BRILL

Timing & Time Perception 5 (2017) 244–259

TIME

brill.com/time

Repetition and a Beat-Based Timing Framework: What Determines the Duration of Intervals Between Repetitions of a Tapping Pattern?

Olivia Murton[1], Lauryn Zipse[2,*], Nori Jacoby[3] and Stefanie Shattuck-Hufnagel[4]

[1]Program in Speech and Hearing Bioscience and Technology, Harvard University, Cambridge, MA, USA

[2]Department of Communication Sciences and Disorders, MGH Institute of Health Professions, Boston, MA, USA

[3]The Center for Science and Society, Columbia University, New York, NY, USA

[4]Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA

Received 17 May 2017; accepted 6 November 2017

Abstract

The production of speech and music are two human behaviors that involve complex hierarchical structures with implications for timing. Timing constraints may arise from a human proclivity to form 'self-organized' metrical structures for perceived and produced event sequences, especially those that involve repetition. To test whether the propensity to organize events in time arises even for simple motor behaviors, we developed a novel experimental tapping paradigm investigating whether participants use the beat structure of a tapped pattern to determine the interval between repetitions. Participants listened to target patterns of 3, 4, or 5 events, occurring at one of four periodic rates, and tapped out the pattern 11 times, creating 10 inter-pattern intervals (IPIs), which participants chose freely. The ratio between mean IPI and mean inter-tap interval (ITI) was used to measure the beat-relatedness of the overall timing pattern; the closer this ratio is to an integer, the more likely the participant was timing the IPI to match a multiple of the target pattern beat. Results show that a beat-based strategy contributes prominently, although not universally, to IPI duration. Moreover, participants preferred interval cycles with even numbers of beats, especially cycles with four beats. Finally, the IPI/ITI ratio was affected by rate, with more beats of silence for the IPI at faster rates. These findings support the idea that people can generate a larger global timing structure when engaging in the repetition of simple periodic motor patterns, and use that structure to govern the timing of those motor events.

Keywords

Beat, meter, motor, rhythm, tapping, prediction, repetition

---

[*] To whom correspondence should be addressed. E-mail: lzipse@mghihp.edu

DOI: 10.1163/22134468-00002095

1. Introduction

Repetition invites temporal organization and grouping. In perception, the effects of repetition are widely acknowledged. A number of perceptual phenomena illustrate the human tendency to impose a metrical structure on patterns of repeated events, particularly in the auditory domain. A commonly cited example is the 'tick-tock' phenomenon, when a listener perceives a ticking clock to produce a sequence of two-event units; the two events sound slightly different, even though each repeating tick is the same (Bolton, 1894). When a motor behavior is repeated, there is an opportunity to organize the repetitions in time. However, less is known about the effects of repetition in production tasks compared to perceptual ones. The tendency of human actors to impose a metrical structure on patterns of repeated movements, particularly in conditions without a metronome, where the participant is free to impose a self-selected temporal pattern on the repetitive behavior, has not been widely explored. This paper reports on a study of how participants organize the timing of repetitions of simple tapping patterns.

Musical meter is often defined as an organizing temporal principle that creates expectation. Expectations may be generated from an array of cues, not only for music but for a wide variety of sensory inputs. Recent work points to a model of how expectancies for simple isochronous sequences interact with the associated sensory input, and highlights the importance of expectation in perception (Di Luca & Rhodes, 2016). Cues may simply be weak and strong beats (simple accent patterns, e.g., Keller & Repp, 2005), or at a more complex level, accent patterns may promote grouping events together, and perhaps generate recurrent grouping patterns (e.g., Povel & Essens, 1985), with multiple hierarchical levels (e.g., Snyder & Krumhansl, 2001). In this vein, London (2002) defines meter as 'a stable and recurring pattern of hierarchically structured temporal expectations' (p. 529). The hierarchical nature of the expectations generated by meter has been demonstrated experimentally, using both perceptual (Palmer & Krumhansl, 1990) and production (Keller & Repp, 2005) paradigms, and the resulting expectations allow a listener to anticipate future auditory events.

A critical idea concerning meter is that it is a mental construct. In fact, musicians can 'hear' the same passage in a different meter if they are instructed to do so. As Repp (2007) explains, different possible metrical interpretations may serve as attractors. If top-down influences, such as instructing the listener to hear the music in a particular meter, are minimized, the interpretation with the most bottom-up factors in its favor is salient. Repp (2007) demonstrated this metrical 'multistability' by prompting musicians to apply different metrical interpretations when tapping along to the same melodic sequence of tones. Tapping accuracy was affected by the perceived meter, demonstrating that meter can play an important role in the production of events within an auditory context.

As a mental framework that can be greatly affected by top-down processes, meter is closely related to the idea of subjective rhythmicization. In this phenomenon, the stimulus itself does not fully determine the grouping of a sequence of auditory events; instead, a structure emerges from an interaction between the sensory input and top-down processes. Bolton (1894) described subjective grouping, in which a listener has a tendency to hear isochronous sequences of events as occurring in groups of twos, threes, or fours. More recently, electrophysiological data, in the form of ERPs, has served as evidence for a default binary metric structure that appears to be attention-based (Brochard et al., 2003; see also Nozaradan et al., 2011).

Bolton (1894) suggested that the ability to form subjective groups from a string of isochronous auditory events is affected by the presentation rate, and he identified limits. At faster rates, with events separated by less than 115 ms, listeners no longer report hearing events in groups, and at slow rates, with events separated by more than 1581 ms, listeners tend to hear individual auditory events as ungrouped. Repp (2006) explains that at faster rates, listeners cannot track events individually, and therefore are unable to group them. Conversely, when inter-event intervals increase to approximately 1800 to 2000 ms (only slightly slower than the limit identified by Bolton), processing changes: synchronization to a sequence of isochronous auditory events becomes difficult, variability of the inter-tap intervals and of the asynchronies between the stimulus events and associated taps becomes high, and people tend to react to the events rather than anticipate them (Repp, 2006). This synchronization limit seems to be due to auditory working memory capacity: after approximately 2000 ms, the preceding event is no longer ‘present’ in auditory working memory and the listener is using a deliberate strategy of interval estimation (London, 2002; Repp, 2006).

Keller and Repp (2005) define metric frameworks as “cognitive/motor schemas that guide musical rhythm in perception and action” (p. 293). The propensity to use these schemas to organize behavior in time has implications for complex motor activities such as musical production and speech. Therefore, the tendency to impose temporal structure when repeating simple motor behaviors is of interest. In the present study, we investigated what governs the amount of time between repetitions when simple, discrete tapping patterns are repeated. One possibility is that the duration of the interval between pattern repetitions is part of a larger timing framework, and therefore should be related to the timing structure of the repeated pattern. Another possibility is that pattern repetitions are not strongly temporally organized. This study was designed to investigate these possibilities.

In evaluating the tendency to produce a metrical structure when repeating simple tapped patters, we identified two types of intervals that are important: the inter-tap interval (ITI) within a target pattern, and the inter-pattern interval (IPI) between repetitions of that pattern (Fig. 1a). We tested whether the preferred IPI

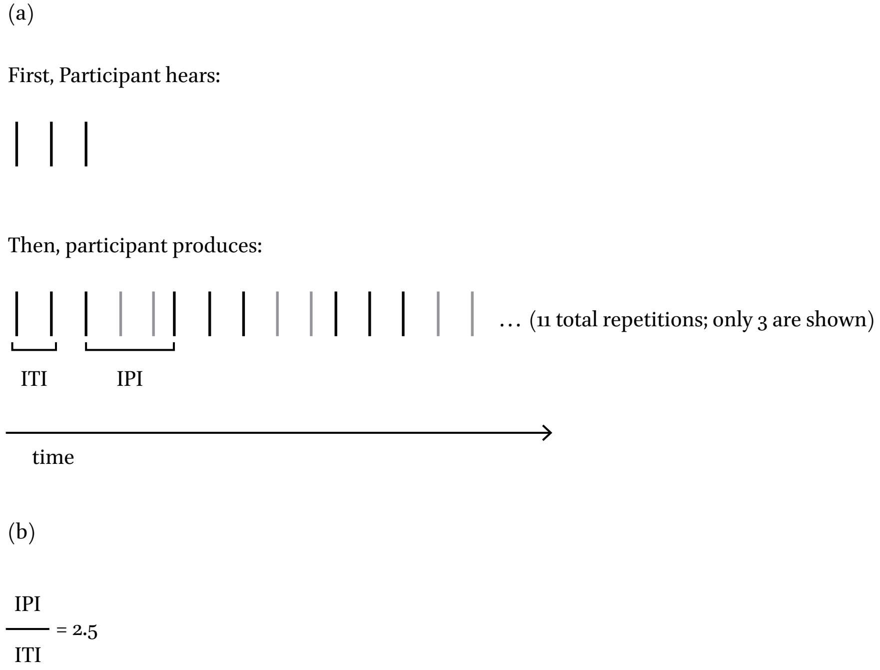

Figure 1. (a) An example of a single trial in a three-click condition. The participant hears the pattern once and then taps it back repeatedly. The stimulus clicks and response taps produced by the participant are both indicated in black, while silent 'beats' are indicated in gray. ITI is the inter-tap interval, while IPI is the inter-pattern interval. In this experiment the target ITI was specified by the target stimulus but the IPI was generated by the participant, since the target pattern for each trial was only presented once. Note that this example shows just one possible IPI timing behavior, i.e., one that is not an integer multiple of the beat. See text for further discussion. (b) An example of an IPI/ITI ratio. This is the ratio associated with the illustration in (a).

would be an integer multiple of the beat structure suggested by the ITI of the target pattern, consistent with the use of a higher-order structure (Fig. 1b). We used simple target patterns with different numbers of events (3, 4, or 5 clicks) and four different rates, produced in long sequences, to determine whether participants tailored their IPIs based on the inter-event intervals in the target patterns. An important distinction between the two types of intervals (ITI and IPI) is that, from the participant's perspective, the ITI was constrained by their attempt to replicate the timing that they heard when the target was presented, while the IPI was freely generated.

Two characteristics of the target pattern were identified as potential influences on the IPI: rate, and the number of events. Rate has an effect on beat salience, as well as on how rhythms are performed and perceived (Parncutt, 1994; Repp, 1995). Furthermore, perceptual constraints on the upper and lower durational limits at which listeners perceive a beat restrict how far beats can be sub-divided (e.g., duple vs. triple subdivisions), which may in turn affect metrical interpretation (London, 2002). We used four different click rates to increase the chances that at least some of the targets would lend themselves to perception (and, thus, generation) of a beat structure. Varying the rate across different trials also eliminated the possibility that the inter-tap interval in the target pattern would, by chance, correspond to a participant's preferred fixed IPI; if this had been the case, the resulting behavior would have appeared to support a beat-related IPI while in fact reflecting the participant's fixed preference. Varying the rate of the target pattern was a way to reveal whether the IPI varied with the rate of the target events, or remained fixed across rate changes.

A second factor that might affect IPI duration is the number of events in the target pattern. For example, if there is a preference for four-beat performance units (hereafter referred to as 'cycles') and the target pattern contains three events, then participants might tend to leave one beat of silence between repetitions, for a total of four beats per cycle. In contrast, for a pattern that contains four events, participants might leave no beats between repetitions, to maintain the four-beat structure. There is some evidence that Western listeners, at least, prefer to organize events into groups of two or four. For example, Repp (2007) presented isochronous monotonic sequences to listeners and asked them to tap on every *n*th tone, with *n* varying from 2 to 9. Accuracy was best for groupings of 2, 3, 4, and 8, and worst for groupings of 5 and 7. Based on subsequent experimental findings, Repp (2007) suggested that a group of two or three tones can be subitized, and that larger divisions may be based on these groups (and integer multiples of them).

Thus, our research questions about the how participants organize the timing of repetitions of tapping sequences included the following: (1) Do people produce IPIs that are integer multiples of a given ITI, suggesting they are using a beat structure related to the ITI? (2) How are IPIs affected by the tapping rate? (3) How are IPIs affected by the number of events in the tapping pattern? For example, do people prefer grouping structures that suggest a 'two attractor' or a 'four attractor'? To address these questions, we used a novel paradigm that is as assumption-free as possible regarding the IPIs, because participants were free to choose their IPIs.

2. Method

2.1. Participants

The participants were 30 undergraduates (11 female), ranging in age from 18–21 years (mean = 19.6, SD = 1.1), and recruited predominantly from various campus-based listservs (e.g., for dormitories). Twenty-three were native speakers of English; all spoke English fluently. Participants' self-reported level of musical experience varied widely, from none to many years (e.g., violin lessons since age 3 for a 19-year-old). Eight were deemed to have substantial musical experience, with ≥10 years that spanned into adulthood. Most of the reported musical experience was instrumental, but participants also reported experience with singing/voice (7 participants), dance (1), and rap (1). This study was approved by the MIT Committee on the Use of Humans as Experimental Subjects (COUHES).

2.2. Procedure

There were 12 target patterns: 3, 4, or 5 clicks presented at 90, 120, 180, or 240 clicks/minute (CPM) (see the Glossary for an explanation of the terminology used). In each trial, a pattern was played once using Praat software (Boersma & Weenink, 2014) on a laptop computer connected to speakers. The participant then used a pen like a drumstick to tap the pattern repeatedly on the tabletop. Responses were recorded with a high-quality microphone connected to the laptop, using Audacity software (Audacity Team, 2013). Participants were instructed to maintain the stimulus inter-click interval, while repeating the pattern at a natural pace (Fig. 1a). The experimenter stopped the participant after 11 repetitions of the pattern. A practice trial before the experiment used a pattern not included in the experiment, three clicks at 150 CPM. Participants completed four trials for each stimulus (48 trials), in a pseudo-random order such that no pattern was presented twice in a row.

Glossary

Click	stimulus event
Tap	response event
Beat	a mental representation of an event
Clicks per minute (CPM)	tapping rate (90, 120, 180, 240)
Clicks per pattern (CPP)	number of events (3, 4, 5)
Pattern	combination of one CPM and CPP
Cycle	tapped pattern + beats of silence
Trial	a response to one pattern (four per pattern, 48 per participant)
Repetition	one group in the response (11 per trial)

2.3. Scoring

Taps were marked by hand in Praat (Boersma & Weenink, 2014), following the rule that the point of highest amplitude was marked for each tap. Timestamps of taps were extracted in Praat and used to calculate ITIs and IPIs.

3. Analysis and Results

3.1. Histograms to Evaluate Use of a Beat-Based Strategy

A histogram was created for each of the 12 patterns, showing IPI/ITI ratios for all trials of each pattern for all participants (Fig. 2). Peaks at integer values indicate that the IPI is a multiple of the ITI for that pattern, implying that participants used a beat-based structure when determining IPIs.

To see whether the peaks tended to occur at certain integer values of IPI/ITI, we identified a window around integer IPI/ITI ratios within which a trial could be considered beat-like. This window allowed us to classify the first tap in a repetition as being on or off the beat in a binary way. We considered three possible ranges for this window: IPIs within 0.05 (i.e., 5%), 0.1, and 0.15 of an integer IPI/ITI value. Given the four stimulus rates we used (ITIs of 250, 333, 500, and

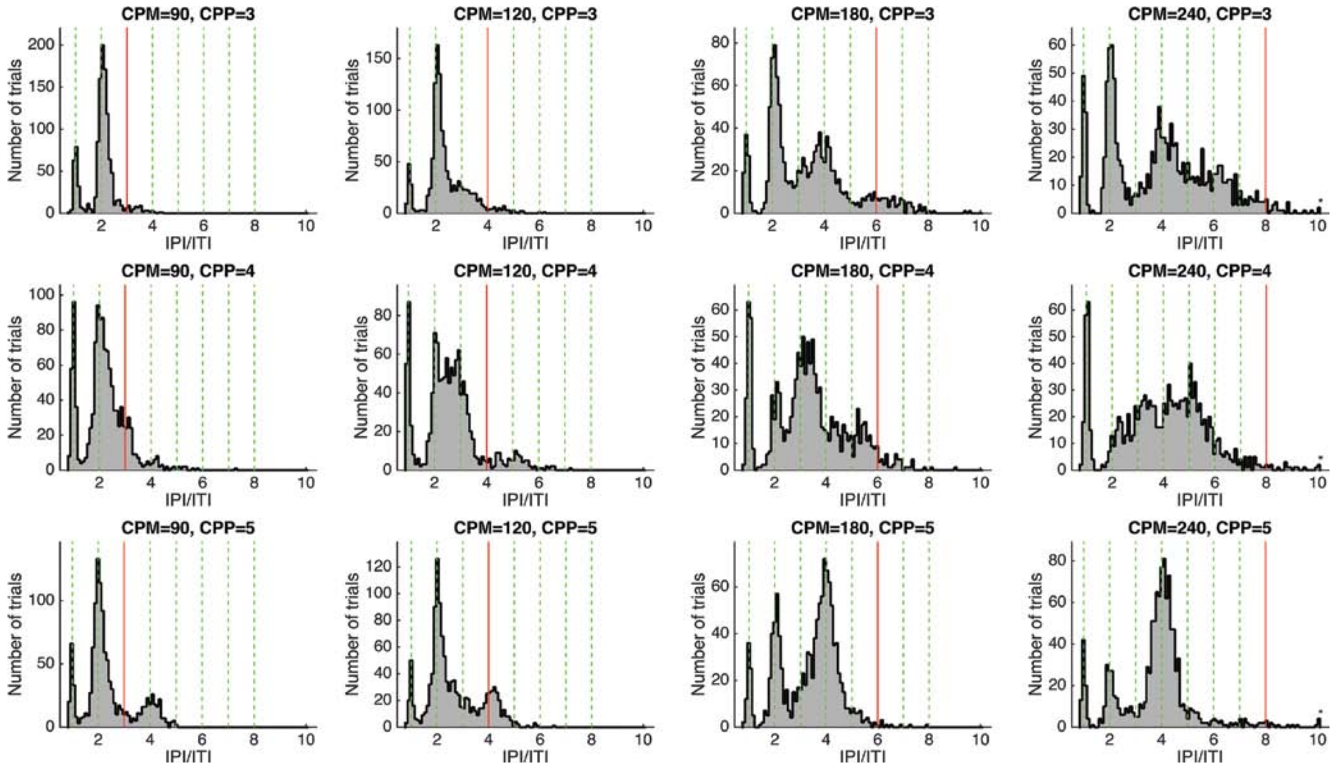

Figure 2. IPI/ITI ratios are shown across all participants, plotted for each unique pattern. Dotted vertical lines indicate integer ratios; solid lines indicate the 2000 ms synchronization limit. Note that the upper limit of the y axis varies across plots, as we are primarily concerned with relative peak heights within each plot.

667 ms), these three candidate ranges (5, 10, and 15%) span 12.5–33.3 ms, 25.0–66.7 ms, and 37.5c–100.0 ms in duration, respectively. These values cover the typical range of the negative mean asynchrony (NMA), the amount of time by which taps anticipate a pacing signal when a person attempts to entrain to an isochronous sequence of auditory events. This anticipatory tapping is a well-known phenomenon, and while the magnitude of the NMA varies based on experimental conditions and between individuals, it typically ranges from 0 to 100 ms (Aschersleben, 2002). Given that taps produced with an NMA are perceived by the person producing them to be 'on the beat,' this is a reasonable starting point for selecting a time window around a beat, within which events can be considered coincident with that event. Three of the authors (OM, LZ, SSH) listened to trials at the outer limit of each candidate threshold, and all agreed that trials 0.05 from an integer IPI/ITI ratio sounded unambiguously beat-like, while trials 0.15 from an integer ratio did not. Our judgments differed or were uncertain for trials where the IPI was 0.1 from an integer ratio. We therefore selected this as the threshold, which is supported by a rationale but ultimately set subjectively.

Across all of the IPIs, 34% fell within 0.1 of an integer IPI/ITI value, indicating that a beat-based strategy was used for a substantial portion of the total trials. Across subjects, the percent of trials falling within 0.1 of an integer ratio ranged from 11.3 to 89.7%. The distribution of individual participants' percentages of beat-based trials was unimodal and positively skewed (Fig. 3). Therefore, it was *not* the case that participants fell into two clear categories of beat-based and non-beat-based strategizers. Instead, all participants showed beat-like behavior on at least some trials, and not on others.

To determine whether more intervals (i.e., IPIs) fell within this 'beat-like' range than would be expected by chance, we ran binomial proportion tests comparing the proportion of intervals that fell within 0.1 of an integer IPI/ITI ratio to the test proportion 0.20 (i.e., the proportion of IPIs that would fall within 0.1 of an integer if the distribution of IPI/ITI ratios was flat). To avoid violating the assumption of independent observations, we ran a separate binomial proportion test for each repetition of each pattern across subjects, for a total of 480 tests (i.e., 12 unique patterns × 4 trials of each pattern × 10 IPIs/trial). Of these tests, 67.5% of the *p*-values attained significance with alpha set to 0.05 and false-discovery rate maintained below 0.05 (i.e., correcting the significance level to control for the multiple comparisons; Benjamini & Hochberg, 1995), indicating that more IPIs than would be expected by chance fell within 10% of an integer IPI/ITI value (Fig. 4). Thus, participants exhibited a tendency to use a beat-based strategy when determining IPIs.

It is noteworthy that all the histograms reveal clustering around an IPI/ITI ratio of 1, and most also show clustering around a ratio of 2 (Fig. 2). This suggests that participants often choose one of two strategies: tapping straight through one

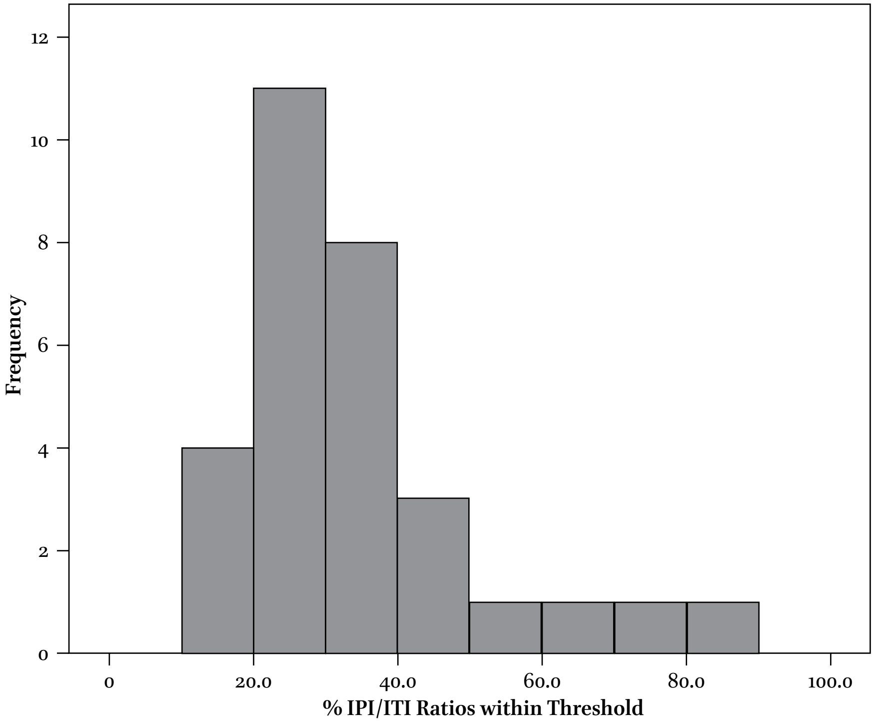

Figure 3. Distribution of individual participants' percentages of beat-based trials (i.e., trials where the IPI/ITI ratio fell within the threshold of 0.1).

pattern to the next (for an IPI/ITI of 1), or leaving one beat of silence before starting the next pattern (for an IPI/ITI of 2). (Note that the IPI/ITI ratio is always one more than the number of beats of silence in a cycle, because when the number of beats is n, this defines $n+1$ intervals. For example, when there are three beats of silence, this defines four inter-beat intervals, resulting in an IPI/ITI of 4.)

3.2. *Effects of Rate*

Considering the effects of rate on IPI/ITI ratio, it is evident that for a given number of clicks in the target pattern (i.e., across a single row in Fig. 2), faster rates show larger IPI/ITI values. For example, three-click patterns show peaks at IPI/ITI ratios of 1 and 2 across all four tapping rates. However, at the two fastest rates, these peaks are lower in amplitude, there are additional peaks around the ratio of 4, and more higher IPI/ITI values. A similar effect of rate is also discernible for the four-click and five-click patterns: peaks at lower IPI/ITI values decrease and additional peaks appear at higher IPI/ITI values as rate increases. It appears that at faster tapping rates, more beats of silence may be left between pattern repetitions. To further investigate the effects of rate, we averaged together all of the IPIs for

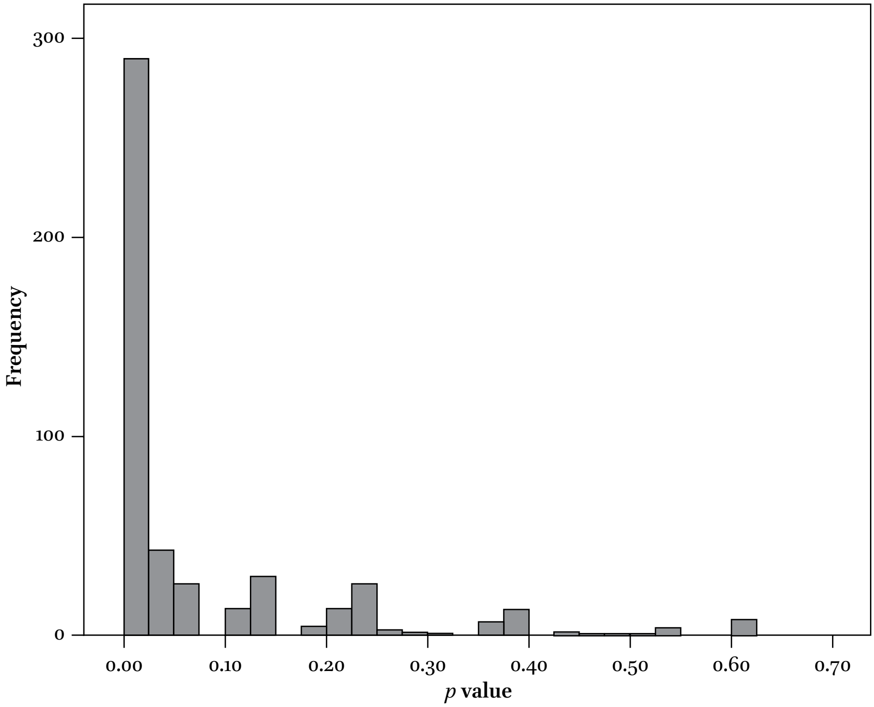

Figure 4. Distribution of p-values from binomial proportion tests comparing the proportion of IPIs that adhered to beat-based behavior, to the test proportion 0.20 (i.e., the proportion of IPIs that would fall in the beat-based range if the IPIs were evenly distributed).

each of the 12 target patterns. To account for the non-normality of the IPI distributions, we ran a Friedman test for each number of clicks per pattern. For each CPP, the main effect of rate was significant, with larger IPI/ITI ratios at faster rates (for CPP = 3, $\chi^2(3) = 58.8$, $p < 0.05$; for CPP = 4, $\chi^2(3) = 73.3$, $p < 0.05$; for CPP = 5, $\chi^2(3) = 69.6$, $p < 0.05$). Follow-up pairwise comparisons (Wilcoxon signed-rank tests) revealed that the differences between all rates within a CPP were significant with $p < 0.05$.

3.3. Grouping Structure Preferences

In assessing whether people prefer certain grouping structures, we noted that peaks at IPI/ITI ratios of 2 are generally higher than peaks at 1 for the three- and five-click patterns, while the opposite is true for the four-click patterns. This reflects a tendency to create cycles containing an even number of beats. Relatedly, there is notable clustering around a ratio of 4 in the five-click patterns (bottom row of Fig. 2). In these repetitions, participants tapped five beats and left three beats of silence, for a total cycle length of eight beats. In fact, at the two faster

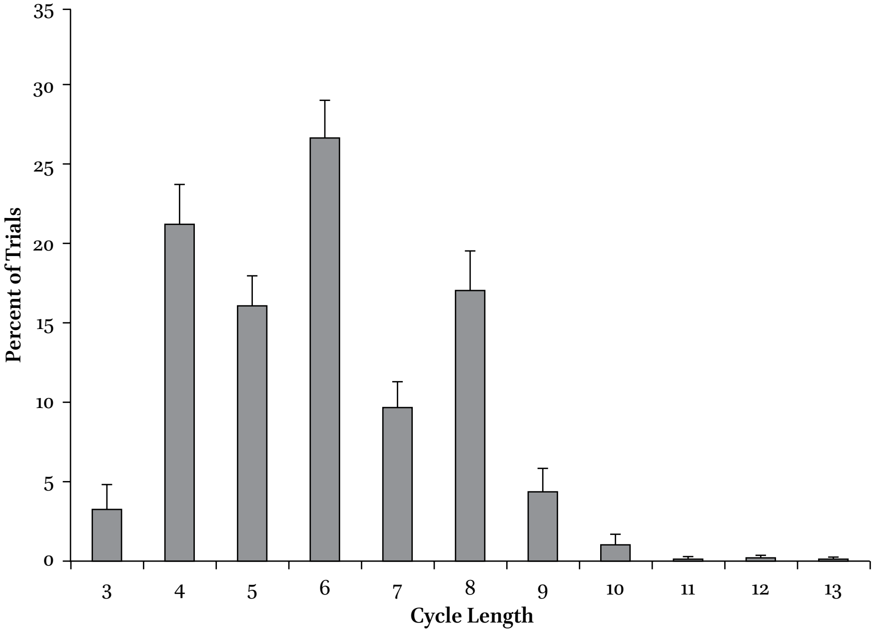

Figure 5. Cycle length in events (i.e., taps + beats of silence) across all participants and all conditions, plotted with integer bin sizes. Error bars are the standard error of the mean, computed from the mean histogram counts across participants.

rates, the peaks at a ratio of 4 (an eight-beat cycle) were considerably larger than those at a ratio of 2 (six-beat cycle), indicating a preference not just for even-number-of-beat cycles, but for cycle lengths that were multiples of four.

There is a particularly interesting preference for eight-beat cycles over seven-beat ones in the 90 CPM/five-click condition. Here, the length of silence in an eight-beat cycle is 667 ms * 4 = 2667 ms, which is more than the 2000 ms synchronization limit (London, 2002). A seven-beat cycle, by contrast, would have a length of silence of exactly 2000 ms and so should be easier for the participant. However, participants produced eight-beat cycles significantly more often than seven-beat ones [$t(29) = -31.74, p < 0.001$].

Collapsing across all patterns (all rates and all CPM) clearly shows the preference for cycles with even numbers of events (Fig. 5).

3.4. General Observations

While all participants displayed some beat-based behavior, other apparent strategies included a fixed duration across all stimulus groups/rates, different strategies for different rate-group combinations, and highly variable IPIs. It was *not* the case that a given participant consistently applied a single strategy, such as always leaving one beat of silence between repetitions. In fact, some participants occasionally showed mixed strategies even within a trial (although this was rare).

4. Discussion

This study tested the hypothesis that, when repeating a simple periodic tapping pattern, participants determine the time interval between successive repetitions (i.e., the inter-pattern interval, IPI) based on the beat structure of the pattern, i.e., the ITI within the pattern. Results revealed a tendency for people to use beat-based strategies when choosing an IPI. Prominent beat-based strategies included leaving no beats of silence between repetitions, leaving one beat of silence between repetitions, or leaving a number of beats of silence that creates a cycle with an even number of beats. A large majority of participants used multiple strategies; it was not the case that some people consistently used beat-based strategies while others consistently used other strategies. The proportion of trials on which a beat-based strategy was used varied across participants along a continuum, similar to the skewed unimodal distribution described in a study that examined beat perception (Grahn & McAuley, 2009). Thus we found no evidence for a bimodal distribution in the population, which would have suggested that some people reliably behave in a beat-based way and others do not.

We found that IPI was affected by tapping rate: at faster rates, participants tended to leave more beats of silence. The 'synchronization limit' of approximately 2000 ms seemed to be a constraint, though a violable one. As a consequence of this limit, more beats of silence could be 'fit in' to the IPI at faster rates, while at slower rates participants left fewer beats of silence between pattern repetitions. At inter-event intervals beyond about 2000 ms, events feel disconnected, perhaps because successive events no longer fall within the same envelope of working memory, and therefore are not perceived as part of the same higher-level unit (London, 2002). While the precise duration of this limit is task- and context-dependent, there is evidence that 1800 to 2000 ms is a relevant limit for tapping tasks such as the one used in this study. In 1:1 sensorimotor synchronization tasks (i.e., tapping along with every metronome click), the distribution of stimulus-tap asynchronies becomes bimodal when the inter-onset interval (IOI) of the stimulus stream is increased beyond 1800 ms. In addition to negative mean asynchronies (NMAs; i.e., taps anticipating the associated clicks, indicating prediction), positive mean asynchronies emerge as a common occurrence, reflecting that participants are no longer successfully predicting the occurrence of the clicks but instead are tapping in response to them (Repp & Su, 2013). When participants are specifically instructed to avoid tapping in response to the stimulus click, they are able to do so for IOIs up until approximately 3500 ms, producing a normal distribution of asynchronies. However, this distribution shows an increasingly large standard deviation as IOIs are increased (Repp & Doggett, 2007). These findings indicate that sensorimotor synchronization becomes difficult when IOIs exceed approximately 1800 ms, resulting in increased response time variability. Switching to a strategy of reacting to the pacing stimulus rather than anticipating it is a common strategy for reducing this variability, unless participants are instructed not to do so (Repp & Su, 2013). For IOIs longer than 1500–1800 ms, sensorimotor synchronization appears to demand more substantial attentional resources (Miyake et al., 2004), and is rated as being more difficult (Bååth & Madison, 2012). In the present study, participants were free to choose their IPI and it is therefore reasonable to assume they would stay within a 'comfortable' estimation time frame (i.e., not beyond 1800 ms), which is largely what we observed.

London (2002) notes that this approximately 2000 ms limit for hearing events as connected in time likely applies "outside of a metric hierarchy" (p. 537); within the framework of a metrical structure, he suggests that this limit might be expanded. Our findings support this idea, as illustrated by data from the five clicks per pattern (CPP) trials at the slowest rate, 90 clicks per minute (CPM). Here, a prominent strategy was to leave three beats of silence between pattern repetitions, yielding an eight-beat cycle. Notably, doing so resulted in an average IPI of 2667 ms, beyond the synchronization limit.

Data from the five CPP trials also indicate that there is a preference not just for creating cycles with even numbers of beats, but for creating cycles with units of four. At the two faster rates, leaving three beats of silence to create an eight-beat cycle was a more common strategy than leaving one beat of silence to create a six-beat cycle. It is possible that this preference for a four-beat cycle is related to the fact that all of the participants are from Western cultures in which musical structures based on four-beat measures predominate.

The propensity of participants to use beat-based strategies is especially worth considering because participants faced an open-ended task. We purposely used very simple stimuli and limited instructions. Participants imposed a temporal strictness on the task that we did not require. In light of this, our data support the hypothesis that participants use a larger organizing structure for their behavior when repeating a regularly occurring sequence of events. They tend to insert an amount of silence between pattern repetitions that makes the whole sequence temporally regular. This finding is consistent with participants using hierarchical metrical structure in motor control, but the use of such a structure cannot be definitively determined from the data presented here. Metrical structure has grouping and prominences. We have demonstrated that people have a tendency to group simple tapped patterns, but a limitation of this study is that we did not collect force data and therefore cannot analyze the existence and placement of prominences.

An intriguing question is *why* there is a tendency to use a larger temporal structure to organize behavior. One consideration is suggested by the preceding discussion of the 2000 ms constraint on perceiving a temporal relationship among successive events. Use of an organizing structure, whether hierarchical or not, may permit the cognition of timing relationships among sequences of events which span longer temporal distances. Such structures may also lie behind the ability to understand and produce systematic departures from temporal regularity. This may allow individual units to be expanded or contracted for expressive or structure-signaling reasons, and yet still be considered as part of the same overall temporally regular structure, because they are still one of the same type of structural component.

While the task used in the present study is simple, it is clear that multiple constraints interacted, resulting in a complex behavioral response. There is a structural similarity between the way that constraints interact in this simple task and in more complex behaviors such as music and speech. The results reported here may relate to the question of periodicity in speech, or perhaps more accurately, speech produced with temporal periodicity. Despite the sense that many listeners have that speech has a rhythmic structure that is periodic, it is growingly acknowledged that careful measurements have not revealed much temporal periodicity in typical speech (e.g., Cummins, 2012). Yet speakers can produce speech in what sounds like a periodic manner, and it would be of interest to determine how that periodicity is determined, i.e., the nature of the unit that is repeated regularly.

The current results raise the question of what happens to timing and rhythmic structure when a spoken element or pattern is repeated a number of times. There is some relevant evidence from the speech timing literature: Cummins and Port (1998) used a speech cycling method to study the tendency of speakers to produce particular timing patterns when repeating a short phrase over and over to a metronome. They found evidence for the *harmonic timing effect*, by which speakers align the vowels of stressed syllables within the repeated phrase so that they occur at 'harmonics' (i.e., simple fractions including 1/3, 1/2, and 2/3) of the 'fundamental' (time between repetitions of the phrase). This suggests a tendency to create beat-based timing patterns for speech produced under conditions of metronome-guided repetition.

4.1. *Conclusions*

When asked to repeat a simple tapped pattern at a comfortable rate, people show a propensity to choose an IPI that fits with the beat structure determined by the pattern. The IPI/ITI ratio was affected by rate, with more beats of silence for the IPI at faster rates. Participants also showed a preference for producing events in groups of two and four, when the silent beats are included in the count. These findings support the idea that people often elect to use a higher-order framework when organizing repeated motor behavior.

Acknowledgements

The first author was supported by the National Institutes of Health (NIH), National Institute on Deafness and Other Communication Disorders (NIDCD) training grant T32 DC000038.

References

Aschersleben, G. (2002). Temporal control of movements in sensorimotor synchronization. *Brain Cogn., 48,* 66–79.

Bååth, R., & Madison, G. (2012). *The subjective difficulty of tapping to a slow beat.* Paper presented at the 12th International Conference on Music Perception and Cognition, Thessaloniki, Greece.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B, 57,* 289–300.

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer [Computer program]. Version 5.3.83, retrieved 16 August 2014 from http://www.praat.org/.

Bolton, T. L. (1894). Rhythm. *Am. J. Psychol., 6*, 145–238.

Brochard, R., Abecasis, D. Potter, D., Ragot, R., & Drake, C. (2003). The 'ticktock' of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychol. Sci., 14*, 362–366.

Cummins, F. (2012). Looking for rhythm in speech. *Empir. Musicol. Rev., 7*, 28–35.

Cummins, F., & Port, R. F. (1998). Rhythmic constraints on stress timing in English. *J. Phon., 26*, 145–171.

Deutsch, D., Henthorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *J. Acoust. Soc. Am., 129*, 2245–2252.

Di Luca, M., & Rhodes, D. (2016). Optimal perceived timing: Integrating sensory information with dynamically updated expectations. *Sci. Rep., 6,* 28563.

Falk, S., Rathcke, T., & Dalla Bella, S. (2014). When speech sounds like music. *J. Exp. Psychol. Hum. Percept. Perform., 40*, 1491–1506.

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage, 47*, 1894–1903.

Keller, P. E., & Repp, B. H. (2005). Staying offbeat: Sensorimotor syncopation with structured and unstructured auditory sequences. *Psychol.l Res., 69,* 292–309.

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

Manning, F. C., & Schutz, M. (2016). Trained to keep a beat: Movement-related enhancements to timing perception in percussionists and non-percussionists. *Psychol. Res., 80*, 532–542.

Manning, F. C., Harris, J., & Schutz, M. (2017). Temporal prediction abilities are mediated by motor effector and rhythmic expertise. *Exp. Brain Res.*, *235*, 861–871.

Miyake, Y., Onishi, Y., & Pöppel, E. (2004). Two types of anticipation in synchronization tapping. *Acta Neurobiol. Exp.* (*Wars.*)*, 64,* 415–426.

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *J. Neurosci., 31*, 10234–10240.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *J. Exp. Psychol. Hum. Percept. Perform., 16*, 728–741.

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

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

Repp, B. H. (1995). Detectability of duration and intensity increments in melody tones: A partial connection between music perception and performance. *Percept. Psychophys., 57,* 1217–1232.

Repp, B. H. (2006). Rate limits of sensorimotor synchronization. *Adv. Cogn. Psychol., 2*, 163–181.

Repp, B. H. (2007). Perceiving the numerosity of rapidly occurring auditory events in metrical and nonmetrical contexts. *Percept. Psychophys., 69*, 529–543.

Repp, B. H. (2010). Do metrical accents create illusory phenomenal accents? *Atten. Percept. Psychophys., 72*, 1390–1403.

Repp, B. H., & Doggett, R. (2007). Tapping to a very slow beat: A comparison of musicians and nonmusicians. *Music Percept., 24*, 367–376.

Repp, B. H., & Su, Y. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonom. Bull.Rev., 20*, 403–452.

Snyder, J., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Percept., 18*, 455–489.