Beat Perception in Polyrhythms: Time is Structured in Binary Units

Cecilie Møller*, Jan Stupacher*, Alexandre Celma-Miralles*, and Peter Vuust

Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music Aarhus/Aalborg, Denmark

* These authors contributed equally to this work.

Author Note

Cecilie Møller https://orcid.org/0000-0002-6524-7608

Jan Stupacher https://orcid.org/0000-0002-2179-2508

Alexandre Celma-Miralles https://orcid.org/0000-0003-2557-6308

Peter Vuust https://orcid.org/0000-0002-4908-735X

This study was preregistered: https://aspredicted.org/yi5si.pdf. We have no known conflict of interest to disclose. Center for Music in the Brain is funded by the Danish National Research Foundation (DNRF117). Jan Stupacher is supported by an Erwin Schrödinger fellowship from the Austrian Science Fund (FWF J4288).

CRediT Authorship Contribution Statement

Cecilie Møller, Jan Stupacher and Alexandre Celma-Miralles: Conceptualization, methodology, data curation, formal analysis, writing – original draft. Peter Vuust: Conceptualization, methodology, writing – review & editing.

Corresponding author

Cecilie Møller, Universitetsbyen 3, Building 1710, 8000 Aarhus C, Denmark, cecilie@clin.au.dk

Word count: 7645 words (excl. title page, abstract, figures, tables, references)

Abstract

In everyday life, we group and subdivide time to understand the sensory environment surrounding us. Our perceptual system establishes hierarchical structures by nesting different groups of time intervals. Organizing time in units such as diurnal rhythms, phrases, and beat patterns, is fundamental to everyday behavior, speech, and music. When listening to music, we extract rhythmic regularities to create a hierarchical metrical structure that enables us to predict the timing of the next events. Foot tapping and head bobbing to musical rhythms are observable evidence of this process. In the special case of polyrhythms, at least two metrical structures compete to become the reference for these temporal regularities, rendering several possible beats with which we can synchronize our movements. While there is general agreement that tempo, pitch, and loudness influence beat perception in polyrhythms, we focused on the yet neglected influence of beat subdivisions. In three online experiments, 300 participants listened to a range of polyrhythms and tapped their index fingers in time with the perceived beat. The polyrhythms consisted of two simultaneously presented isochronous pulse trains with different ratios (2:3, 2:5, 3:4, 3:5, 4:5, 5:6) and different tempi. For ratios 2:3 and 3:4, we additionally manipulated the pitch of the pulse trains. Results showed a highly robust influence of subdivision grouping on beat perception manifested as a propensity towards beats that are subdivided into two or four equally spaced units, as opposed to beats with three or more complex groupings of subdivisions. Additionally, lower pitched pulse trains were more likely to be perceived as the beat. Our findings suggest that subdivisions, not beats, are the basic unit of beat perception, and that the principle underlying the binary grouping of subdivisions reflects a propensity towards simplicity. This preference for simple grouping is widely applicable to human perception and cognition of time.

Keywords: rhythm perception, beat and meter, sensorimotor synchronization, polyrhythm, music

Beat Perception in Polyrhythms: Time is Structured in Binary Units

In speech, music, and natural environments, we automatically group, subdivide and structure sound sequences evolving in time. The function of such hierarchical structures is to scaffold and anticipate upcoming auditory events, and to facilitate detection of unexpected events through a process that has been termed "predictive timing" (Friston, 2005; Fujioka et al., 2014). This perceptual grouping of temporal events is a cognitive mechanism essential for reducing complexity and making sense of the vibrant sensory environment surrounding us.

In search of a universal principle that can explain the ubiquity of rhythms in nature, physiology, attention, speech, poetry, and music, Bolton (1894) performed one of the earliest investigations of human rhythm processing. He showed that when listening to unaccented, equally spaced events, such as the isochronous ticks of a clock, listeners tend to subjectively accentuate every fourth or second tick and just rarely accentuate every third tick. Subsequent studies used various paradigms to explore this "tick-tock effect" (Fraisse, 1982; Temperley, 1963) and its neurophysiological correlates (Abecasis et al., 2005; Brochard et al., 2003). A general preference for binary or quaternary over ternary grouping is also evident in other tasks involving rhythm perception and production, such as in music (Bergeson & Trehub, 2006; Drake, 1993; Fujioka et al., 2010; Repp, 2003, 2007).

The spontaneous clapping, tapping, swaying, and nodding in time with music is a universal human behavior. It provides evidence of our ability to extract and perceive a regular pulse and its underlying hierarchically organized metrical structure. This capacity for beat perception is a fundamental human cognitive skill (Honing, 2012; Merchant et al., 2015) and present from infancy (Zentner & Eerola, 2010). Even when listening to complex musical rhythmic structures, which do not accent the beat itself, most people can extract a regular pulse and synchronize their movements to it, showing that beat perception is a constructive and endogenous process (Patel & Iversen, 2014).

The regular pulse we emphasize when synchronizing with music represents only one level in a more complex metrical structure. Figure 1 explains the concept of a metrical structure and illustrates its perceptual and behavioral consequences. Figure 1A illustrates how the subdivisions mark the points of the metrical grid, which is established on the basis of the smallest interval between perceptible events of the stimulus. The beat level is the level with which we usually synchronize our body movements. The cycle level marks the onset of the whole repeating pattern. Although the beat level is often the most perceptually salient level, we exhibit high flexibility with regards to synchronizing with any level in the metrical structure (Large et al., 2002). What is considered moving "in time with music" can relate to any level of the metrical structure. Which level of the metrical structure we synchronize our movements with may depend on a number of factors, including stimulus rate, dynamically changing rhythmic accents, and spontaneous motor tempo (Drake, Jones, et al., 2000; McKinney & Moelants, 2006).

In the special case of polyrhythms, two or more metrical structures co-exist and as such, polyrhythms are often used to create tension and increase expressiveness in musical performances. A polyrhythm is created by presenting at least two pulse trains containing coprime numbers of beats within the same periodic cycle, e.g., in ratios of 2:3, 3:4, or 3:5. A listener may perceive one or the other pulse train as representing the underlying beat and extract the corresponding metrical structure. The example depicted in Figure 1 is a 2:3 polyrhythm. What distinguishes the two manifestations of metrical structure is the beat level, which is defined by the grouping of elements at the subdivision level. On the left of Figure 1A, the ternary subdivision grouping results in two perceived beats per cycle. On the right, the binary subdivision grouping results in three perceived beats per cycle. The resulting metrical structures are organized differently and give rise to distinct and mutually exclusive perceptual experiences depending on how the elements at the subdivision level are grouped (Figure 1B). Two of these perceptual experiences are illustrated with speech examples in Figure 1C.

Previous polyrhythm studies have primarily focused on the constituent pulse trains (Beauvillain, 1983; Fidali et al., 2013; Jagacinski et al., 1988, 2016; Klapp et al., 1985, 1998; Pressing et al., 1996; Summers, 2002), neglecting the polyrhythms' metrical structures and the subdivisions underlying each of the pulse trains. Most of the research has aimed at assessing whether polyrhythms are perceived as integrated or segregated streams (Bregman, 1994), while some studies have made efforts to describe the factors that influence beat perception in polyrhythms. Tempo is consistently reported to strongly influence whether the faster or the slower pulse train represents the beat, which is also affected by the density, pitch, accentuation of elements and the relative timing between them (Handel, 1984; Handel & Lawson, 1983; Handel & Oshinsky, 1981; Moelants & van Noorden, 2005; Oshinsky & Handel, 1978). Importantly, these studies focused on characteristics of the constituent pulse trains, taking no notice of the characteristics of the two competing metrical structures that emerge when the pulse trains are superimposed on each other. In order to elucidate how we organize temporal auditory patterns, it is necessary to explicitly consider the hierarchical relationships between metrical levels (London, 2002).

Beat perception studies that do take metrical hierarchies into account tend to focus on the beat and meter levels, e.g., by assessing sensitivity to various manipulations of events at strong and weak beat positions (Bouwer et al., 2014; Drake, 1993; Fujioka et al, 2010). Yet, beat perception entails perception of subdivisions (London, 2002). Unfolding the empirically established temporal relation between beats and subdivisions, London (2002) pointed to the fact that the shortest inter-onset-interval (IOI) that can be perceived as representing beats is approximately 200 ms. In comparison, the shortest IOIs necessary for subjective rhythmization, such as the "tick-tock" effect (Bolton, 1894) and interval discrimination (Hirsh et al., 1990) is approximately 100 ms, i.e., corresponding to subdividing the beat by a factor of two. As such, we only perceive a regular beat if the cognitive constraints on temporal perception allow us to perceive the subdivisions of that beat, at least potentially. In a tapping study investigating the benefits and costs of explicitly subdividing the beat, similar thresholds for synchronization rates were found (Repp, 2003). Because participants were required

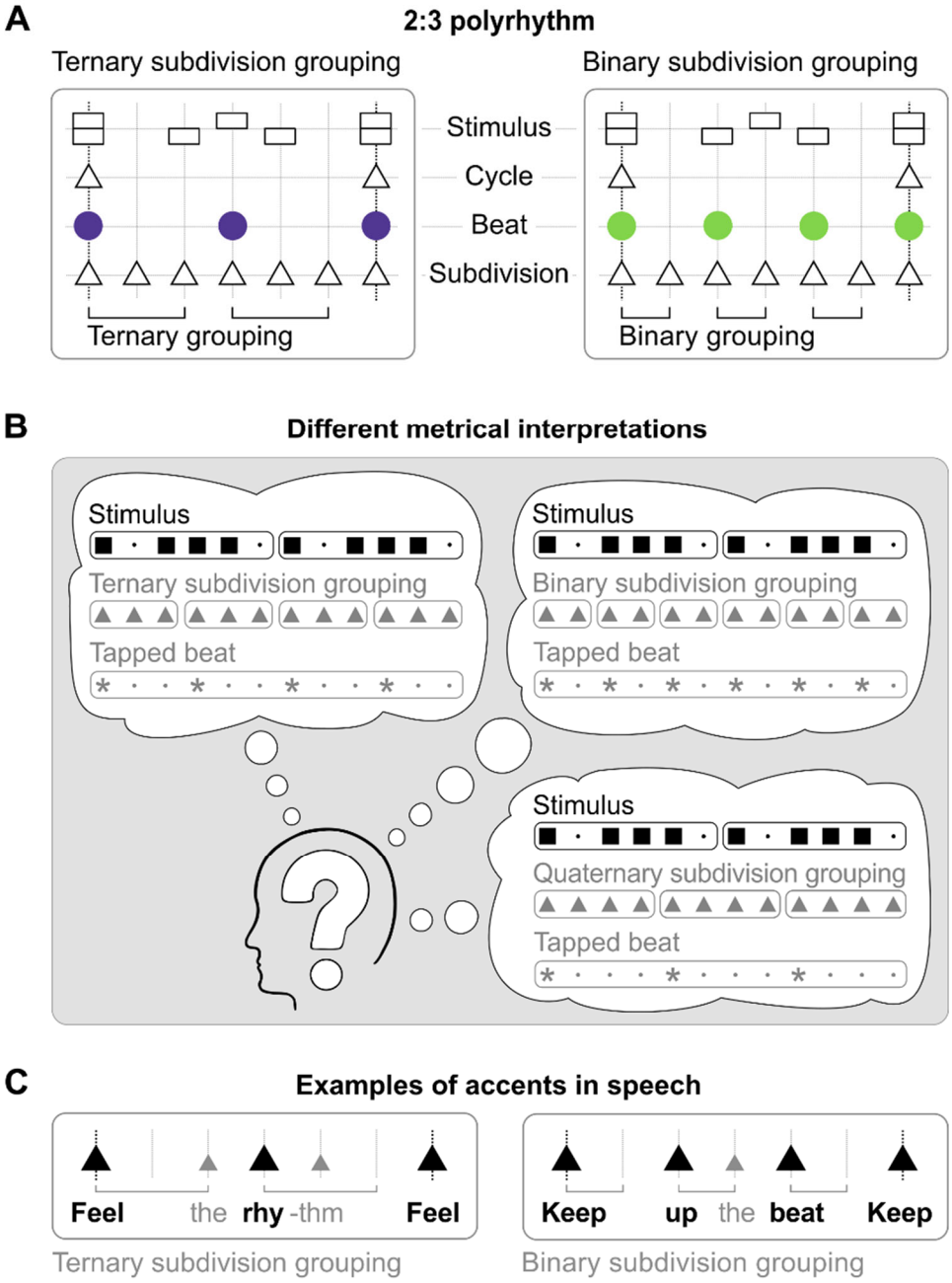
to tap only to every 2nd, 3rd, or 4th element of pulse trains presented at different rates, Repp's study also ruled out the possibility that motor constraints influenced participants' ability to make judgements of the quantity of the faster subdivisions. Assuming that successful grouping of subdivisions is necessary for beat perception to occur (London, 2002), we have to move our research focus from the beat level to the subdivision level of the metrical structure—especially when assessing beat perception and sensorimotor synchronization in polyrhythms.

To provide a comprehensive account of beat perception that takes into account the most basic level of the metrical hierarchy, the purpose of this online finger tapping study was to assess how subdivision grouping biases listeners to adopt one rather than another possible metrical structure inherent in a given polyrhythm consisting of two pulse trains. Owing to their ambiguous nature, polyrhythms are ideal stimuli for assessing rhythmic interpretations in tapping studies. Such studies rest on the assumption that subjects synchronize their taps with the perceived beat (Beauvillain, 1983; McKinney & Moelants, 2006). The paradigm in the present study allowed for categorization of tapping responses into all possible metrical levels, including half and double tempo in relation to the constituent pulse trains. Overall, we hypothesized that participants would prefer to tap to a beat with binary rather than ternary or irregular subdivision grouping, and ternary rather than irregular subdivision grouping. This hypothesis was based on the general propensity for binary grouping of isochronous auditory stimuli and their subdivisions (e.g., Bolton, 1894; Brochard et al., 2003; Repp, 2008).

Participants were recruited worldwide via social media. In three separate online experiments we manipulated tempo (N = 100), ratio (N = 120), and pitch (N = 80) of pulse trains in various polyrhythms. Tempo was manipulated to assess transition points of tapping preference within and between metrical structures. Ratio was manipulated to investigate different types of subdivision grouping (binary, ternary, irregular) in the slow and the fast pulse trains at different tempi. Pitch manipulations allowed assessing the effect of low-frequency notes on beat perception relative to the effect of subdivision grouping. We report the results of preregistered main analyses (https://aspredicted.org/yi5si.pdf).

Figure 1

Examples of Metrical Structures in Rhythms



Note. A) Left and right panels show the same 2:3 polyrhythm with two different underlying metrical structures, corresponding to the two-beat pulse train with ternary grouped subdivisions (left) and the three-beat pulse train with binary grouped subdivisions (right). B) Three different examples of interpretations of the 2:3 polyrhythm that lead to three different behavioral outcomes when synchronizing body movements—here finger tapping—to the stimulus. The subjective experience of the rhythm's 'feeling' depends on the perceived beat, which in turn depends on the grouping of subdivisions. C) Stressing the bold syllables of the speech examples induces ternary grouped two-beat (left) or binary grouped three-beat (right) interpretations of the 2:3 polyrhythms.

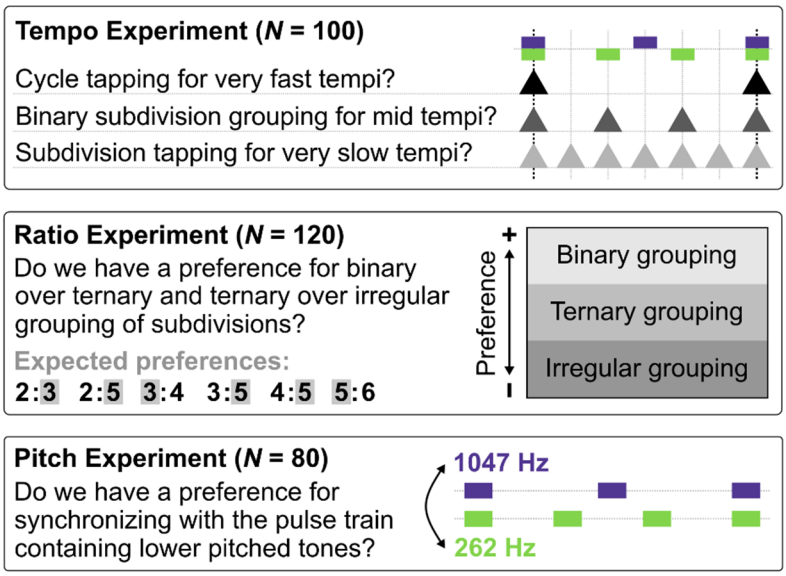
Methods

Participants

The study included data of 300 participants (159 female, age range 18-75 years, *M* = 31 years, *IQR* = 10.5 years). Additional incomplete or duplicate responses were excluded. The majority of respondents grew up in Denmark (32.7%) followed by Spain (12.3%), UK (7.3%), Germany (5.3%), and the US (4.7%). Forty-four youth countries accounted for the remaining 62.3 % of responses. Musicianship was assessed with one item from Ollen's Musical Sophistication Index (Ollen, 2006). 11% considered themselves nonmusicians, 29%, music-loving nonmusicians, 24% amateur musicians, 18% serious amateur musicians, 11% semi-professional musicians, and 8% professional musicians. Participants were randomly assigned to complete either the Tempo (*N* = 100), Ratio (*N* = 120) or Pitch (*N* = 80) Experiment (Figure 2). Sample sizes were determined a priori based on pilot studies and preregistered. Participants were informed that their data would be used for scientific purposes. They were not offered any kind of payment for their participation. The study was conducted in accordance with the guidelines from the Declaration of Helsinki and the Danish Code of Conduct for Research Integrity and Aarhus University's Policy for research integrity, freedom of research and responsible conduct of research. In Denmark, research that does not collect nor store personally identifiable or sensitive information are exempt from IRB approval, which we confirmed in correspondence with the local IRB. All collected data included no personally identifiable information.

Figure 2

Experimental Overview and Main Research Questions for the Three Individual Experiments



Procedure

Participants were recruited through social media and directed to a webpage containing a Shiny app, which was developed using the JavaScript library jspsych (de Leeuw, 2015) embedded into psychTestR (Harrison, 2020). Headphones and touch screens were recommended, though the experiment could also be run on computers using internal speakers. After initial assessment and testing of their device, participants were randomly assigned to the three different experiments, Tempo, Ratio, or Pitch. To ensure a proper representation of low-frequency tones, for the Pitch dataset, we only included participants wearing headphones. After a spontaneous motor tempo assessment, which familiarized participants with tapping on their device, the experimental tapping task was explained. Participants' task was to listen to the polyrhythm and to start tapping with the index finger of their dominant hand when they could clearly feel the beat. They continued tapping until the sound stopped. Stimuli were presented once and in random order. After each trial, a 100-point slider allowed participants to rate how difficult it was to find the beat. Finally, participants filled out a short questionnaire assessing their musical and linguistic background. Questionnaire data, spontaneous motor tempo, and difficulty ratings were not analyzed in the present work, which focuses specifically on reporting the tapping data as specified in the preregistration (https://aspredicted.org/yi5si.pdf).

Finger Tapping Analyses

We obtained a high proportion of tapping responses. Only 1.3% of the tapping trials were missing in Tempo, 2.0% in Ratio, and 1.0% in Pitch. To remove involuntary double taps and device artefacts, we calculated all the time intervals between consecutive taps (Inter Tapping Intervals, ITIs) and removed the second tap of each ITI shorter than 150 ms (see figure S1A in Supplementary Material). We additionally removed the first 2 taps of each trial. If less than five taps remained, we removed the trial from the analysis. The resulting percentages of missing or excluded trials were 4.9% in Tempo, 6.6% in Ratio, and 4.0% in Pitch.

The timing of the taps was converted into angular measures, i.e., phase in radians. This means that the taps are measured as a circular angle in relation to the timing of a periodic reference event at angle 0. The periodic references were defined as the different metrical levels in each polyrhythm: *cycle*, *slow pulse train*, *slow pulse train - double tempo*, *slow pulse train - half tempo*, *fast pulse train*, *fast pulse train - double tempo*, *fast pulse train - half tempo*, and the common *subdivision* level (see Figure S1B). This method allowed us to obtain the consistency of the taps at each metrical level, even if the device or the participant missed some taps.

Using circular statistics (Berens, 2009) we computed the mean resultant vector of the tapping responses to each stimulus for every metrical level. The tapping consistency is reflected by the length of the mean vector ranging from 0 to 1. To determine the metrical level tapped by the participant in each trial, we took the longest vector length among all the metrical levels and checked whether the taps were uniformly distributed using the Rao's Spacing test and the Rayleigh test. Only when both tests were significant ($p \leq .05$), the tapping responses were assigned to a metrical category. Finally, we confirmed the selection of the categorized metrical level by

only accepting those metrical tapping responses (i) whose mean of the ITIs fell in the range of ±15 % of the inter-onset interval (IOI), i.e., the tempo in milliseconds, of the categorized meter and (ii) whose standard deviation of the ITIs was smaller than 66% of the IOI. These means and standard deviations were obtained after removing ITIs that fell beyond two standard deviations from the mean ITI of each trial. The combination of linear and circular analyses resulted in inclusion of the following percentages of trials with consistent tapping at one of the metrical levels: 66.2% in Tempo, 54.1% in Ratio, and 76.2% in Pitch.

Tempo Experiment

The perception of a regular beat in isochronous sequences of sounds is possible if the tempo is within a range of approximately 30-300 BPM / 2000-200 ms (Drake, Penel, et al., 2000; Duke, 1989; London, 2002). Particularly salient pulses are perceived at tempi between 80-160 BPM / 750-375 ms, corresponding to our preferred spontaneous motor tempo (Fraisse, 1982; Michaelis et al., 2014; Parncutt, 1994). Previous studies suggested that, when synchronizing with polyrhythms, individuals tap in time with the pulse train closest to the human preferred tempo, i.e., the faster pulse train in slower tempi and vice versa (Moelants & van Noorden, 2005; Oshinsky & Handel, 1978). In contrast, we expected that individuals synchronize with the pulse train that can be subdivided into binary groups. We assessed 2:3 and 3:4 polyrhythms in a wide range of tempi. The former contains binary subdivision grouping in the fast pulse train (the 3 in a 2:3 polyrhythm) whereas the latter contains binary subdivision groupings in the slow pulse train (the 3 in a 3:4 polyrhythm).

Hypotheses

We expected that any difference in distribution of tapping preference between the two polyrhythms can be explained by differences in subdivision grouping and not by the relative timings of the slow and the fast pulse trains. At moderate tempi, we expected taps to occur in time with the pulse train in which subdivisions could be grouped in pairs (binary). At faster tempi, we expected the pulse train itself to be perceived as subdivisions - again with preferences for binary grouping of subdivisions. At extremely slow and extremely fast tempi, we expected taps to shift towards the subdivisions and the cycle, respectively.

Stimuli

All stimuli in this study were created with Ableton Live 8 (Ableton, Berlin, Germany; audio files in https://researchbox.org/278). The Tempo stimuli consisted of 2:3 and 3:4 polyrhythms ranging from very slow (approx. 40 BPM) to very fast tempi (approx. 450 BPM; see Table 1). The two pulse trains in each of the polyrhythms were presented with the same cowbell sound and the same amplitude. The duration of the 15 stimuli was between 18 and 28 s, depending on the ratio and tempo. Every stimulus was presented once in random order and consisted of at least six repetitions of a whole polyrhythm cycle. Additional stimuli for control analyses are shown in Figures S2 and S3 in Supplementary Material.

Statistical Analyses

Tapping responses were categorized as one of the following metrical categories: *cycle*, *slow pulse train*, double and the half tempo of the *slow pulse train*, *fast pulse train*, double and half tempo of the *fast pulse train*, and the common *subdivisions*. Unclear tapping performances that could not be categorized were not analyzed. Within each metrical level, Cochran's Q tests were used to investigate the effect of tempo. McNemar's tests were used to analyze differences between neighboring tempo pairs. All analyses were Bonferroni-corrected for multiple comparisons.

Table 1

The 15 Stimuli of the Tempo Experiment

2:3 Polyrhythm	
Tempo in ms	**Tempo in BPM**
1500:1000	40:60
1000:667	60:90
667:444	90:135
444:296	135:202
296:198	202:304
198:132	304:456
3:4 Polyrhythm	
Tempo in ms	**Tempo in BPM**
1579:1184	38:51
1186:889	51:67
889:667	67:90
667:500	90:120
500:375	120:160
375:281	160:213
281:211	213:284
211:158	284:379
158:119	379:506

Ratio Experiment

Although the literature acknowledges that rhythmic interpretation depends on the configuration and in turn on the structure of the polyrhythm (Handel, 1984; Handel & Lawson, 1983; Pitt & Monahan, 1987), no efforts have yet been made to directly assess which particular aspects of the polyrhythm configurations drive tapping preference. In the Ratio Experiment, we assessed preference for metrical structure by focusing on the polyrhythm subdivision level, rather than preference for the constituent pulse trains, as in previous studies (Handel, 1984; Handel & Lawson, 1983; Handel & Oshinsky, 1981; Oshinsky & Handel, 1978). The paradigm included polyrhythms ranging from simple (i.e., 2:3) to complex (i.e., 5:6; see Table S1 in Supplementary Material for a definition of complexity). Different configurations of polyrhythms give rise to different possible subdivision groupings (binary, ternary, irregular). For example, in a 2:3 polyrhythm, a binary subdivision grouping subserves the three-beat and a ternary subdivision grouping subserves the two-beat (Figure 1A). Accordingly, with respect to subdivision grouping, the 2:3 polyrhythm can be denoted ternary:binary, while for instance the 2:5 polyrhythm can be denoted irregular:binary.

Hypothesis

We expected that the metrical structure containing simpler subdivision grouping would be preferred over those containing more complex subdivision grouping. This means that we expected the following preferences: binary grouping (2 or 4) preferred over ternary grouping (3), and ternary grouping (3) preferred over irregular grouping (5).

Stimuli

The Ratio stimuli consisted of 2:3, 2:5, 3:4, 3:5, 4:5 and 5:6 polyrhythms (audio files in https://researchbox.org/278). The two pulse trains in each polyrhythm were presented with the same cowbell sound and the same amplitude. The tempo of the polyrhythms were based on the duration of their subdivisions, i.e., their least common denominator (Table 2). Two 3:4 polyrhythms with the subdivision tempi 167 and 125 ms, corresponding to pulse train tempi of and 90:120 and 120:160 BPM (667:500 and 500:375 ms) were used as anchors. To make the tempo of the pulse trains comparable across the different ratios, 2:3 and 2:5 polyrhythms were additionally slowed down to half tempo, whereas 3:5, 4:5, and 5:6 polyrhythms were additionally speeded up to double tempo (Table 2). Following the temporal constraints on beat perception described by London (2002) and Repp (2003), it is reasonable to assume that beat perception is only possible when the tempo allows for grouping of subdivisions. Consequently, in a metrical structure containing groupings of a large number of subdivisions, the tempo of the pulse train must be slowed down to allow beat perception to occur, and to make balanced comparisons between different ratios possible. In other words, comparisons should be made between subdivision tempi, not pulse train tempi. The duration of the 22 stimuli was between 15 and 24 s, depending on the ratio and tempo. Every stimulus consisted of at least five repetitions of a whole polyrhythm cycle.

Table 2

The 22 Stimuli of the Ratio Experiment

Ratio	Duration of subdivisions in ms	Duration of cycle in ms	Tempo of pulse trains in ms	Tempo of pulse trains in BPM
2:3	333, 250, 167, 125	2000, 1500, 1000, 750	1000:667, 750:500, 500:333, 375:250	60:90, 80:120, 120:180, 160:240
2:5	333, 250, 167, 125	3330, 2500, 1670, 1250	1667:667, 1250:500, 833:333, 625:250	36:90, 48:120, 72:180, 96:240
3:4	167, 125	2000, 1500	667:500, 500:375	90:120,120:160
3:5	167, 125, 84, 63	2500, 1875, 1260, 945	833:500, 625:375, 420:252, 316:189	72:120, 96:160, 143:238, 190:317
4:5	167, 125, 84, 63	3340, 2500, 1680, 1260	833:667, 625:500, 420:335, 316:252	72:90, 96:120, 143:179, 190:238
5:6	167, 125, 84, 63	5010, 3750, 2520, 1890	1000:833, 750:625, 504:420, 377:316	60:72, 80:96, 119:143,159:190

Statistical Analyses

Tapping responses were coded as 1 when falling into one of the *fast pulse train* categories (*fast pulse train*, *fast pulse train - double tempo*, or *fast pulse train - half tempo*) and coded as 0 when falling into one of the *slow pulse train* categories (*slow pulse train*, *slow pulse train - double tempo*, or *slow pulse train - half tempo*). These values were averaged across all tempi in each of the polyrhythm ratios (Figure 5A). For statistical analysis, we computed two means: 1) the mean of all polyrhythm ratios in which the slow:fast pulse train relation allowed for ternary:binary subdivision grouping (2:3), irregular:binary subdivision grouping (2:5, 4:5) or irregular:ternary subdivision grouping (3:5), i.e., ratios with simpler subdivision grouping in the faster pulse train, and 2) the mean of all polyrhythm ratios in which the slow:fast pulse train relation allowed for binary:ternary subdivision grouping (3:4) or binary/ternary:irregular subdivision grouping (5:6), i.e., ratios with simpler subdivision grouping in the slower pulse train. These two means were compared using a paired sample Wilcoxon test.

Pitch Experiment

The pitch of elements in a musical rhythm is an important factor for beat perception. Low-pitched rhythmic elements increase the sensitivity to timing variation on behavioral and neural levels (Hove et al., 2014), and EEG activity at meter-related frequencies increases with low-pitch sounds (Lenc et al., 2018). In general, high energy in bass frequencies are important for inducing movements, such as tapping in time with the beat and dancing (Hove et al., 2020; Stupacher at al., 2016; van Dyck et al., 2013; Varlet et al., 2020). When investigating the effect of pitch on beat perception in polyrhythms, Handel and Oshinsky (1981) found that

participants tended to perceive the lower pitched pulse train as the beat, and that this preference counteracted preferences related to the timing of the pulse trains. Here, we investigated this effect of lower pitch in more detail by not only varying pitch, but also loudness between the slow and fast pulse trains in 2:3 and 3:4 polyrhythms. This manipulation was informed by pilot studies showing that without a loudness manipulation the vast majority of participants tap in time with the pulse train with binary subdivision grouping. The resulting design allowed us to assess whether preferences for binary subdivision grouping have stronger effects on beat perception than loudness or bass frequencies.

Hypotheses

We expected participants' tapping responses to reflect a preference for lower pitched pulse trains. This means that the preference for binary subdivision grouping should be strengthened when coinciding with the low-pitched pulse train and weakened when coinciding with the high-pitched pulse train.

Stimuli

The Pitch stimuli consisted of 2:3 and 3:4 polyrhythms created with marimba sounds at the tempi of 90:135 BPM (667:444 ms) and 90:120 BPM (667:500 ms), respectively (audio files in https://researchbox.org/278). Table 3 details the *pitch* and *loudness* manipulation. In each polyrhythm, one of the pulse trains was pitched low with a peak frequency of 262 Hz (C4) and the other pulse train was pitched higher with a peak frequency of 1047 Hz (C6). This manipulation was counterbalanced. The loudness of the two pulse trains was either the same, moderately louder for the pulse train with ternary subdivisions, or markedly louder for the pulse train with ternary subdivisions. Loudness was measured in Loudness K-weighted Full Scale (LKFS) with the Orban Loudness Meter (version 2.9.6; www.orban.com/meter/). The duration of the 12 stimuli was 17 and 18 s for ratios 2:3 and 3:4, respectively. To exclusively investigate the effect of amplitude, we presented six additional control stimuli with the same loudness manipulation as the experimental stimuli but using the same pitch in both pulse trains (C5 with a peak frequency of 524 Hz; Figure S4 in the Supplementary Material).

Statistical Analyses

The dependent variable was defined as the tapping consistency related to the *slow pulse train* minus the tapping consistency related to the *fast pulse train* in each trial. As the tapping consistency is defined as the vector length in circular statistics, this procedure resulted in a value ranging between -1 and 1. Values close to -1 indicate that participants consistently tapped in time with the *slow pulse train*, whereas values close to 1 indicate that participants consistently tapped in time with the *fast pulse train*. Data for the individual factor combinations (2 *pitch* x 3 *loudness* factor levels) were not normally distributed (Shapiro-Wilk p-values < .001). In the preregistration for this study, we planned to use linear mixed effects models for analyzing the pitch data. However, the residuals of these models were not normally distributed, as indicated by visual inspections of Q-

Q plots and Shapiro-Wilk tests (all p-values < .001). Consequently, we used two Wilcoxon tests for paired samples to investigate the effect of *pitch* (low pitch in *slow* vs. *fast pulse train*) and two Kruskal-Wallis tests to investigate the effect of *loudness* in 2:3 and 3:4 polyrhythms, separately. We report the results of the nonparametric tests, which were anyway supported by the linear mixed effects models.

Table 3

The 12 Pitch Stimuli

2:3 Polyrhythm (90:135 BPM)			
2 pitch	**3 pitch**	**2 amplitude**	**3 amplitude**
low	high	=	=
low	high	+ 3 LKFS	- 3 LKFS
low	high	+ 6 LKFS	- 6 LKFS
high	low	=	=
high	low	+ 3 LKFS	- 3 LKFS
high	low	+ 6 LKFS	- 6 LKFS
3:4 Polyrhythm (90:120 BPM)			
3 pitch	**4 pitch**	**3 amplitude**	**4 amplitude**
low	high	=	=
low	high	- 3 LKFS	+ 3 LKFS
low	high	- 6 LKFS	+ 6 LKFS
high	low	=	=
high	low	- 3 LKFS	+ 3 LKFS
high	low	- 6 LKFS	+ 6 LKFS

Note. Low pitch refers to C4 (peak frequency 262 Hz) and high pitch to C6 (peak frequency 1047 Hz). In the amplitude columns, "=" refers to equal loudness measured in LKFS; +3/+6 LKFS mark increases of the original loudness by 3 and 6 LKFS, respectively; -3/-6 LKFS mark decreases of the original loudness by 3 and 6 LKFS, respectively.

Results and Discussion

In the following section, we report and discuss the results of the three experiments Tempo, Ratio, and Pitch separately, to unravel the individual effects of modulating these three features on beat perception in

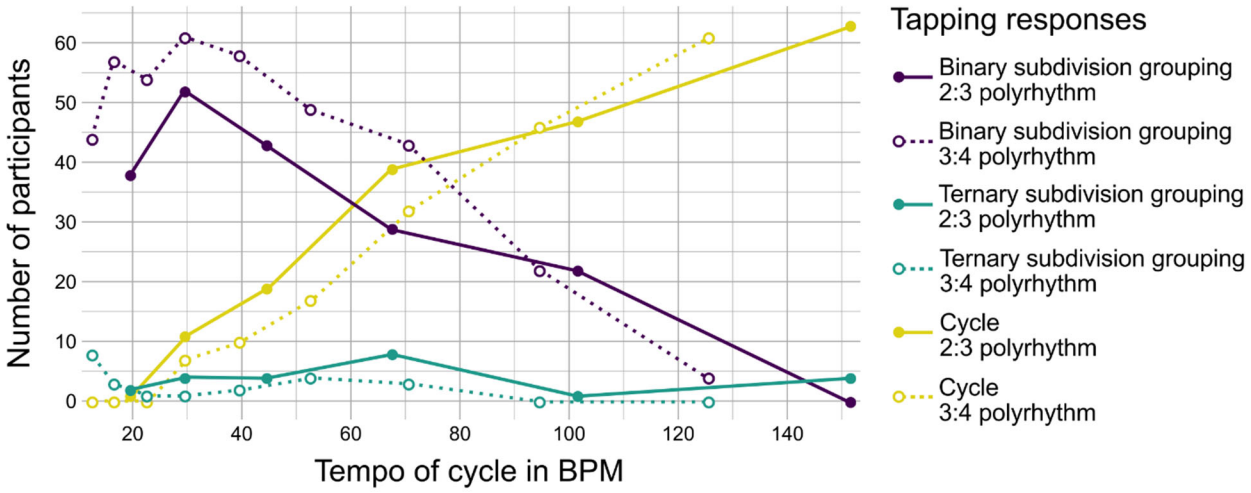
polyrhythms. In the subsequent General Discussion, we focus on the converging evidence for a preference for binary grouping of subdivisions in musical rhythms, and broaden the perspective to binary grouping of temporal units in the perception of time in everyday life.

Tempo Experiment

Consistent with our hypothesis, most participants' responses fell into one of the response categories with binary subdivision groupings (Figure 3). The pulse trains with binary subdivision grouping, i.e., *fast pulse train* in 2:3 and *slow pulse train* in 3:4 polyrhythms, showed statistically significant differences in the proportion of responses to the different tempi ($X^2(5) = 128.50$, $p < .001$, $\eta^2_Q = 0.22$, and $X^2(8) = 251.32$, $p < .001$, $\eta^2_Q = 0.28$, respectively; Figure 4). For the pulse trains with ternary subdivision grouping, i.e., *slow pulse train* in 2:3 and *fast pulse train* in 3:4 polyrhythm, no statistically significant differences in the proportion of responses to the different tempi of the stimuli were found (both p-values > .5). In any one of the six tempi in 2:3 and 9 tempi in 3:4 polyrhythms, less than 6% of participants tapped in time with the pulse train with ternary subdivision grouping.

Figure 3

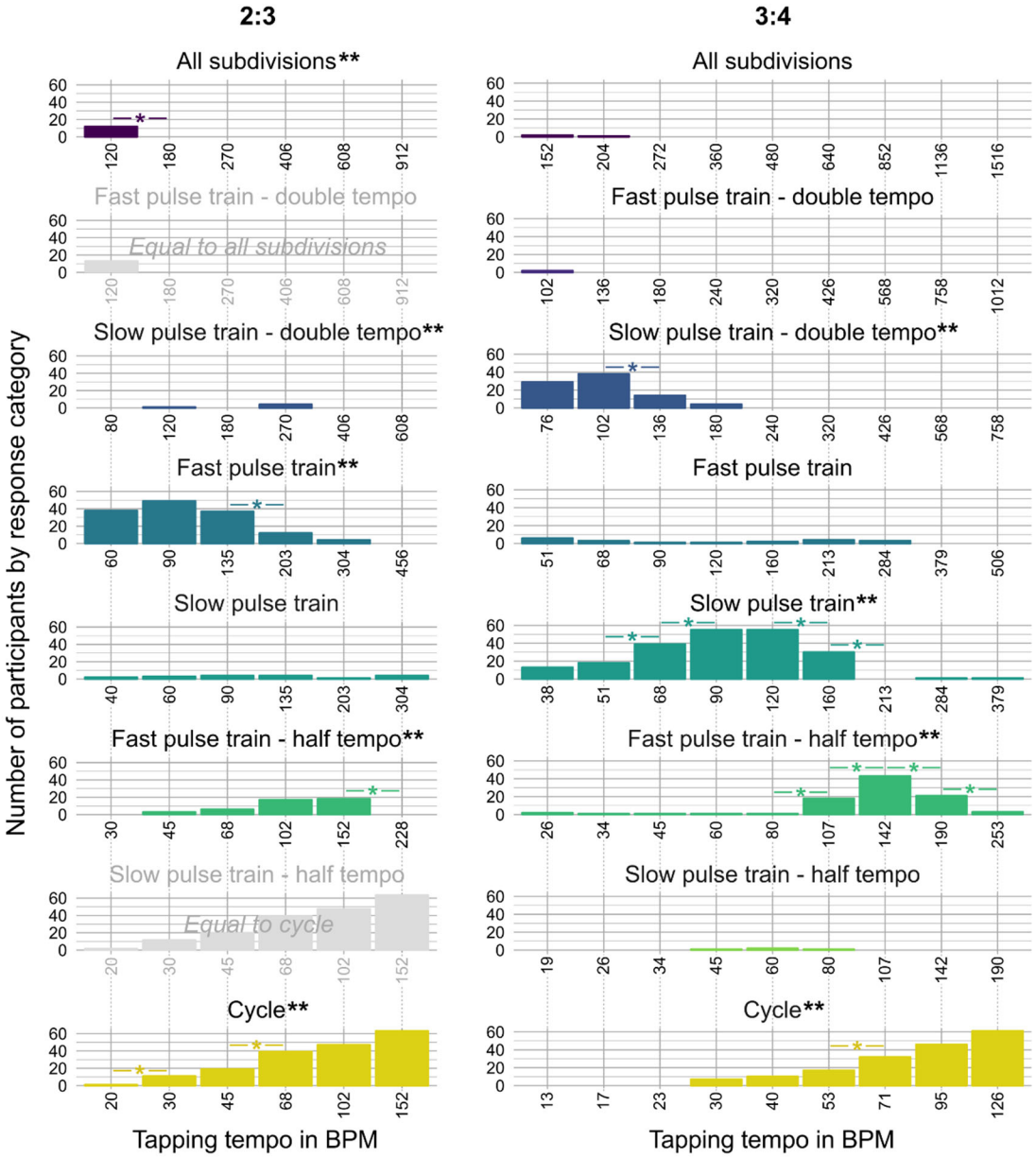
Participants' Preference for Subdivision Grouping in the Experiment Manipulating the Tempo of 2:3 and 3:4 Polyrhythms



Note. At slow and moderate tempi, most participants tapped in time with one of the response categories with binary subdivision groupings regardless of tempo. At faster tempi, participants switched to tapping in time with the cycle. Only a few participants' tapping behavior corresponds to ternary subdivision grouping. Tapping responses related to groupings of binary and ternary subdivisions are concatenated separately: binary grouping in 2:3 (*fast pulse train* including *half tempo*), binary grouping in 3:4 (*slow pulse train,* including *half tempo* and *double tempo,* and *fast pulse train - half tempo*), ternary grouping in 2:3 (*slow pulse train,* including *double tempo*), and ternary grouping in 3:4 (*fast pulse train*, including *double tempo*). Subdivision grouping is indiscernible at the *all subdivisions* (not shown) and *cycle* levels.

Figure 4

Tapping Behavior as a Function of Tempo for the Individual Metrical Categories of 2:3 and 3:4 Polyrhythms



Note. Bars represent the number of participants tapping in time with each of the eight different metrical categories ranging from the fastest category, *all subdivisions* (top) to the slowest category, *cycle* (bottom). Responses to the 2:3 and the 3:4 polyrhythms are shown at different tempi ranging from slow (left) to fast (right). Individual x-axes show the tapping tempo of responses included in that metrical category. For instance, a participant listening to the slowest 2:3 polyrhythm may choose to tap in time with either *all subdivisions* at 120 BPM, the *slow pulse train - double tempo* at 80 BPM, the *fast pulse train* at 60 BPM, the *slow pulse train* at 40 BPM, the *fast pulse train - half tempo* at 30 BPM, or the *cycle* level at 20 BPM. Double asterisks (**) mark significant Cochran's Q tests comparing responses over different tempi within each metrical category. Asterisks (*) mark significant McNemar's tests for pairwise comparisons between neighboring tempi within each metrical category. All tests are Bonferroni-corrected for multiple comparisons.

In 2:3 polyrhythms, most participants chose the *fast pulse train* with binary subdivision grouping at slower and moderate tempi. At faster tempi, most participants tapped in time with the *cycle* instead of switching to the *slow pulse train*. Notably, the second most popular response at faster tempi was to group two cycles to form a 4:6 polyrhythm and to tap every other beat of the *fast pulse train*, effectively treating the *fast pulse train* itself as the subdivision level (Duke, 1989). We interpret this response, denoted *fast pulse train - half tempo* in Figure 4, as the participants' successful attempt to ensure the maintenance of a metrical structure containing binary subdivision grouping.

In 3:4 polyrhythms, most participants tapped the *slow pulse train - double tempo* at slower tempi and switched to *slow pulse train* at intermediate tempi from 68 BPM. Both of these responses indicate that participants perceived beats with binary subdivision grouping. At tempi faster than 160 BPM, most participants tapped the *fast pulse train - half tempo*. As in the 2:3 polyrhythm, tapping every other event in the *fast pulse train* of the 3:4 polyrhythm ensured the maintenance of a metrical structure containing binary subdivision grouping.

As expected, in both polyrhythms, *cycle* tapping increased with tempo and was the preferred response at the fastest tempi, consistent with previous reports of a preference for higher metrical levels at faster tempi (Duke, 1989; Parncutt 1994). We also expected tapping in time with *all subdivisions* in the slowest end of the tempo range. This hypothesis was only confirmed for the 2:3 polyrhythm ($X^2(5) = 60.00$, $p < .001$, $\eta^2_Q = 0.10$). Even the slowest *all subdivisions* tempi in the experiment were quite fast (120 and 152 BPM in the 2:3 and 3:4 polyrhythms), which likely explains why only few participants preferred to synchronize with the subdivision level, especially in the 3:4 polyrhythm.

It is particularly interesting to inspect responses to the polyrhythms with subdivision tempi of approximately 100 ms / 600 BPM (see tempo of *all subdivisions* in the top row of Figure 4), as this tempo marks the lower limit of subjective grouping of sounds (London, 2002; Repp, 2003). It is reasonable to assume that a shift in response pattern around this location reflects a change in the perceived metrical structure that occurs simply as a consequence of the fact that grouping of subdivisions is no longer possible. In other words, the pulse train itself becomes the subdivision level when subdivisions are too fast to be subjectively grouped. Participants discarded the most common tapping preference, i.e., *fast pulse train* in 2:3 and *slow pulse train* in 3:4 around this point. Furthermore, a significant shift was observed from *fast pulse train - half tempo* to *cycle* at the fastest tempo of the 3:4, where even the pulse train (at 506 BPM) approached the cognitive limit of subjective grouping of short sounds.

The distribution of tapping preferences in the present study contrasts with results of Handel and Oshinsky (1981), who argued for a general preference for faster pulse trains at slower presentation rates and vice versa. They reported that the 4-pulse train of a 3:4 polyrhythm was preferred in cycle durations from 1.4 seconds to 2.4 seconds (with 2.0 seconds corresponding to 90:120 BPM). In our study, this tempo range represented the peak of 3-pulse train responses, i.e. *slow pulse train*. Our results also differ from Handel and Oshinsky's findings for 2:3 polyrhythms. Handel and Oshinsky show an increase in the 2-pulse train and a

decrease in the 3-pulse train at faster tempi. We see a similar decrease in 3-pulse tapping; however, this decrease was only in favor of *cycle* and *fast pulse train - half tempo* tapping, and not in favor of the 2-pulse train tapping.

One advantage of using polyrhythms to investigate beat perception is that one can assess the effect of metrical structure and tempo at the same time. Our findings show that participants' responses followed the common distributions of preferred motor tempo around 80-160 BPM (Moelants, 2002). However, it is important to note that if preferred motor tempo was the only factor driving the results, we would see similar distributions of tapping responses in the slow and fast pulse trains, with the *fast pulse train* reaching its peak slightly earlier, i.e., at slightly slower stimulus tempi, than the *slow pulse train*. This is not the case. Rather, the tapping responses indicate a clear preference for tapping to a beat with binary rather than ternary subdivision grouping. This preference for simpler subdivision grouping is further explored in the Ratio Experiment.

Ratio Experiment

Consistent with our hypothesis, results reflected tapping preferences for pulse trains with simpler subdivision groupings, regardless of whether this was the faster or the slower pulse train in the polyrhythm. Binary subdivision grouping was preferred over ternary grouping and ternary grouping was preferred over irregular grouping. A Wilcoxon test for paired samples indicated a significant difference between tapping to polyrhythms with simpler subdivision grouping in the *slow pulse train* (3:4, 5:6; Mdn = 0.0, IQR = 0.25) and polyrhythms with simpler subdivision grouping in the *fast pulse train* (2:3, 2:5, 3:5, 4:5; Mdn = 0.83, IQR = 0.24; $Z = -7.96$, $p < .001$, $r = .73$; Figure 5B). As shown in Figure 5A, the propensity towards simpler subdivisions is particularly strong in the four simpler polyrhythms, i.e., 2:3, 2:5, 3:4, and 3:5, whereas the 4:5 and 5:6 polyrhythms appear to be more perceptually ambiguous.

A more nuanced picture emerges when the metrical categories of the tapping responses for the individual tempi in each of the polyrhythm are taken into account (Figure 5C). The results of the 3:4 (binary:ternary) polyrhythm show a clear preference for the *slow pulse train* with binary subdivision grouping. As expected, this preference was flipped in favor of the *fast pulse train* of the 2:3 (ternary:binary), 2:5 (irregular:binary), and 3:5 (irregular:ternary) polyrhythms, mirroring the shift to simpler subdivision grouping in the fast pulse trains. In these polyrhythms, only seven or less participants synchronized with the slower pulse trains with ternary or irregular subdivision groupings. Notably, increasing the tempo within each of these polyrhythms was not associated with a clear increase in *slow pulse train* responses, as a pure effect of tempo would predict. Rather, it resulted in an increase in tapping the *fast pulse train - half tempo*, ensuring the maintenance of a metrical structure containing simpler grouped subdivisions.

In the 3:5 polyrhythm, only 21% of responses were categorized as being related to one of the two pulse trains or its half- and double-tempo equivalents, making this the polyrhythm with the least consistent pulse train tapping. This may be due to the fact that, compared to the other ratios, neither of the pulse trains in the 3:5 polyrhythm contains binary subdivision grouping—the grouped subdivisions are irregular:ternary. At the

slowest tempo, all but two of the 41 participants who engaged in consistent tapping synchronized with the *fast pulse train* with simpler subdivision grouping. With increases in tempo, the number of consistent responses decreased and the proportion of *fast half tempo* responses increased. *Fast half tempo* tapping indicates that participants skipped every other beat of the 5-pulse train to construct an additional metrical structure. This task is more complex compared to *fast half tempo* tapping in a 2:5 polyrhythm, which maintains a metrical structure with binary subdivision grouping. The new metrical structure in 3:5 is likely preferred over the two metrical structures already inherent in the 3:5 polyrhythm cycle, because it contains binary subdivision groupings.

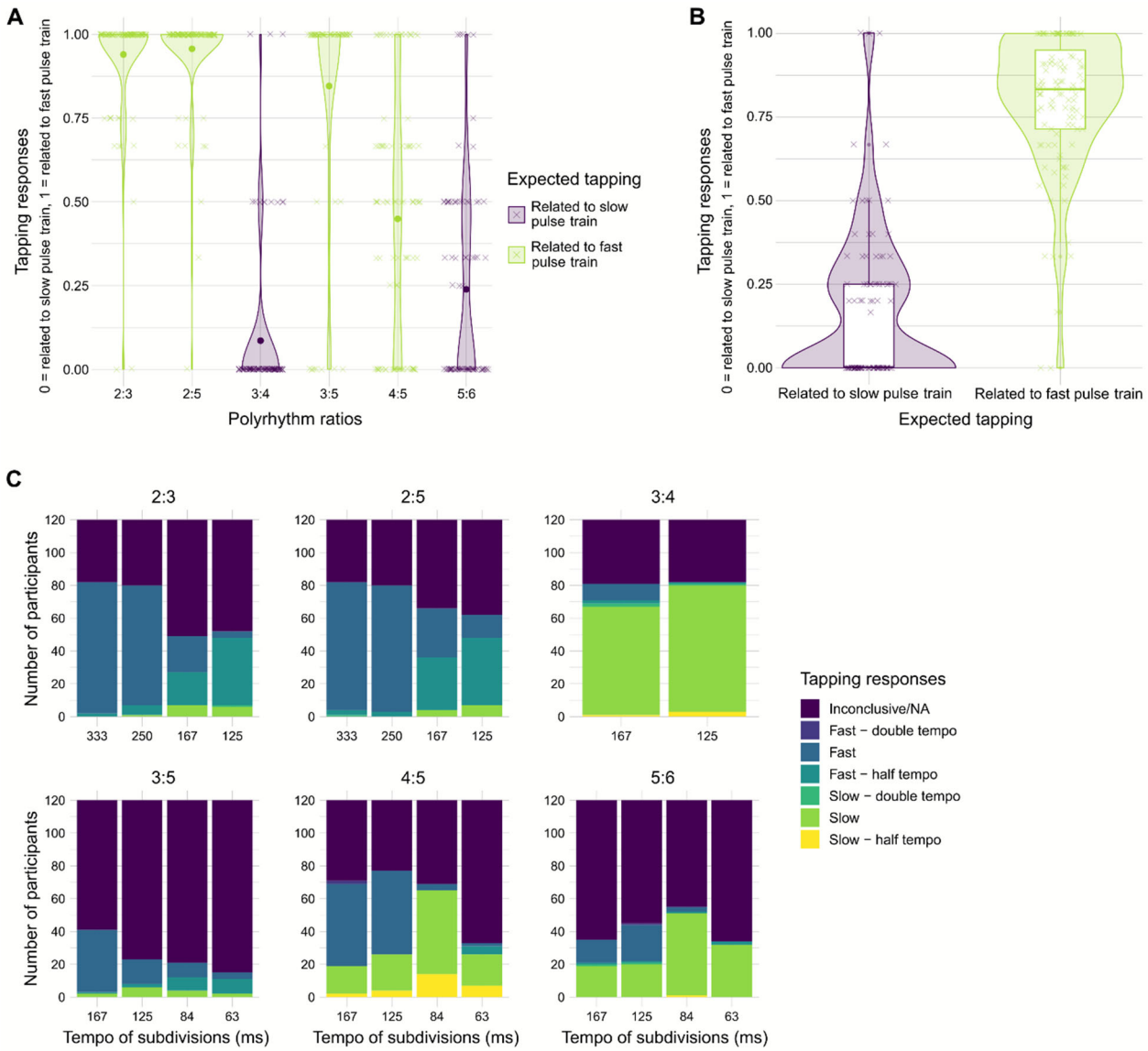
In the 4:5 polyrhythm with subdivision tempi of 167 ms and 125 ms most participants engaged in *fast pulse train* tapping, which is consistent with our hypothesis. However, the number of participants who preferred the pulse train with more complex grouping of subdivisions was larger compared to the polyrhythms described above. With subdivisions shorter than 84 ms, a clear preference shift was observed in favor of the *slow pulse train*. We propose two possible explanations for this shift. The first is that at these faster tempi, pulse trains are perceived as subdivisions, not because the tempo of the *fast pulse train* is too fast to tap, but because the tempo at the subdivision level is too fast to allow the grouping necessary to establish a sense of a beat. The second explanation is that, since the second beat of the *fast pulse train* precedes the second beat of the *slow pulse train* by only 84 ms and 63 ms respectively, the *fast pulse train* may serve as an up-beat to the *slow pulse train* which in turn becomes the perceived beat (Povel & Okkerman, 1981; Repp et al., 2012; see also the iambic-trochaic law, Hayes, 1995). With the present tapping data, we cannot determine the extent to which one or both of these explanations apply.

In the slower tempi of the 5:6 polyrhythm, results showed no preference for either of the pulse trains. Responses to the faster tempi of the 5:6 polyrhythm, however, mirrored responses to the faster tempi of the 4:5 polyrhythm. This result suggests that subdivision grouping does not determine beat perception at the faster tempi of these two polyrhythms. This may be explained by the fact that the metrical structures inherent in the 4:5 (irregular:binary) and 5:6 (binary/ternary:irregular) polyrhythms are more complex. In both polyrhythms, the two constituent pulse trains contain larger groups of subdivisions (≥ 4), and consequently, the subdivision levels are represented by a lower proportion of physical elements (see Table S1). These findings suggest a limit for subdivision grouping as predictor for beat perception and indicate that the propensity towards simpler subdivision grouping may only apply in cases where the auditory stimulus induces a strong feeling of a beat. This is less likely in complex metrical structures containing large groups of subdivisions.

In sum, our findings demonstrate that neither the number of elements in a pulse train nor the overall cycle duration determines the point at which participants switch tapping strategy. Rather, the transition points are tightly linked to the tempo of the subdivisions, and the most salient beat is represented by the pulse train with the simplest grouping of subdivisions. However, for more complex ratios, such as 4:5 and 5:6, this subdivision effect is less clear.

Figure 5

Tapping Responses of the Experiment Presenting Polyrhythms of Different Ratios



Note. A) The dependent variable coded preferences for tapping in relation to the slow (0) or fast (1) pulse train. Mean across tempi are shown per participant (×) and as grand means for each ratio (•). Tapping preferences were expected to be related to the slow pulse train in polyrhythms with simpler subdivision grouping in the slow pulse train (dark purple) and to the fast pulse train in polyrhythms with simpler subdivision grouping in the fast pulse train (bright green). B) Mean per participant for polyrhythms with simpler subdivision grouping in the slow pulse train (3:4, 5:6) in dark purple and polyrhythms with simpler subdivision grouping in the fast pulse train (2:3, 2:5, 3:5, 4:5) in bright green. C) Metrical categories of the tapping responses for the individual tempi in each of the polyrhythm ratios. The category "Inconclusive/NA" includes consistent tapping to the cycle and subdivision levels and to the polyrhythm itself as well as inconsistent and no tapping.

Pitch Experiment

As hypothesized, our results showed that pulse trains with lower pitch were more likely to be perceived as the beat. This was indicated by two individual Wilcoxon tests with significant effects of *pitch* in 2:3 and 3:4 polyrhythms (Z = -4.98, p < .001, r = .56; Z = -5.52, p < .001, r = .62, respectively; Figure 6). Comparable to the previous two experiments, participants preferred to tap in time with the pulse train with binary subdivision grouping, i.e., the *fast pulse train* in 2:3 polyrhythms and the *slow pulse train* in 3:4 polyrhythms. Attenuating the loudness of these pulse trains did not significantly affect these preferences. Figure 6 shows that participants tapped more consistently with the pulse train with binary subdivision grouping, especially when the lower pitch was in this pulse train. When the lower pitch was in the pulse train with ternary subdivisions, i.e., *slow* in 2:3 and *fast* in 3:4 polyrhythms, participants' tapping responses were split. Our interpretation is that participants either chose to tap in time with the pulse train with binary subdivision grouping, or to tap in time with the lower pitched pulse train. The preference for synchronizing with the lower pitched pulse train is in line with findings from studies showing that when separating the pulse trains of polyrhythms by a musical fourth, the beat is more likely to be perceived in the lower pitched pulse train (Handel, 1984; Handel & Lawson, 1983; Handel & Oshinsky, 1981). However, the stimuli in these studies were composed of 440 and 586 Hz pulse trains, whereas the current study used an interval of two octaves with peak frequencies of 262 and 1047 Hz. With the increased interval and the low-pitched pulse train reaching frequencies in the bass range, our stimuli more clearly speak for an effect of bass superiority.

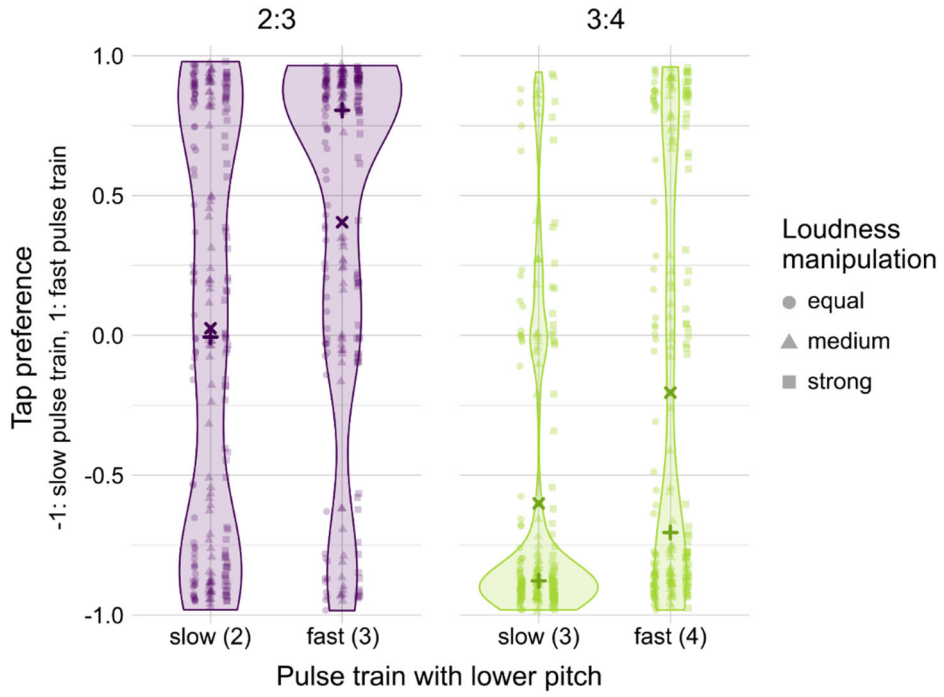
Two individual Kruskal-Wallis tests revealed no significant effect of *loudness* in 2:3 and 3:4 polyrhythms, respectively (X^2 = 1.32, p = .516 and X^2 = 3.93, p = .140). However, the control stimuli with equal pitch but the same *loudness* manipulation as the main stimuli revealed a significant effect of *loudness* (Figure S4). We conclude that in the current paradigm, beat perception is more strongly affected by the combined preferences for lower pitch and simpler subdivisions than by the loudness of the individual pulse trains.

Taken together, our findings suggest that we prefer to synchronize movements with polyrhythm pulse trains that consist of lower pitched sounds and simpler metrical subdivisions. Compared to these effects of low pitch and subdivisions, subtle to moderately strong loudness differences between pulse trains seem to have a smaller effect on beat preferences. Whether pitch or subdivision grouping has a stronger effect on beat perception in polyrhythms likely depends on individual factors, such as taste and familiarity, as well as on physical attributes of the stimulus, such as the size of the pitch interval between the two pulse trains, their peak frequencies, and the ratio of the polyrhythm (Moelants & van Noorden, 2005). Physiological not mutually exclusive factors that drive the bass superiority effect include tactile stimulation (Hove et al., 2020), vestibular stimulation (Todd & Lee, 2015), and encoding in the auditory pathway (Hove et al., 2014). Other factors influencing the close connection between bass, beat, and movement might be learned by exposure to music

with important rhythmic information produced by bass instruments. Whether other factors promoting the bass superiority effect are innate and potentially driven by evolutionary pressures remains an open question.

Figure 6

Tapping Responses of the Experiment Manipulating the Pitch of Individual Pulse Trains



Note. Individual data points per participant, *pitch*, and *loudness* condition of tapping preferences with 2:3 and 3:4 polyrhythms. Medians (+) and means (x) are computed over all *loudness* conditions. Values close to -1 indicate that participants consistently tapped in time with the slow pulse train, whereas values close to 1 indicate that participants consistently tapped in time with the fast pulse train.

General Discussion

In three online experiments (Tempo, Ratio, and Pitch), we investigated beat perception by using polyrhythms and demonstrated a yet neglected compelling influence of beat subdivisions on beat perception. Specifically, participants preferred to tap in time with beats that could be subdivided into binary groups as compared to ternary or irregular groups. We showed that this preference for binary subdivision grouping is stable across various polyrhythmic ratios and a wide range of tempi, but influenced by a preference for low pitch. Our findings highlight the importance of metrical structures, which need to be taken into account in order to understand beat perception—an insight that has been overlooked in previous polyrhythm studies which mostly focused on characteristics of the constituent pulse trains (Beauvillain, 1983; Beauvillain & Fraisse, 1984; Handel, 1984; Moelants & van Noorden, 2005; Pressing et al., 1996; van Noorden & Moelants, 1999). In

contrast to studies focusing only on pulse train and cycle tapping responses, we focused on subdivision grouping and analyzed tapping responses at half, original, and double tempo of the polyrhythms' pulse trains to account for all levels of the competing metrical structures. As such, our findings refine the understanding of temporal grouping principles in rhythm perception by suggesting that we lean towards simplicity when spontaneously grouping subdivisions, the most basic units within metrical structures of music.

We propose that subdivisions are the basic unit of beat perception, and that the principle underlying the grouping of subdivisions reflects a propensity towards simplicity. Preferences for structuring time in binary units are not only found in polyrhythms, but also in more general aspects of timing in our daily lives. Binary rhythms are pervasive in human behavior and perception, such as bipedal walking (Larsson, 2014; Larsson et al., 2019) and the "tick-tock" effect (Bolton, 1894; Brochard et al., 2003). Preferences for binary groupings of beats appear in early stages of development (Bergeson & Trehub, 2006) and in adults (Drake & Bertrand, 2001), and simpler grouping of beats are statistical universals across musical cultures (Savage et al., 2015). A binary categorization of events emerges even spontaneously when irregular rhythms are imitated from participant to participant in iterated learning paradigms (Ravignani et al., 2016), likely due to an automatic subdivision of the beat into two intervals (Repp, 2008). Indeed, sensorimotor synchronization to a periodic beat is generally more consistent when it is subdivided into two events, compared to three, four or five events); an advantage that can be observed in the brain when measuring the steady-state evoked potentials related to duplets (Celma-Miralles et al., 2021). All these studies align well with our interpretation that beat perception follows a subdivision hierarchy in favor of binary groupings, the simplest organization of temporal information.

Our findings are consistent with the general conclusion drawn from the seminal studies on polyrhythm perception by Handel and colleagues: rhythmic interpretation is contextual and depends on a number of factors, including tempo, ratio, relative intensity, and pitch (Handel, 1984; Handel & Lawson, 1983; Handel & Oshinsky, 1981; Oshinsky & Handel, 1978). However, we do not agree with Handel's notion that "There are no basic units of rhythm, and no theories based on grouping principles or tree representations seem able to encompass rhythmic diversity" (Handel, 1984, p. 481). Our Tempo Experiment showed that binary grouping of subdivisions was preferred regardless of whether the beat corresponded to the fast or the slow pulse train (Figures 3 and 4). Thus, a principle of simpler grouping pertains to the subdivisions, not the beats, making the subdivision the basic unit of rhythm. Our Ratio Experiment expanded the validity of this grouping principle to more complex polyrhythms and outlined its limits. When binary grouping of subdivisions was not present in the polyrhythm, listeners preferred ternary to irregular grouping of subdivisions or they constructed a binary subdivision level by concatenating two cycle repetitions (Figure 5). The preference for simpler subdivision grouping became less clear for our most complex polyrhythms (4:5, 5:6). Our Pitch Experiment showed that the preference for simpler grouping of subdivisions decreased when the pulse train with more complex grouping of subdivisions was presented at a lower pitch (Figure 6). The 6 and 12 LKFS loudness differences between the pulse trains had a surprisingly small effect on beat preferences compared to the preference for

lower pitched pulse trains and for grouping the basic rhythm units in pairs. Therefore, our three experiments support the notion that rhythmic interpretation is contextual. Importantly, our findings demonstrate that we have to extend the list of contextual factors to include grouping at the subdivision level as a critical determinant of beat perception in polyrhythms.

The finding that people tap in a way that reflects the simplest possible grouping of subdivisions may explain discrepancies between findings in previous beat perception studies. For instance, using real musical excerpts, McKinney and Moelants (2006) observed that the most saliently perceived tempo for some excerpts deviated largely from the preferred tempo of 120 BPM (Parncutt, 1994; van Noorden & Moelants, 1999). Our results showed that tapping preference transition points were tightly linked to the tempo of the subdivisions, rather than to the tempo of the pulse trains. Similarly, temporal preferences in beat perception in real music may not be constrained by the tempo of the beat itself, but by the tempo and grouping of the underlying subdivisions.

The knowledge about the universality and limits of the human propensity towards simple subdivision grouping in temporal processing still has to be refined. To this aim, future research should assess the generalizability and exceptions of binary subdivision preferences in beat perception in real music, sensorimotor synchronization with rhythms perceived in different modalities, and timing in social interactions involving movement, dance, music or language. In future studies, we will expand the current findings in a number of exploratory analyses mentioned in the preregistration for the present study (https://aspredicted.org/yi5si.pdf). We will examine the influences of musical training, individual spontaneous motor tempo, native languages, and cultural background on subdivision grouping preferences in rhythm processing.

Studying the hierarchical organization of temporal patterns is an interdisciplinary effort. From the perspective of music perception and cognition, our findings reveal the importance of lower hierarchical levels, such as subdivisions, to grouping processes at higher levels, such as beat and meter. Our study adds to evidence suggesting that in everyday behavior, we prefer to structure time in binary units."

References

Abecasis, D., Brochard, R., Granot, R., & Drake, C. (2005). Differential brain response to metrical accents in isochronous auditory sequences. *Music Perception*, *22*(3), 549–562. https://doi.org/10.1525/mp.2005.22.3.549

Beauvillain, C. (1983). Auditory perception of dissonant polyrhythms. *Perception & Psychophysics*, *34*(6), 585–592. https://doi.org/10.3758/BF03205915

Beauvillain, C., & Fraisse, P. (1984). On the temporal control of polyrhythmic performance. *Music Perception*, *1*(4), 485–499. https://doi.org/10.2307/40285274

Berens, P. (2009). CircStat: A MATLAB toolbox for circular statistics. *J Stat Softw*, *31*(10), 1–21.

Bergeson, T. R., & Trehub, S. E. (2006). Infants' perception of rhythmic patterns. *Music Perception*, *23*(4), 345–360. https://doi.org/10.1525/mp.2006.23.4.345

Bolton, T. L. (1894). Rhythm. *The American Journal of Psychology*, *6*(2), 145–238. JSTOR. https://doi.org/10.2307/1410948

Bouwer, F. L., van Zuijen, T. L., & Honing, H. (2014). Beat processing is pre-attentive for metrically simple rhythms with clear accents: An ERP study. *PloS One, 9*(5), e97467–e97467. https://doi.org/10.1371/journal.pone.0097467

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT press.

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. *Psychological Science*, *14*(4), 362–366. https://doi.org/10.1111/1467-9280.24441

Celma-Miralles, A., Kleber, B. A., Toro, J. M., & Vuust, P. (2021). Neural entrainment facilitates duplets: Frequency-tagging differentiates musicians and non-musicians when they tap to the beat. *BioRxiv*, 2021.02.15.431304. https://doi.org/10.1101/2021.02.15.431304

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Perception & Psychophysics*, *53*(1), 25–33. https://doi.org/10.3758/BF03211712

Drake, C., & Bertrand, D. (2001). The quest for universals in temporal processing in music. *Annals of the New York Academy of Sciences*, *930*(1), 17–27. https://doi.org/10.1111/j.1749-6632.2001.tb05722.x

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*(3), 251–288. https://doi.org/10.1016/S0010-0277(00)00106-2

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception*, *18*(1), 1–23. https://doi.org/10.2307/40285899

Duke, R. A. (1989). Musicians' perception of beat in monotonic stimuli. *Journal of Research in Music Education*, *37*(1), 61–71. https://doi.org/10.2307/3344953

Fidali, B. C., Poudrier, È., & Repp, B. H. (2013). Detecting perturbations in polyrhythms: Effects of complexity and attentional strategies. *Psychological Research*, *77*(2), 183–195. https://doi.org/10.1007/s00426-011-0406-8

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *Psychology of Music* (pp. 149–180). Academic Press. https://doi.org/10.1016/B978-0-12-213562-0.50010-3

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–836.

Fujioka, T., Fidali, B. C., & Ross, B. (2014). Neural correlates of intentional switching from ternary to binary meter in a musical hemiola pattern. *Frontiers in Psychology*, *5*, art. no. 1257. https://doi.org/10.3389/fpsyg.2014.01257

Fujioka, T., Zendel, B. R., & Ross, B. (2010). Endogenous neuromagnetic activity for mental hierarchy of timing. *Journal of Neuroscience*, *30*(9), 3458–3466. https://doi.org/10.1523/JNEUROSCI.3086-09.2010

Handel, S. (1984). Using polyrhythms to study rhythm. *Music Perception*, *1*(4), 465–484. https://doi.org/10.2307/40285273

Handel, S., & Lawson, G. R. (1983). The contextual nature of rhythmic interpretation. *Perception & Psychophysics*, *34*(2), 103–120. https://doi.org/10.3758/BF03211335

Handel, S., & Oshinsky, J. S. (1981). The meter of syncopated auditory polyrhythms. *Perception & Psychophysics*, *30*(1), 1–9. https://doi.org/10.3758/BF03206130

Harrison, P. M. (2020). psychTestR: An R package for designing and conducting behavioural psychological experiments. *Journal of Open Source Software*, *5*(49), 2088. https://doi.org/10.21105/joss.02088

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.

Hirsh, I. J., Monahan, C. B., Grant, K. W., & Singh, P. G. (1990). Studies in auditory timing: 1. Simple patterns. *Perception & Psychophysics*, *47*(3), 215–226. https://doi.org/10.3758/BF03204997

Honing, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences*, *1252*, 85–91. https://doi.org/10.1111/j.1749-6632.2011.06402.x

Hove, M. J., Marie, C., Bruce, I. C., & Trainor, L. J. (2014). Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proceedings of the National Academy of Sciences*, *111*(28), 10383–10388. https://doi.org/10.1073/pnas.1402039111

Hove, M. J., Martinez, S. A., & Stupacher, J. (2020). Feel the bass: Music presented to tactile and auditory modalities increases aesthetic appreciation and body movement. *Journal of Experimental Psychology: General*, *149*(6), 1137–1147. https://doi.org/10.1037/xge0000708

Jagacinski, R. J., Marshburn, E., Klapp, S. T., & Jones, M. R. (1988). Tests of parallel versus integrated structure in polyrhythmic tapping. *Journal of Motor Behavior*, *20*(4), 416–442. https://doi.org/10.1080/00222895.1988.10735455

Jagacinski, R. J., Rizzi, E., Kim, T. H., Lavender, S. A., Speller, L. F., & Klapp, S. T. (2016). Parallel streams versus integrated timing in multilimb pattern generation: A test of Korte's Third Law. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(11), 1703.

Klapp, S. T., Hill, M. D., Tyler, J. G., Martin, Z. E., Jagacinski, R. J., & Jones, M. R. (1985). On marching to two different drummers: Perceptual aspects of the difficulties. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(6), 814–827. https://doi.org/10.1037/0096-1523.11.6.814

Klapp, S. T., Nelson, J. M., & Jagacinski, R. J. (1998). Can people tap concurrent bimanual rhythms independently? *Journal of Motor Behavior*, *30*(4), 301–322.

Large, E. W., Fink, P., & Kelso, J. A. S. (2002). Tracking simple and complex sequences. *Psychological Research*, *66*(1), 3–17. https://doi.org/10.1007/s004260100069

Larsson, M. (2014). Self-generated sounds of locomotion and ventilation and the evolution of human rhythmic abilities. *Animal Cognition*, *17*(1), 1–14. https://doi.org/10.1007/s10071-013-0678-z

Larsson, M., Richter, J., & Ravignani, A. (2019). Bipedal steps in the development of rhythmic behavior in humans. *Music & Science*, *2*. https://doi.org/10.1177/2059204319892617

Lenc, T., Keller, P. E., Varlet, M., & Nozaradan, S. (2018). Neural tracking of the musical beat is enhanced by low-frequency sounds. *Proceedings of the National Academy of Sciences*, *115*(32), 8221–8226. https://doi.org/10.1073/pnas.1801421115

London, J. (2002). Cognitive constraints on metric systems: Some observations and hypotheses. *Music Perception*, *19*(4), 529–550. https://doi.org/10.1525/mp.2002.19.4.529

McKinney, M. F., & Moelants, D. (2006). Ambiguity in tempo perception: What draws listeners to different metrical levels? *Music Perception*, *24*(2), 155–166. https://doi.org/10.1525/mp.2006.24.2.155

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: A neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140093.

Michaelis, K., Wiener, M., & Thompson, J. C. (2014). Passive listening to preferred motor tempo modulates corticospinal excitability. *Frontiers in Human Neuroscience*, *8*, art. no. 252. https://doi.org/10.3389/fnhum.2014.00252

Moelants, D., & van Noorden, L. (2005). The influence of pitch interval on the perception of polyrhythms. *Music Perception: An Interdisciplinary Journal*, *22*(3), 425–440. https://doi.org/10.1525/mp.2005.22.3.425

Møller, C., Stupacher, J., Celma-Miralles, A., & Vuust, P. (2020). Polyrhythms - Online tapping study, MIB, Aarhus 2020 [preregistration]. https://aspredicted.org/yi5si.pdf

Ollen, J. E. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings* [The Ohio State University]. https://etd.ohiolink.edu/apexprod/rws_olink/r/1501/10?clear=10&p10_accession_num=osu1161705351

Oshinsky, J. S., & Handel, S. (1978). Syncopated auditory polyrhythms: Discontinuous reversals in meter interpretation. *The Journal of the Acoustical Society of America*, *63*(3), 936–939. https://doi.org/10.1121/1.381773

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*(4), 409–464. https://doi.org/10.2307/40285633

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*, art. no. 57. https://doi.org/10.3389/fnsys.2014.00057

Pitt, M. A., & Monahan, C. B. (1987). The perceived similarity of auditory polyrhythms. *Perception & Psychophysics*, *41*(6), 534–546. https://doi.org/10.3758/BF03210488

Povel, D.-J., & Okkerman, H. (1981). Accents in equitone sequences. *Perception & Psychophysics*, *30*(6), 565–572. https://doi.org/10.3758/BF03202011

Pressing, J., Summers, J., & Magill, J. (1996). Cognitive multiplicity in polyrhythmic pattern performance. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(5), 1127–1148. https://doi.org/10.1037/0096-1523.22.5.1127

Ravignani, A., Delgado, T., & Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*, *1*(1), 1–7. https://doi.org/10.1038/s41562-016-0007

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*(4), 355–370. https://doi.org/10.1080/00222890309603156

Repp, B. H. (2007). Perceiving the numerosity of rapidly occurring auditory events in metrical and nonmetrical contexts. *Perception & Psychophysics*, *69*(4), 529–543. https://doi.org/10.3758/BF03193910

Repp, B. H. (2008). Multiple temporal references in sensorimotor synchronization with metrical auditory sequences. *Psychological Research*, *72*(1), 79–98. https://doi.org/10.1007/s00426-006-0067-1

Repp, B. H., London, J., & Keller, P. E. (2012). Distortions in reproduction of two-interval rhythms: When the "attractor ratio" is not exactly 1:2. *Music Perception*, *30*(2), 205–223. https://doi.org/10.1525/mp.2012.30.2.205

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, *112*(29), 8987–8992. https://doi.org/10.1073/pnas.1414495112

Stupacher, J., Hove, M. J., & Janata, P. (2016). Audio features underlying perceived groove and sensorimotor synchronization in music. *Music Perception*, *33*(5), 571–589. https://doi.org/10.1525/mp.2016.33.5.571

Summers, J. (2002). Practice and training in bimanual coordination tasks: Strategies and constraints. *Brain and Cognition*, *48*(1), 166–178. https://doi.org/10.1006/brcg.2001.1311

Temperley, N. M. (1963). Personal tempo and subjective accentuation. *The Journal of General Psychology*, *68*(2), 267–287. https://doi.org/10.1080/00221309.1963.9920534

Todd, N. P., & Lee, C. S. (2015). The sensory-motor theory of rhythm and beat induction 20 years on: A new synthesis and future perspectives. *Frontiers in Human Neuroscience*, *9*, art. no. 444. https://doi.org/10.3389/fnhum.2015.00444

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2013). The impact of the bass drum on human dance movement. *Music Perception*, *30*(4), 349–359. https://doi.org/10.1525/mp.2013.30.4.349

van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, *28*(1), 43–66. https://doi.org/10.1076/jnmr.28.1.43.3122

Varlet, M., Williams, R., & Keller, P. E. (2020). Effects of pitch and tempo of auditory rhythms on spontaneous movement entrainment and stabilisation. *Psychological Research*, *84*(3), 568–584. https://doi.org/10.1007/s00426-018-1074-8

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. Proceedings of the National Academy of Sciences of the United States of America, 107(13), 5768–5773. https://doi.org/10.1073/pnas.1000121107