, with the experimenter (BHR) sitting next to them (except for author JRI, who ran himself in California using very similar equipment). BHR carefully coached them through each task and made sure that it was understood, performed correctly (if not always accurately), and not accompanied by overt movements such as tapping on the beat with the other hand or the foot. Slight movements of the head, mouth, or body were not proscribed, as they are natural accompaniments of trying to keep an imposed beat in mind. The position of the arrow in the display panel, indicating the tapping phase, changed at the beginning of each trial, and the panel remained in view throughout. Participants started each trial by pressing the space bar and started tapping on (T1) or immediately after (T2, T3, T4) the third low tone. In view of the length and difficulty of the experiment, participants were instructed to repeat a trial only if they found themselves tapping in an incorrect phase relative to the imposed beat, but not necessarily if they felt that the imposed beat had shifted to a different phase. The crucial importance of tapping in the correct phase relative to the imposed beat was emphasized. BHR and JRI completed all eight test blocks in two sessions. Four other participants completed seven blocks each (three in the first session and four in the second session), and one (the graduate student percussionist) completed only six, due to time constraints. One block took about 17 minutes.

Results

No participant had to be excluded because of inadequate performance. Repetition occurred in a total of 20 trials (3.3%), some of which were attempted repeatedly, indicating great difficulty. Only the final attempt was analyzed. In four trials the tapping phase was incorrect from the start, suggesting a misreading of the tapping phase display, and four other trials had a large number of taps missing, probably because the taps were too gentle to be registered by the electronic pad. These eight trials were excluded.

Other errors (phase slips or drift) occurred only during the continuation part of SBI sequences, where the imposed beat had to be maintained mentally. Three participants committed no or hardly any errors. The distribution of errors is shown in Table 4. It can be seen that nearly all errors occurred in the T2 and T4 conditions, which had been expected to be challenging. Moreover, errors in these off-beat tapping conditions were more than three times as frequent when the imposed beat was B3 than when it was B1, as predicted. Two participants (BHR being one) had consistent difficulties with the B3/T2 and B3/T4 conditions. In the B3/T4 condition, BHR's actual tapping phase relative to B3 was often T2, whereas in the B3/T2 condition it was sometimes T4. Given that BHR, to the best of his knowledge, always maintained the subjective beat phase, this pattern suggests that his mentally imposed beat had slipped from B3 to B1. Other participants'

TABLE 4. Distribution of Errors Across Continuation Conditions in Experiment 3.

Designated relative tapping phase	Number of errors (actual relative tapping phase)											
	Imposed beat = B1						Imposed beat = B3					
	T1	T2	T3	T4	?	Total	T1	T2	T3	T4	?	Total
T1	—	0	1	0	0	1	—	0	0	0	0	0
T2	1	—	0	0	4	5	0	—	4	5	5	14
T3	0	0	—	0	0	0	0	0	—	0	0	0
T4	0	0	1	—	3	4	1	10	1	—	7	19

Note: The question mark refers to phase drift and multiple phase slips.

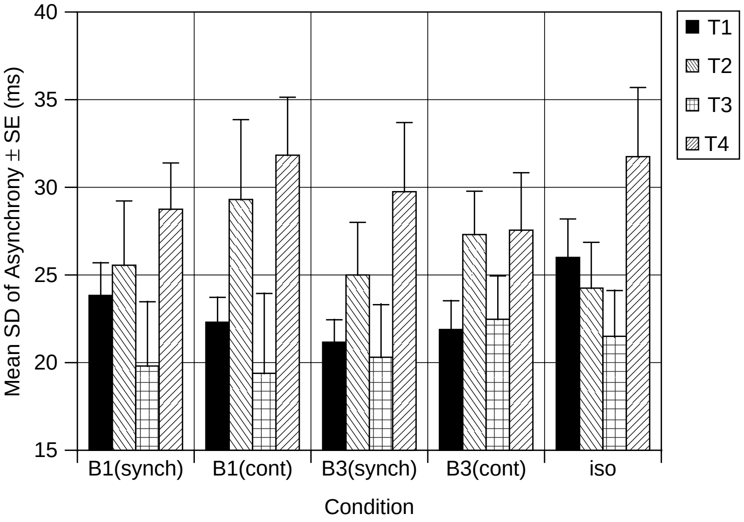

FIGURE 5. Mean standard deviation of asynchronies in the various conditions of Experiment 3, with standard error bars. B1, B3 = phase of imposed beat; synch = synchronization, cont = continuation; iso = isochronous sequence; T1, T2, T3, T4 = tapping phase relative to imposed beat.

errors did not show such a clear pattern of phase slips. All error trials (7.2% of all trials) were omitted from analysis. As a result, there were three empty cells in the data matrix, which in the ANOVAs reported below were filled by duplicating the corresponding data from the synchronization part (the first half) of the same tapping condition.

The standard deviations of the asynchronies in the various conditions are shown in Figure 5. We focus first on the predicted interaction between imposed beat and tapping phase in the T1 and T3 conditions. These conditions were analyzed in an ANOVA with the variables of imposed beat (B1 vs. B3), tapping condition (synchronization vs. continuation), and designated relative tapping phase (T1 vs. T3); the isochronous condition was not included. The predicted interaction was indeed significant, $F(1, 6) = 12.89$, $p < .02$. It can be seen in Figure 5 that tapping in phase with B3 (B3/T1) was less variable than tapping in phase with B1 (B1/T1), as in Experiments 1 and 2, but tapping in anti-phase with B3 (B3/T3) was more variable than tapping in anti-phase with B1 (B1/T3). In addition, the analysis yielded a significant main effect of tapping phase, $F(1, 6) = 7.36$, $p < .04$, with in-phase tapping (T1) being more variable overall than anti-phase tapping (T3), and a significant interaction between imposed beat and tapping condition, $F(1, 6) = 8.95$, $p < .03$, because absence of explicit beat markers during synchronization led to a decrease in variability for B1 but to an increase for B3 relative to continuation tapping. The triple interaction was not significant. Nevertheless, separate ANOVAs were conducted on the synchronization and continuation conditions to confirm that the interaction between imposed beat and tapping phase was stronger in the latter condition, $F(1, 6) = 14.38$, $p < .009$, than in the former, $F(1, 6) = 6.39$, $p < .05$. These results are quite in accord with predictions.

Clearly, variability in the T2 and T4 conditions was greater than in the T1 and T3 conditions, although there were considerable individual differences. An overall ANOVA on all four tapping phases (still not including the isochronous sequence condition) yielded a significant main effect of tapping phase, $F(3, 18) = 7.82$, $p < .02$. A separate ANOVA on the T2 and T4 data yielded no significant effects, however. Thus, the prediction that variability in these conditions would be greater for B3 than for B1 was not confirmed. Note, however, that the increased error rate for B3/T2 and B3/T4 in some participants was in accord with the prediction, and that some variability data from the continuation condition were missing for these participants because of their high error rates.

As in Experiment 1, variability of tapping in phase with an isochronous sequence (T1) tended to be *larger* than the variability of tapping on the favored beat (B1/T1), and was significantly larger than tapping on the nonfavored beat (B3/T1), $F(1, 6) = 10.86$, $p < .02$. Interestingly, these differences were shown most clearly by the three percussionists among the participants. The only participant who showed a difference in the opposite direction was BHR, whose very low standard deviation in the isochronous T1 condition (14 ms) probably reflected his extensive experience with simple synchronization tasks. The variability differences among the different tapping phase conditions were similar to those in the B1 synchronization condition, except that the T2 condition was as stable as the T1 condition here. Tapping in anti-phase (T3) was least variable.

An ANOVA on the asynchronies yielded no significant effects. The grand mean asynchronies in the various conditions ranged from −5 to −25 ms, but there were large individual differences.

Discussion

Experiment 3 replicated once again the counterintuitive finding of Experiments 1 and 2 that tapping on the favored beat of SBI sequences (B1/T1) tends to be more variable than tapping on an imposed B3 beat (B3/T1). Even though the difference never reached statistical significance because not all participants showed it, it was consistent across three experiments. A new finding is that tapping in anti-phase with B1 tends to be less variable than tapping in anti-phase with B3. This interaction between beat phase and tapping phase was predicted on the assumption that on-beat tapping is governed mainly by rhythmic context (explicit subdivisions of the beat), whereas off-beat tapping depends more strongly on tones coinciding with the beat, which are temporally closer to off-beat than to on-beat taps. Anti-phase tapping was also less variable than in-phase tapping with isochronous sequences, which can be explained by the smaller temporal distance of the taps from the beat tones that serve as the temporal references (see also Repp & Doggett, 2007). The results thus suggest that B1, the favored beat that was consistently marked by tones, was a stronger beat than B3, after all; however, B3 had a firmer rhythmic context, precisely because the regularly occurring tones in the B1 phase were part of that context.

The results of Experiment 3 also give some support to the related prediction that off-beat tapping in phases T2 and T4 should be more difficult and more variable when the imposed beat is B3, compared to B1. Two participants had serious difficulties with the B3/T2 and B3/T4 tasks when no beat markers were present, which is in agreement with the prediction. On correct trials and on trials with beat markers, however, these tasks did not exhibit greater variability than the B1/T2 and B1/T4 tasks or, for that matter, than the T2 and T4 tasks with isochronous sequences (where the beat was B1). This suggests that as long as the B3 beat could be maintained mentally, producing the relative tapping phases posed no serious problem to the present participants. Clearly, however, the T2 and T4 tasks were generally more difficult than the T1 and T3 tasks, and T4 (tapping upbeats to the imposed beat) tended to be more variable than T2 (tapping afterbeats), most clearly so in isochronous sequences.

For the majority of participants, especially the three percussionists, tapping in phase with an isochronous sequence was more variable than tapping in phase with B1 and, especially, B3 in the rhythm sequences. One reason may be that in-phase synchronization of actions with a simple beat, which is such a popular task in psychological studies of timing, is relatively rare in actual music making. Furthermore, percussionists typically provide the beat for other musicians rather than follow their beat. It is also possible that the percussionists did not find the simplest task challenging enough and therefore did not try hard enough to be maximally accurate. Furthermore, the low pitch and loudness level of the beat marker tones could have played a role.

General Discussion

Tracking a Beat Within a Metrical Grid

The main hypothesis tested in Experiment 1 was that the beat favored by a rhythmic structure through consistent marking by tones and frequent grouping accents (Povel & Essens, 1985) would afford tighter, less variable sensorimotor coupling than a beat that has less structural support. Our earlier results (Patel et al., 2005) had confirmed that hypothesis, but there we had compared beats that were strongly supported by the structure of the rhythmic sequences (SBI and WBI) and differed only in degree of support. The present study instead compared a favored beat that was less strongly supported (perhaps more comparable to the favored beat of the previous WBI sequences) with other beats that were physically or mentally imposed and had no specific structural support at all, apart from the basic pulse implied by the rhythmic sequences. Here we failed to find a consistent difference. This finding would be unremarkable if synchronization had been much poorer overall, but this was not the case: Musically trained participants were able to impose and tap accurately with any arbitrary beat, as long as it coincided with the basic pulse.[9] The coefficients of variation

---

[9] And, we should add, as long as it had an 800-ms period. Other beat periods were not investigated here, but are conceivable for these rhythms. A 400-ms period should present no difficulty, and similarly for a 1600-ms period. We do not know at this time, however, whether beats with 600-ms or 1000-ms periods (which imply different meters) could be imposed successfully on these rhythms.

of the asynchronies were between 2% and 2.5% across all conditions; this represents very stable synchronization performance.

Using a beat-finding task, Experiment 2 demonstrated that the favored beat of SBI sequences could still be detected most of the time, even though its structural support had been weakened by deletion of the empty 800-ms interval from the original SBI sequences. So, a difference in perceptual salience still existed between the favored beat and the other imposed beats; yet this difference was not reflected in synchronization performance. Some methodological differences from our earlier study (Patel et al., 2005) could possibly account for the difference in results. The method of beat induction differed: Patel et al. employed a simple induction sequence that ended when the sequence started, whereas in the present Experiments 1 and 3 there was not only a (probably superfluous) induction sequence but also an explicit beat marker during the entire first half of each long sequence; thus, it could be argued that beat induction was stronger than previously, and that this reduced the difference in relative strength between the favored beat and various other beats when they had to be maintained mentally. Still, this cannot account for the absence of any difference in favor of the favored beat, and moreover the tapping results of Experiment 2 resembled those of Experiment 1, even though the beats were entirely self-chosen in Experiment 2. Another, perhaps more important methodological difference was that in Patel et al. (2005) most sequences were not perceived as syncopated: Only WBI sequences were so perceived, and they occurred in the context of not only SBI sequences but also five other kinds of nonsyncopated sequence (i.e., in only one out of seven trials). In contrast, the majority of the sequences in Experiment 1 were perceived as highly syncopated relative to the beat (seven out of ten trials), and this may have improved participants' performance with arbitrary beats relative to the favored beat. In Experiment 3, however, the proportion of syncopated sequences was lower (four out of twelve of trials); yet, no difference in performance favoring the B1 beat emerged for on-beat tapping. Finally, it should be pointed out that the present participants also had more music training, on average, than the participants in our earlier study.

Thus, although there are some methodological differences that may have contributed to the different findings of the present study compared to Patel et al. (2005), we believe that the most important factor was the deletion of the empty 800-ms interval from the rhythmic sequences. The recurrence of this interval in the original sequences may have drawn participants' attention to the metrical level of the beat and away from lower metrical levels created by subdivision of the beat. In addition, it created a slow 3200-ms hypermetric period that was reflected in the timing of the taps, as noted by Patel et al. In contrast, the present sequences may well have focused participants' attention more on the subdivisions of the beat, particularly on the basic pulse having a period of 200 ms, but also perhaps on the intermediate 400-ms level. All rhythms defined a tight metrical grid at the basic pulse level with which taps could be coordinated in various ways once attention was focused on it. This might be described as "locking into the metrical grid." A metrical grid can also be generated mentally, as when tapping in different beat phases with isochronous beat sequences in Experiment 3, but it is most effective when it has (at least partial) physical support. Musicians often use pulse-based subdivision strategies when they have to execute complex rhythms in synchrony with a beat (Weisberg, 1993), a task complementary to the present one of tapping a beat in synchrony with a complex rhythm. A focus on the basic pulse is also essential in various kinds of non-Western music, such as African drumming and Balinese gamelan, where a complex interlocking of rhythms must be achieved.

Temporal References for Phase Correction

It is well known that sensorimotor synchronization cannot be achieved unless some form of error correction takes place (Vorberg & Wing, 1996). The traditional view is that phase correction is based on perception of asynchronies between taps and tones. However, various recent findings (reviewed in Repp, 2005b) have led to the alternative view, originally suggested by Hary and Moore (1985, 1987), that sensorimotor synchronization involves phase resetting with reference to multiple reference points provided by preceding tones and taps. For example, Repp (2008) found that the corrective response to an asynchrony created by artificial perturbation of a beat tone depends not only on that asynchrony but also on subsequent tones that subdivide the inter-beat interval. Conversely, perturbation of a subdivision tone can result in a phase correction response, even though there is no tap coinciding with that tone. In other words, the temporal placement of each tap depends on its immediate rhythmic context.

This insight can help explain the curious tendency, observed in Experiments 1 and 2, for tapping to be less variable in phase B3 than in phase B1. When the task was to tap on B1, the tones that were consistently present in phase B1 constituted previous target tones. When the task was to tap on B3, they provided rhythmic context and were temporally closer. It is well known that interval-based responses are less variable for short than for long intervals. Therefore, a phase reset based on a close tone (within limits) will be more accurate than one based on a distant tone. Tapping on B3 thus benefited from a stable rhythmic context in the immediate past, whereas tapping on B1 had to rely on a stable context in the more distant past as well as on more sparse and unpredictable immediate context.

In Experiment 3, however, when anti-phase tapping (T3) was required relative to B1 and B3, the situation was different. Off-beat tapping by definition has a preceding beat as its primary reference, but that beat can stabilize tapping only when it is marked by a tone. In the case of B1/T3, the consistent beat tones on B1 are temporally close and provide a solid reference for phase resetting. In the case of B3/T3, the beat tones are intermittent and less predictable, and the B1 tones are temporally distant. This eliminates the advantage of B3 tapping over B1 tapping that was seen for on-beat (T1) tapping. On the whole, then, the present results are consistent with the view that error correction in synchronization is based on context-based phase resetting rather than (only) on asynchronies between taps and target tones.

Metrical Interpretation

Beyond these specific results, the present study provides another demonstration that musically trained individuals can perceive the same rhythmic sequence in a variety of different ways, by imposing different metrical interpretations on it. Compared to other recent studies that had a similar goal (Repp, 2005b, 2007), the current experiments used much more complex rhythms and a different method of beat induction (previous metrical context containing explicit beat markers rather than mere verbal instructions or musical notation). When explicit beat markers were present in Experiments 1 and 3, they easily induced a corresponding beat in participants, not only because of their perceptual salience but also because instructions emphasized that they were to be considered the beat. However, it was the mental continuation of this beat that was of primary interest here. The noteworthy result is that participants had no difficulty with this task and performed essentially as well as when the explicit beat markers were present, despite the fact that the sequences sounded strongly syncopated in most cases. It can be concluded that on-beat tapping with syncopated rhythms is no more difficult than on-beat tapping with nonsyncopated rhythms, at least for musically trained participants and for the kinds of rhythmic sequences used here. The story may well be different for participants with less music training and for rhythms that more strongly favor a particular beat (cf. Patel et al., 2005).

The present results thus show that the metrical structure perceived in a rhythmic sequence is quite malleable (London, 2004), as long as there are no overpowering cues favoring a particular beat. Metrical interpretation can be seen as an example of perceptual multistability (Attneave, 1971) that is governed by both bottom-up (exogenous) cues and top-down (endogenous) processes (Leopold & Logothetis, 1999; Repp, 2007).

Author Note

This research was supported by NIH grant MH-51230 (Bruno Repp, P.I.). Additional support for BHR from NIH grants HD-01994 and DC-03782 (Carol Fowler, P.I.) and DC-03663 (Elliot Saltzman, P.I.) is gratefully acknowledged. The research was also supported by the Neurosciences Research Foundation as part of its research program on music and the brain at The Neurosciences Institute, San Diego, where ADP is the Esther J. Burnham Fellow and JRI is the Karp Foundation Fellow.

Correspondence concerning this article should be addressed to Bruno H. Repp, Haskins Laboratories, 300 George Street, New Haven, CT 06511-6624. E-MAIL: repp@haskins.yale.edu

References

ATTNEAVE, F. (1971). Multistability in perception. *Scientific American, 225,* 63-71.

DEMANY, L., & SEMAL, C. (2002). Limits of rhythm perception. *Quarterly Journal of Experimental Psychology, 55A,* 643-657.

DESAIN, P., & HONING, H. (1999). Computational models of beat induction: The rule-based approach. *Journal of New Music Research, 28,* 29-42.

FITCH, W. T., & ROSENFELD, A. J. (2007). Perception and production of syncopated rhythms. *Music Perception, 25,* 43-58.

HANNON, E. E., SNYDER, J. S., EEROLA, T., & KRUMHANSL, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 956-974.

HARY, D., & MOORE, G. P. (1985). Temporal tracking and synchronization strategies. *Human Neurobiology, 4,* 73-77.

HARY, D., & MOORE, G. P. (1987). Synchronizing human movement with an external clock source. *Biological Cybernetics, 56,* 305-311.

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

LARGE, E. W. (2001). Periodicity, pattern formation, and metric structure. *Journal of New Music Research, 30,* 173-185.

LARGE, E. W., FINK, P., & KELSO, J. A. S. (2002). Tracking simple and complex sequences. *Psychological Research, 66,* 3-17.

LARGE, E. W., & KOLEN, J. F. (1994). Resonance and the perception of musical meter. *Connection Science, 6,* 177-208.

LARGE, E. W., & PALMER, C. (2002). Perceiving temporal regularity in music. *Cognitive Science, 26,* 1-37.

LEOPOLD, D. A., & LOGOTHETIS, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences, 3,* 254-264.

LONDON, J. (2004). *Hearing in time: Psychological aspects of musical meter.* New York: Oxford University Press.

LONGUET-HIGGINS, H. C., & LEE, C. S. (1982). Perception of musical rhythms. *Perception, 11,* 115-128.

MCAULEY, J. D., & SEMPLE, P. (1999). The effect of tempo and musical experience on perceived beat. *Australian Journal of Psychology, 51,* 176-187.

PATEL, A. D., IVERSEN, J. R., CHEN, Y., & REPP, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research, 163,* 226-238.

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

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

REPP, B. H. (2002). Phase correction following a perturbation in sensorimotor synchronization depends on sensory information. *Journal of Motor Behavior, 34,* 291-298.

REPP, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior, 35,* 355-370.

REPP, B. H. (2005a). Rate limits of on-beat and off-beat tapping with simple auditory rhythms: 2. The role of different kinds of accent. *Music Perception, 23,* 167-189.

REPP, B. H. (2005b). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review, 12,* 969-992.

REPP, B. H. (2007). Hearing a melody in different ways: Multistability of metrical interpretation, reflected in rate limits of sensorimotor synchronization. *Cognition, 102,* 434-454.

REPP, B. H. (2008). Multiple temporal references in sensorimotor synchronization with metrical auditory sequences. *Psychological Research, 72,* 79-98.

REPP, B. H., & DOGGETT, R. (2007). Tapping to a very slow beat: A comparison of musicians and non-musicians. *Music Perception, 24,* 367-376.

SLOBODA, J. A. (1985). Expressive skill in two pianists: Metrical communication in real and simulated performances. *Canadian Journal of Psychology, 39,* 273-293.

SNYDER, J., & KRUMHANSL, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception, 18,* 455-489.

STEVENS, L. T. (1886). On the time-sense. *Mind, 11,* 393-404.

TOIVIAINEN, P., & SNYDER, J. S. (2003). Tapping to Bach: Resonance-based modeling of pulse. *Music Perception, 21,* 43-80.

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.

WEISBERG, A. (1993). *Performing twentieth-century music: A handbook for conductors and instrumentalists.* New Haven, CT: Yale University Press.

WING, A. M., & KRISTOFFERSON, A. B. (1973). The timing of interresponse intervals. *Perception and Psychophysics, 13,* 455-460.