Attention modulates neural measures associated with beat perception

Gibbings, Aaron; Henry, Molly J; Cruse, Damian; Stojanoski, Bobby; Grahn, Jessica A

Link to publication on Research at Birmingham portal

Attention modulates neural measures associated with beat perception

Aaron Gibbings[1,2], Molly J. Henry[1,2,3], Damian Cruse[1,4], Bobby Stojanoski[1], & Jessica A. Grahn[1,2*]

[1]The Brain and Mind Institute, University of Western Ontario, London, ON, N6A 3K7, Canada

[2]Department of Psychology, University of Western Ontario, London, ON, N6A 3K7, Canada

[3]Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

[4]School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

---

[*] Authors' present addresses:
MJH: Toronto Metropolitan University, 350 Victoria St. Toronto, ON, Canada; Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main, Germany
BS: University of Ontario Institute of Technology, 2000 Simcoe Street North, Oshawa, ON L1G 0C5

Abstract

A growing body of evidence suggests that steady-state evoked potentials may be a useful measure of beat perception, particularly when obtaining traditional, explicit measures of beat perception is difficult, such as <u>with</u> infants or <u>non-human</u> animals. Although attending to a stimulus is not necessary for most traditional applications of steady-state evoked potentials, it is unknown how attention affects steady-state evoked potentials that arise in response to beat perception. Additionally, most applications of steady-state evoked potentials to measure beat perception have used repeating rhythms or real music. Therefore, it is unclear how the steady-state response relates to the robust beat perception that occurs with non-repeating rhythms. Here, we used electroencephalography to record participants' brain activity as they listened to non-repeating musical rhythms while either attending to the rhythms or while distracted by a concurrent visual task. Non-repeating auditory rhythms elicited steady-state evoked potentials at perceived beat frequencies (perception was validated in a separate sensorimotor synchronization task) that were larger when participants attended to the rhythms compared to when they were distracted by the visual task. Therefore, although steady-state evoked potentials appear to index beat perception to non-repeating musical rhythms, this technique may be limited to when participants are known to be attending to the stimulus.

Keywords: Beat perception, Rhythm, Attention, Steady-state evoked potential, Electroencephalography

Listening to musical rhythm is typically accompanied by the perception of a "beat", a regular underlying pulse that stands out to the listener and to which many listeners spontaneously synchronize body movements, like swaying or clapping (McAuley, 2010; van Noorden & Moelants, 1999). In music, the beat is often emphasized acoustically by accenting certain events (e.g., making them louder or adding a drum strike), but the perception of a beat also emerges in rhythms in which no acoustic cues indicate the location of the beat (Parncutt, 1994; Povel & Essens, 1985). In this case, the beat percept arises based only on the temporal structure of the rhythm (Bouwer et al., 2018), and therefore the beat percept involves active construction on the part of the listener (Bouwer & Honing, 2015; Grahn, 2012; Grahn & Rowe, 2009, 2013; Lenc et al., 2020). Although the appreciation of musical rhythm and the perception of a beat seem to be culturally universal, two intriguing scientific questions in this field are the developmental trajectory and species-specificity of beat perception in musical rhythms devoid of explicit acoustic accents (Honing et al., 2014; Merchant et al., 2015).

A major impediment to answering these questions is that we cannot generally ask infants or non-human animals to behaviorally indicate whether they perceive a beat, either by producing movements synchronized to the beat (Hagmann & Cook, 2010; Rouse et al., 2016) or rating the strength of the perceived beat. This impediment could be overcome by a neural measure of beat perception (Honing et al., 2009, 2012, 2018; Winkler et al., 2009), evoked independently of the listener's ability to follow instructions, produce an appropriate output, or even maintain attention to the rhythm. Electroencephalography (EEG) measures have been proposed to be instrumental in probing beat perception, as EEG is non-invasive, is sensitive to the temporal dynamics of neural activity, and is silent (Bouwer et al., 2014). Previous work has focused on the mismatch negativity (MMN), an EEG event-related potential that occurs in response to violations of expectations (Bouwer & Honing, 2015). One major advantage of the MMN is that it is elicited pre-attentively, that is, independent of whether the listener attends to

a stimulus, and for that reason can be assessed in sleeping infants (Winkler et al., 2009) or non-human primates (Merchant et al., 2015). On the other hand, the absence of an MMN in the context of beat perception cannot easily differentiate between the possibilities that a) the listener did not perceive a beat or b) the beat-based expectation that was violated required attention to be processed. Moreover, the necessity to include events (or omissions thereof) that violate expectations may actually interfere with the beat percept itself. The current study evaluated another possible neural measure that may be used to probe beat perception: steady-state evoked potentials at beat-relevant frequencies (Gilmore & Russo, 2021; Lenc et al., 2018; Nave et al., 2022; Nozaradan et al., 2011, 2012b; Tierney & Kraus, 2015).

Steady-state evoked potentials are electrophysiological neural responses that faithfully track periodic environmental stimuli such as flickering lights or amplitude-modulated sounds (Regan, 1977). In response to musical rhythms, steady-state evoked potentials occur at the frequency of the perceived beat (as well as its harmonics and subharmonics). Therefore, recent studies have explored the potential of steady-state evoked potentials to index beat perception (Gilmore & Russo, 2017; Lenc et al., 2018; Nozaradan et al., 2011, 2012b, Nave et al., 2022). These studies have revealed that steady-state evoked response magnitudes do not simply reflect the physical stimulus content (de Fleurian et al., 2017), but – like beat perception – are modulated by musical experience (Doelling & Pöeppel, 2015; Tierney & Kraus, 2013), modality (Gilmore & Russo, 2021), and prior auditory or motor context (Lenc et al., 2018; Nozaradan et al., 2016). However, one question that to our knowledge has not been clearly answered is the extent to which steady-state evoked potentials as a metric of beat perception are affected by *attention* being directed towards a rhythm stimulus. Answering this question is relevant for assessing the potential to use steady-state evoked potentials to measure beat perception in infants or non-human primates, as attention cannot always be guaranteed. Here, we asked whether steady-state evoked potentials to auditory rhythms are sensitive to whether

attention is directed towards the rhythm versus towards a concurrent visual task. We also assess whether these attention effects are general to all rhythms or are specific to when a beat is being perceived.

Lack of attention to a rhythm might affect the magnitude of steady-state evoked potentials in one of two ways. It may reduce the overall magnitude of auditory steady-state responses (Keitel et al., 2011; Saupe et al., 2009), regardless of whether a beat is being perceived or not. Alternatively, reduced attention to a rhythm may weaken or eliminate beat perception (Bouwer & Honing, 2015), reducing neural response magnitudes at beat-related frequencies, but only for rhythms in which a beat is perceived in the first place. Thus, our experiment compared EEG power at beat-related frequencies for strong- and weak-beat rhythms when participants were either attending to or distracted from the rhythms. We assessed whether diversion of attention reduced power for strong- and weak-beat rhythms alike, or whether lack of attention specifically reduced the EEG power for strong-beat rhythms, due to diminished strength of the perceived beat. If attention acts on the overall magnitude of steady-state evoked potentials, overall power should be reduced for both strong- and weak-beat rhythms. In this case, we would argue that steady-state evoked responses might be a productive noninvasive measure of beat strength for use with infants or non-human primates, because, although lack of attention reduces neural responses, it does so for all rhythms. However, if diverting attention reduces the strength of the perceived beat, then beat-related power for strong-beat rhythms should be selectively affected. In this case, we would caution against relying on steady-state evoked potentials in cases where attention cannot be guaranteed, as beat perception abilities would likely be underestimated.

We also addressed a secondary theoretical question. Most previous work on electrophysiological markers of beat perception has used rhythms made up of short, repeating patterns (Cirelli et al., 2016; Lenc et al., 2018; Nozaradan et al., 2012, 2016). However, robust

beat perception occurs in response to non-repeating rhythms (Lenc et al., 2018), which are easier than music to acoustically control but have the benefit over repeating rhythms that they allow for manipulation of specific parameters of theoretical interest (e.g., different types of temporal, intensity, or pitch accents, hysteresis effects, etc.). Therefore, we provide an important proof-of-principle by examining neural responses, and attentional modulation thereof, in response to non-repeating rhythms.

Thus, the current EEG study evaluated steady-state evoked potentials elicited at beat-related frequencies by non-repeating musical rhythms. We compared EEG power to rhythms that elicited a strong beat percept to weak-beat control rhythms while participants either attended to the rhythms or were distracted by a concurrent visual task. Importantly, to validate our strong and weak beat rhythm manipulation, and to select beat rates of interest in an assumption-free way, we collected finger-tapping data to identify perceived beat rates in a separate behavioral testing session.

Materials and Methods

Participants

Thirty-nine adults (M_{age} = 21.10 years, SD = 2.96; 24 female) participated in the study. The study involved one EEG session and one behavioral session, conducted a minimum of three days apart. Thirty-one participants completed the EEG session, and 34 participants completed the behavioral testing; the samples did not fully overlap (27 participants completed both sessions; M_{age} = 21.37 years, SD = 3.45; 16 female) because of either technical problems during testing or attrition between testing sessions. Fourteen participants had musical training (M_{Music} = 6.71 years, SD = 3.42). None of the participants had a history of hearing disorder, and all reported being in good health on the day(s) of testing. The Health Sciences Research Ethics Board at Western University approved the study.

Auditory stimuli

Auditory stimuli were created and presented with Matlab 7.14.0 (MathWorks) using the Psychtoolbox extension (Brainard, 1997; Pelli, 1997). All sequences were composed of brief tones (50-ms, 990-Hz sine tones, amplitude modulated at 40 Hz, 5-ms linear onset/offset ramps); The original intent of the 40-Hz modulation was to test for differences in the 40-Hz auditory steady-state response that might index differences in attention. However, the short durations of the tones proved too brief to measure the 40-Hz steady-state response reliably, thus this line of analysis was not explored. Stimuli for the "strong-beat" condition consisted of 40 unique rhythmic sequences, each lasting 33.64 seconds. Inter-onset intervals (IOIs) between tones were 227, 454, 682, or 909 ms (values derived from 1.1 Hz = 909.09 ms IOI, all intervals subsequently rounded to the nearest millisecond), which are related to each other by integer ratios (1:2:3:4). The minimum IOI (227 ms) corresponds to a frequency of 4.4 Hz. The events were grouped such that there was a tone onset every 909 ms (i.e., on the "downbeat" assuming a 4 /4-time signature), which leads to a perceived beat at either 1.1 Hz (assuming a 4/4 time signature) or 2.2 Hz (assuming a 2/4 time signature). We were therefore interested in neural responses at 1.1 Hz, 2.2 Hz, and 4.4 Hz. We confirmed that these "frequencies of interest" were present in the frequency-domain representations of the rhythms themselves: the average frequency spectrum of the rhythm envelopes (obtained via Hilbert transform, Fig. 1B) shows prominent peaks at these frequencies, and behavioral testing confirmed that participants perceived a beat at these expected rates.

Forty "weak-beat" rhythms were derived from the 40 strong-beat rhythms as follows: each 909-ms downbeat-to-downbeat interval was subdivided into four 227-ms periods that marked the possible tone-onset positions in the beat versions of the rhythms. Rhythms were modified by adjusting the onset times of the tones in the second and fourth position when those

tones were present (i.e., tones with onsets at 227 ms or 682 ms after the downbeat event). The tone onsets were jittered by one of five durations: 31.8, 50.0, 72.3, 100.0, or 131.8 ms. Random selection during the stimulus creation determined which of the five jitter durations was applied to each jittered onset. The jittered tones alternated between being moved early or late by the randomly selected duration (Fig. 1A). Using five jitter durations kept participants from becoming accustomed to a single jitter amount, and interpreting the mis-timing as "swing" (Merker, 2014; Prögler, 1995). Critically, the timing of tones in the first and third position was not altered. Thus, strong-beat rhythms and their weak-beat counterparts had identical onset times for tones that occurred in the first and third positions, which corresponded to the beat locations for beat rates of 1.1 Hz or 2.2 Hz, respectively. Shifting the intervening second and fourth onsets was intended to disrupt subjective perception of that beat (Demany & Semal, 2002) while maintaining periodicity at 1.1 Hz (marked by the unjittered first onsets) and 2.2 Hz (marked by the unjittered first and third onsets). The full set of stimuli are available at http://www.jessicagrahn.com/GibbingsHenryCruseStojanoskiGrahn.html.

Interestingly, the spectral power of the stimulus at the 1.1 Hz beat-rate did not significantly differ between strong-beat and weak-beat rhythms ($t(78) = 0.62$, $p = .54$, Fig. 1B), but did differ at 2.2 Hz: the weak-beat rhythms had a significantly *greater* 2.2-Hz power than the strong-beat rhythms ($t(78) = -15.21$, $p = 4.72e-25$, Fig. 1B). At 4.4 Hz, a beat-related frequency (but not one posited to be the perceived beat) the strong-beat rhythms had significantly greater spectral power than weak-beat rhythms ($t(78) = 26.09$, $p = 2.78e-40$, Fig. 1B).

A. Strong- and weak-beat rhythms

227.3 ms

Strong

Weak

Time (s)

909.0 ms

■ 50-ms sine tone
━ perceived beat location (1.1 Hz)
➤ temporal jitter (± 31.8, 50.0, 72.3, 100.0, 131.8 ms)
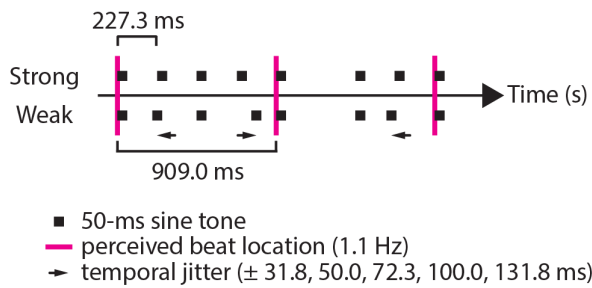
B. Beat-rate spectral stimulus power

0.015   0.017 ★   0.250 ★

Power (a.u.)   0.16

0

1   2   3   4   5   6
Frequency (Hz)

■ Strong beat   ■ Weak beat

C. Task schematic for attended and distracted conditions

Attended

Strong beat

or

Weak beat

+

Distracted

Strong beat

or

Weak beat

+1

145 ms

r A G x w ••• How many targets?

Time
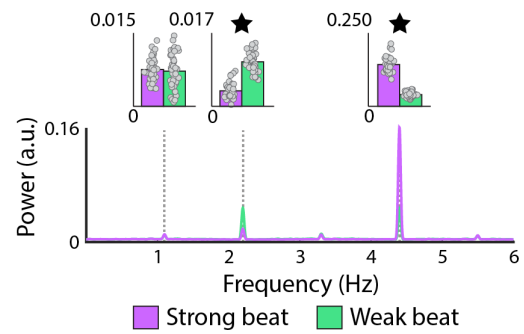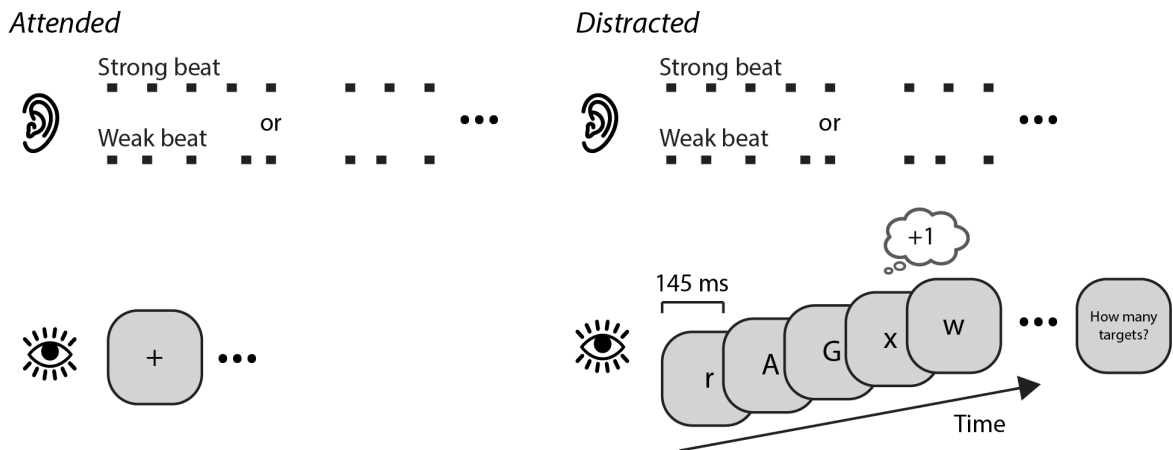
Figure 1. **Stimulus and task schematics. A) Strong-beat and weak-beat rhythms.** Short segments (2 bars) of an example strong- and weak-beat stimulus pair. Top shows two bars of a strong-beat rhythm, bottom shows two bars of a weak-beat rhythm, the latter of which was created by alternately shorting or lengthening the IOI preceding any tones occurring at the second or fourth positions by one of five jitter amounts. Arrows indicate which tones were adjusted and whether IOIs were shortened or lengthened. **B. Beat-rate spectral power**. Frequency spectra for beat (purple) and non-beat (green) rhythms. Spectral power for strong-beat and weak-beat rhythms was similar at 1.1 Hz, higher at 2.2-Hz for weak-beat rhythms, and higher at 4.4-Hz for strong-beat rhythms; note that inset bar graphs each have independent y-axis scales to emphasize condition differences. **C. Task schematic for attended and distracted conditions.** Rhythms were presented in two conditions: attended (left) and distracted (right). In the attended condition, listeners either fixated a cross or closed their eyes while listening attentively to the strong-beat and weak-beat rhythms. In the distracted condition, listeners performed a rapid serial visual presentations (RSVP) task in which they counted how many lower-case x's occurred in each trial while strong-beat and weak-beat rhythms were presented.

Procedure

Behavioral and EEG testing were completed in separate sessions, conducted at least 3 days apart. The order in which participants completed the behavioral and EEG testing sessions was counterbalanced across those participants that completed both sessions.

Behavioral session. Participants listened to the strong-beat and weak-beat auditory rhythms and tapped along with the beat as soon as they perceived it after the stimulus started. Auditory stimuli were presented at a comfortable volume for the participant using QuietComfort® 3 Acoustic Noise Cancelling® headphones. Finger taps were recorded using an ErgoDex DX1 Input System (South Dakota, USA). After completing four practice trials, participants completed four experimental blocks with 20 trials in each block; both strong-beat and weak-beat trials were presented within each block in pseudo-random order. The behavioral experiment lasted approximately one hour.

EEG session. During EEG recording, participants completed four types of listening blocks: *attended strong-beat*, *attended weak-beat*, *distracted strong-beat*, and *distracted weak-beat*, during which they heard the same rhythms that were used in the behavioral experiment (Fig. 1C). In the attended blocks, all participants were instructed to focus their attention on the auditory stimuli, and told they could either close their eyes or focus on a fixation cross (they could choose) presented on the computer screen while listening. In distracted blocks, participants performed a rapid, serial, visual presentation (RSVP) task while the auditory stimuli were presented simultaneously. During the RSVP task, participants monitored a series of visually presented letters for the lower-case target letter 'x'. The number of targets ranged between 4 and 8 (uniformly distributed). Both upper- and lower-case letters were used as targets and distractors. Lower case letters were 1.5 cm tall, and the upper-case letters were 2.25 cm tall. The visual stimuli were presented at a frequency of 6.896 Hz (every 145 ms), which is not a natural harmonic of the beat frequency in the auditory stimuli and is outside the range in which human listeners can perceive a beat (Repp, 2005). Therefore, we assumed that any entrainment of neural activity associated with the visual onset of the letters in the RSVP task would not interfere with the potential neural entrainment response to the auditory stimuli at <u>the</u>

frequencies of interest. At the end of each trial, the participant reported, by button press, if they saw an even or odd number of target letters.

All auditory stimuli were presented over earphones at a comfortable volume using E.A.R.Tone, 3A insert earphones (E.A.R. Auditory Systems). Each 33.64-second rhythm was preceded by 10 seconds of silence. After each attended trial there was a two-second inter-trial interval, while in the distracted condition the inter-trial interval ended when participants provided their response regarding whether an even or odd number of targets were presented.

Each participant completed 8 blocks (2 per condition) with 10 trials per block. Each block was approximately 8 minutes long. Participants were given a short break between blocks. The presentation order of each block was counterbalanced across participants, and the order of the stimuli within each block was randomized. The entire EEG session lasted approximately two hours.

EEG Recording. Participants were seated comfortably in a chair in front of a computer monitor. Participants were instructed to minimize any head or body movements, and reduce eye blinking and jaw clenching as much as possible during the recording. EEG was recorded using a 128-channel high-density electrode array from Electrical Geodesics Inc. (EGI; Eugene, OR, USA). Electrode impedances were kept below 50 kΩ, which is a standard level for this type of high-impedence system (EGI, Eugene, OR, USA). The signals were amplified and digitized using a sampling rate of 250 Hz, and referenced to the vertex electrode. Recording of the EEG signal was done with a MacPro 5.1 running NetStation 4.5.4.

Data Analysis

Behavioral tapping data

First, we discarded taps that occurred less than 454 ms after the start of a rhythm (before the second possible tone was heard). Then we discarded "double taps", defined as produced

intervals shorter than 50% of the minimum IOI (Drake & Botte, 1993; Martens, 2011), and missed taps, defined as produced intervals that were longer than twice the median interval produced by an individual. Overall, a total of 2.46% of inter-tap intervals (ITIs) were removed.

Tapping frequency was calculated as the mean inter-tap interval (ITI) of the cleaned interval series. Rarely, participants began tapping at one rate, but switched to tapping twice as fast (*i.e.*, "double-time") or half as fast (*i.e.,* "half-time") mid-trial. Trials in which participants switched tapping rates, termed "switch trials", were identified as trials in which mean ITI was between 510 and 850 ms (between the predicted beat rates of 454 and 909 ms). Switch trials were discarded prior to further analysis; the median proportion of discarded trials was 2.5% per condition.

We computed the *coefficient of variation* (CoV) as a measure of how consistently participants produced intervals. For each trial, CoV was computed by dividing the standard deviation of the ITIs by the mean ITI for each trial. CoV is thus a normalized measure of ITI variability. *Coefficient of asynchrony* (CoA) was calculated to determine how accurately each participant was able to synchronize with the beat. Asynchrony is the difference between each tap time and the nearest beat time in the stimulus, measured in milliseconds; when participants tapped at 1.1 Hz, beat times were defined as the onsets of tones in the first position of each bar, and when participants tapped at 2.2 Hz, beat times were defined as the onset times of tones that occurred in the first and third positions of each bar. CoA was normalized by dividing the mean asynchrony by the mean ITI. Thus, CoA indexes how well participants synchronized to the beat, as a proportion of the mean ITI for each trial. All dependent measures were averaged across trials separately for strong-beat and weak-beat rhythms.

RSVP Analysis

To determine whether participants successfully attended to the RSVP task, the proportion of correctly reported "odd" vs. "even" responses, averaged over trials, was compared to chance-level performance (50% correct) via a one-sample *t*-test. To determine if the type of auditory stimuli (strong- vs. weak-beat) affected attention to and performance of the RSVP task, a paired samples *t*-test compared proportion correct in the strong- vs. weak-beat conditions.

EEG Analysis

EEG data were analyzed using custom Matlab scripts and the Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). EEG data were re-referenced offline to an average reference, then high-pass filtered (0.5 Hz, 912 points, Hamming window) and low-pass filtered (50 Hz, 95 points, Blackman window). Trials were epoched from –1.5 to +35 s with respect to the onset of the 33.64-s auditory stimulus. Because of a technical error, the final trial of every participant's recording session was lost, leaving 79 total trials (lost trials were evenly distributed across conditions). Blinks, muscle activity, and electrical heart activity were removed from the signal using independent component analysis (ICA) using the Fieldtrip-implemented runica method (Makeig, Bell, Jung, & Sejnowski, 1996), which performs ICA decomposition using the logistic infomax algorithm (Bell & Sejnowski, 1995) with principal component dimension reduction. Noisy electrodes were interpolated (median number of interpolated electrodes = 6 out of 128, sIQR = 5.5), and individual trials were removed if the amplitude range exceeded 200 µV. At this stage, data for 7 participants were discarded because of a large number of artifactual trials (>30%). For the final sample (n=24), the median number of rejected trials was 5.5 out of 79 (± 1.5 sIQR).

Following artifact rejection, full stimulus epochs were analyzed in the frequency domain. First, EEG time courses were averaged across trials to calculate the evoked (i.e., phase-locked) power spectrum (Ding & Simon, 2014). Then, average time-domain signals from each sensor were baseline corrected by subtracting the average in the time window ranging from –1 to 0 s with respect to stimulus onset, zero-padded and multiplied with a Hann window before transformation to the frequency domain using a FFT. The resulting frequency resolution was 0.0025 Hz. The resultant power spectra were normalized by subtracting from each frequency bin the median power of the 16 neighboring frequency bins (8 on each side, i.e., the median over 0.0175 Hz on either side of the center frequency bin); this normalization was performed separately at each electrode (Nozaradan et al., 2012a). Consistent with previous studies (Cirelli, et al., 2016; Gilmore & Russo, 2021; Nave et al., 2022; Nozaradan et al., 2011, 2012a), spectral power values were averaged over all sensors.

Separately for each frequency of interest (1.1 Hz, 2.2 Hz, and 4.4 Hz), we computed linear mixed-effects models predicting normalized spectral power, averaged over all sensors, from fixed factors Rhythm Type, Attention, and Rhythm Type x Attention and random factor Participant ID; we included both a random intercept for each participant as well as a random slope for each fixed factor. When the Rhythm Type × Attention interaction was significant, we followed up with separate linear mixed-effects models for the two Attention conditions to test for simple effects of Rhythm Type.

Control analyses: One potential shortcoming of the current design was that participants performed a visual task during distracted blocks but did not perform any task and were moreover free to close their eyes during the attended blocks. Both performance of a visual task as well as closure of the eyes affect brain activity in the alpha frequency band (~7–13 Hz) over parietal and occipital electrodes (Mulholland, 1972). Thus, it was important to rule out that any "enhancements" of spectral power by attention were not unintentionally driven by systematic

task- or eye-closure-related changes in alpha power. As a first step, we tested whether eye-closure behavior indeed differed between the attended and distracted blocks. To do so, we estimated the number of trials during which each participant had their eyes open versus closed within each condition. We did this by inspecting the data from the ICA. Based on visual inspection of components corresponding to blinks and eye movements, we determined whether each trial contained clear blinks or not. Although this technique is not perfect, blink identification was done blind to condition, and non-blink trials were tallied for each condition only after the fact. Note that the absence of a blink does not necessarily mean that the participant's eyes were closed during that trial, but the presence of blinks does mean that the eyes were open. Thus, this is a conservative approach. We then tested whether the proportion of trials without blinks differed across conditions.

Given substantial differences in eye-closure behavior between attended and distracted conditions (see Results), we also included alpha power in our statistical models testing beat-related spectral power. We computed frequency spectra separately for each participant, trial, and electrode as described above, and then averaged over trials in the frequency domain. Alpha power was averaged between 7 and 13 Hz and over a parietal/occipital electrode cluster (E53, E54, E59, E60, E61, E62, E65, E66, E67, E70, E71, E72, E75, E76, E77, E78, E79, E83, E84, E85, E86, E90, E91). We then added alpha power to the linear mixed-effects models predicting power at beat-related frequencies. In the case that a mixed-effects model for beat-related spectral power indicated a significant effect of or interaction with Attention, we computed a second mixed-effects model with alpha power as an additional fixed factor. We checked for the significance of the alpha power predictor, and compared the goodness-of-fit of the two models using a likelihood-ratio test, which penalizes for the additional parameter added to the model containing alpha power as a predictor.

Results

Tapping was less variable and more accurate for beat rhythms than non-beat rhythms

Paired-sample t-tests were conducted for mean ITI, CoV, and CoA to test for differences in tapping performance for beat and non-beat rhythms. Mean ITI was not significantly different between strong-beat ($M = 0.83$, $SD = 0.34$) and weak-beat ($M = 0.83$, $SD = 0.33$) rhythms ($t(33) = 0.35$, $p = .73$, two-tailed), indicating that participants tapped at similar frequencies for both conditions. The two most prominent tapping frequencies for both rhythm types were 1.1 Hz and 2.2 Hz (Fig. 2a), which is unsurprising as both rhythm types were constructed to have energy primarily at those two frequencies (as well as 4.4 Hz, though [as expected] participants did not tap at this fast frequency, as it falls outside the typical beat range; Repp, 2005a). The agreement between the temporal structure of the rhythms (Fig. 1b) and the frequencies most commonly tapped to the rhythms (Fig. 2a) informed our decision to treat these frequencies as beat-related in the analysis of EEG data.

CoV was significantly lower for tapping to strong-beat rhythms ($M = 0.13$, $SD = 0.09$) than weak-beat rhythms ($M = 0.14$, $SD = 0.09$; $t(33) = -2.38$, $p = 0.02$, two-tailed), indicating that tapping to strong-beat rhythms was significantly less variable than tapping to weak-beat rhythms (see Fig. 2b). Proportional asynchronies were also smaller in the strong-beat than weak-beat condition (see Fig. 3b) (CoA: $M_{strong} = 0.08$, $SD_{strong} = 0.04$; $M_{weak} = 0.10$, $SD_{weak} = 0.05$; $t(33) = -3.26$, $p = 0.003$, two-tailed). This indicates that participants were significantly less accurate at aligning their taps with the unjittered first and third tones in weak-beat rhythms than in strong-beat rhythms, even though those events happened at exactly the same time points in the two rhythm types.

RSVP task was equally distracting for beat and non-beat rhythms

Overall, participants performed better than chance when determining whether RSVP trials contained an even or odd number of target letters, $t(23) = 33.49$, $p <.001$. Importantly, accuracy for the strong-beat trials did not differ from accuracy on the weak-beat trials, $t(23) = 0.67$, $p = .51$. Thus, participants were actively attending to the RSVP task during the "distracted" trials, but they were distracted from the auditory stimuli to a similar degree during both strong-beat and weak-beat trials.

A. Histograms of tapping frequencies   **B. Tapping performance**
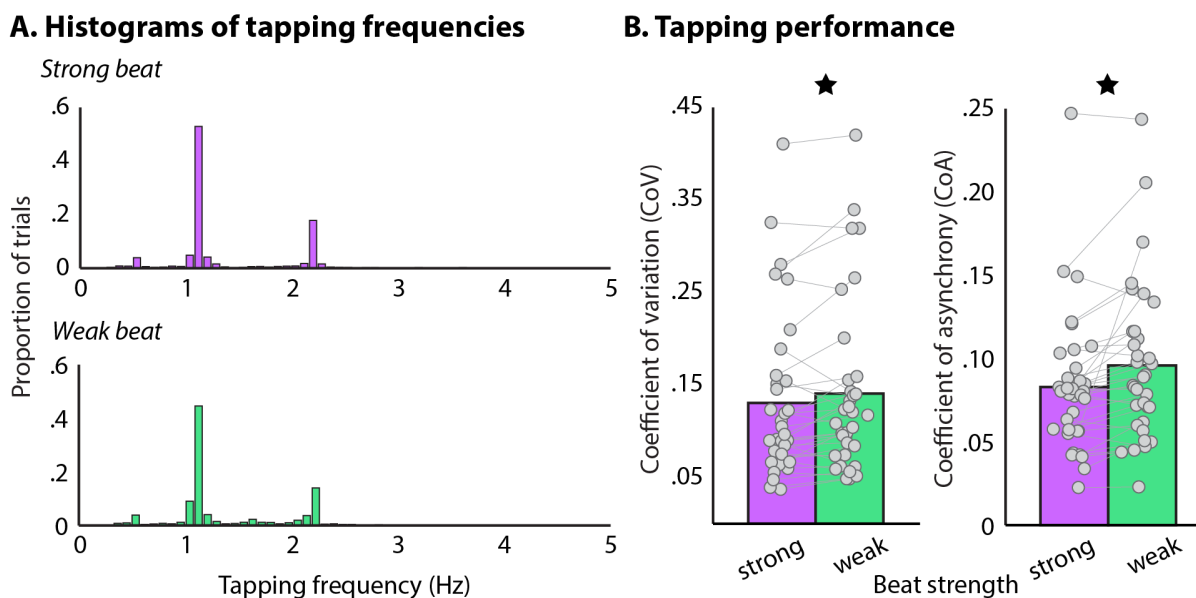
Figure 2: A. Histograms showing proportion of tapping trials in which particular frequencies were tapped for strong-beat (top, purple) and weak-beat (bottom, green) rhythms. The most commonly tapped frequencies were 1.1 Hz and 2.2 Hz, the intended "beat frequencies". **B.** Comparisons of tapping variability (coefficient of variation, CoV) and tapping accuracy (coefficient of asynchrony, CoA) for strong-beat (purple) and weak-beat (green) trials; single-participant data are shown as individual gray circles, and bar graphs show the mean across participants. Stars indicate statistically significant differences between strong-beat and weak-beat rhythms.

Attention enhanced EEG spectral power in response to strong-beat, but not weak-beat, rhythms

The mixed-effects model predicting spectral power at 1.1 Hz revealed nonsignificant main effects and interaction, (Rhythm Type: t(92) = –0.62, p = .53; Attention: t(92) = –0.21, p = .83; interaction: t(92) = 0.05, p = .96). The mixed-effects model predicting spectral power at 2.2 Hz revealed a main effect of Attention, t(92) = –3.09, p = .003, and a Rhythm Type × Attention interaction, t(92) = 2.47, p = .02. For strong-beat rhythms, attention significantly enhanced 2.2

Hz power, t(46) = –2.29, p = .03, whereas for weak-beat rhythms, 2.2-Hz power did not depend on attention, t(46) = 1.21, p = .23. For 4.4-Hz power, the main effects of <u>R</u>hythm <u>T</u>ype (t(92) = –5.62, p = 2.015e-07), and <u>A</u>ttention (t(92) = –2.69, p = .009), as well as their interaction (t(92) = 3.14, p = .002), were significant. For strong-beat rhythms, attention significantly enhanced 4.4-Hz power, t(46) = –2.06, p = .04, whereas for weak-beat rhythms, 4.4-Hz power did not depend on attention, t(46) = 1.28, p = .21.

To summarize, for the beat-related frequencies 2.2 Hz and 4.4 Hz, attention enhanced auditory steady-state responses compared to when attention was diverted to a concurrent task. Critically, this was *only* true for strong-beat rhythms; attention did not significantly influence the magnitude of the steady-state response to weak-beat rhythms at any beat-related frequency.

A. Steady-state evoked responses to strong-beat and weak-beat rhythms

Strong-beat rhythms
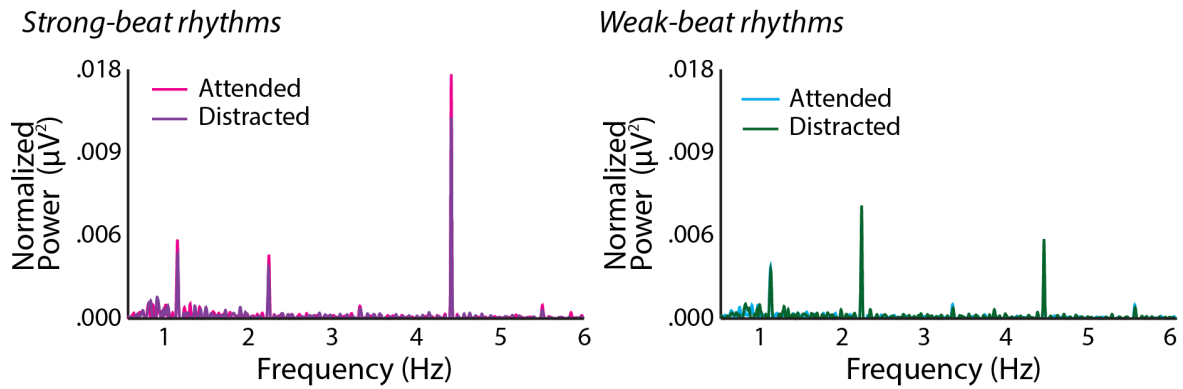
Weak-beat rhythms

B. Attention influences beat-rate steady-state evoked potentials
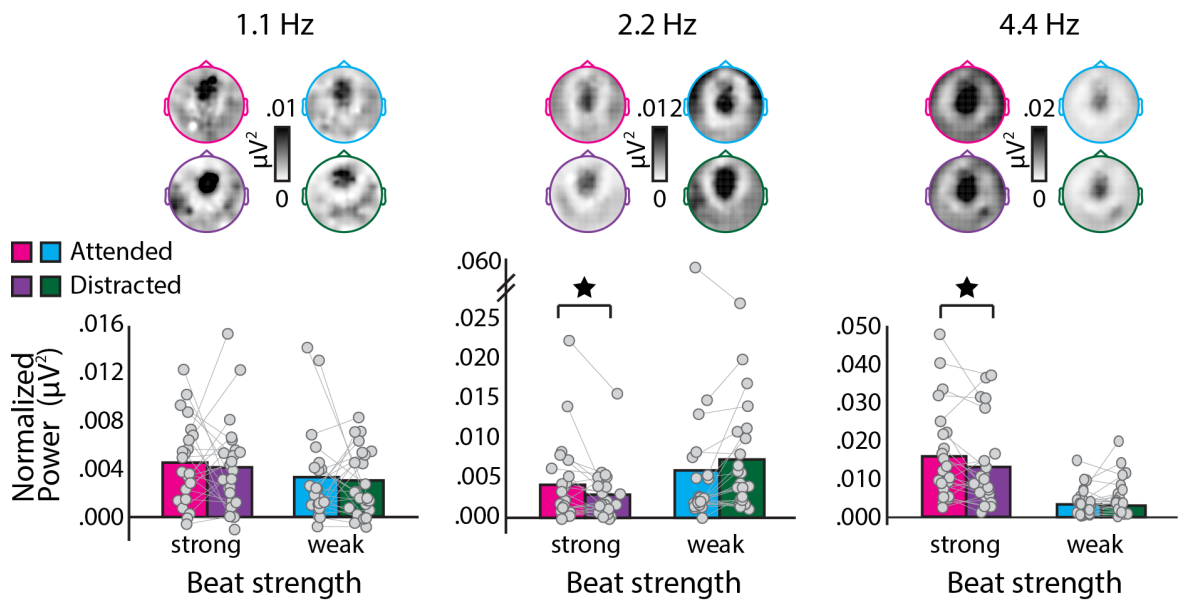
Figure 3: A. Steady-state evoked potentials to beat and non-beat rhythms. Normalized EEG power spectra to strong-beat *(left)* and weak-beat *(right)* rhythms for attended (*magenta*, *cyan*) and distracted *(purple, green)* trials. **B. Attention influences beat-rate steady-state evoked potentials.** Bar plots show comparisons of power at beat (1.1 Hz, 2.2 Hz) and stimulus presentation (4.4 Hz) frequencies to strong-beat *(magenta, purple)* and weak-beat *(cyan, green)* rhythms in the attended and distracted conditions stimuli. Topographies show distributions of normalized power at beat-related frequencies (1.1 Hz, 2.2 Hz, 4.4 Hz).

Participants' eye-closure behavior did not affect beat-related power enhancements

Unsurprisingly, we identified very few trials without blinks during performance of the visual task, and the proportions were very similar across strong-beat and weak-beat conditions (strong beat: mean = .14, median = 0; weak beat: mean = .13, median = 0; t(23) = 1.73, p = .10). Most participants seemed to close their eyes while listening and attending to the auditory rhythms,

but again the proportions of trials without blinks were very similar across strong-beat and weak-beat conditions (strong beat: mean = .78 median = .95; weak beat: mean = .72 median = .98; $t(23) = 0.25$, $p = 0.81$). Thus, eyes were mostly open during performance of the visual task (during the "distracted" blocks) and mostly closed during attention to the auditory rhythms, but there were critically no systematic differences between strong-beat and weak-beat rhythm blocks (Supporting Figure S1).

To ensure that the effects of attention we observed for beat-related frequencies could not have been inadvertently driven by changes in alpha-power due to eye-closure during attended blocks, we repeated the three mixed-effects models we described above (one per frequency of interest), including alpha power as a predictor of beat-related spectral power. In no case did alpha power predict beat-related power, and in no case did the inclusion of alpha power in the model either a) change the statistical results for beat-related power, or b) improve the model performance overall (Supporting Information). We would here take the opportunity to emphasize that attention increased spectral power *only* for strong-beat, but not for weak-beat rhythms. The two rhythm types did not differ in terms of eye-closure behavior, so we interpret this effect as being related to beat perception per se. To summarize, increased 2.2- and 4.4-Hz spectral power during attention compared to distraction was not attributable to differences in alpha power that arise in the EEG spectrum because participants needed to keep their eyes open during the visual task (under distraction) but not during the attended blocks.

Different patterns of tapping behavior were not related to neural activity
We have interpreted attention effects on auditory steady-state responses at 2.2 Hz and 4.4 Hz as reflecting negative effects of distraction on beat perception. However, it is notable that, when participants were asked to tap the beat of the rhythms, most of them tapped at 1.1 Hz, and activity at this frequency was *not* modulated by attention. Although we suspected the analysis

would be underpowered, we nonetheless attempted to a) analyze auditory steady-state responses for only those participants that tapped the beat at 1.1 Hz, and b) compare auditory steady-state responses for participants that tapped the beat consistently at 1.1 Hz to those that did not.

The participants that took part in the tapping experiment were not all the same participants that took part in the EEG experiment (see section: *Participants*). Nonetheless, we were able to identify 11 EEG participants who also participated in the tapping experiment, and who tapped at 1.1 Hz on greater than 70% of all trials (see Fig. 4a). Using the same criterion, we were only able to identify 1 participant that tapped at 2.2 Hz on >70% of trials. For that reason, we focused only on the 1.1-Hz tappers, and repeated our main statistical analysis on their data (linear mixed-effects models including alpha power as a predictor). Similar to the results with the full sample, the Rhythm Type x Attention interaction did not reach significance ($t(39) = 1.20$, $p = .24$).

We also compared the 11 participants that tapped consistently at 1.1 Hz ("1.1-Hz tappers") to everyone else ("non-1.1-Hz tappers", n=13; Fig. 4a). For this analysis, we were simply interested in whether individuals who tapped consistently at 1.1 Hz tended to have a higher ratio of 1.1 to 2.2 Hz power compared to individuals that did not tap at 1.1 Hz consistently. For each individual, we calculated a difference-over-sum metric for 1.1-Hz vs. 2.2-Hz power ([1.1-Hz power – 2.2-Hz power] / [1.1-Hz power + 2.2-Hz power]). Negative values mean that 1.1-Hz power was stronger, positive values mean that 2.2-Hz power was stronger. Figure 4b shows those values for each condition separated by into 1.1-Hz tappers and non-1.1-Hz tappers. There were no statistical differences in the ratio between 1.1- and 2.2-Hz power between participants that tapped consistently at 1.1 Hz and those that did not (linear mixed-effects model, main effect of tapping group: $t(94) = 1.34$, $p = .18$). Thus, in our hands, there was no discernable relationship between tapping behavior and the neural data. We note

that the 1.1-Hz tappers had numerically higher 1.1-Hz power in 3 of 4 experimental conditions (and in particular for strong-beat rhythms), but our current design was admittedly underpowered to fully pursue questions about individual differences.
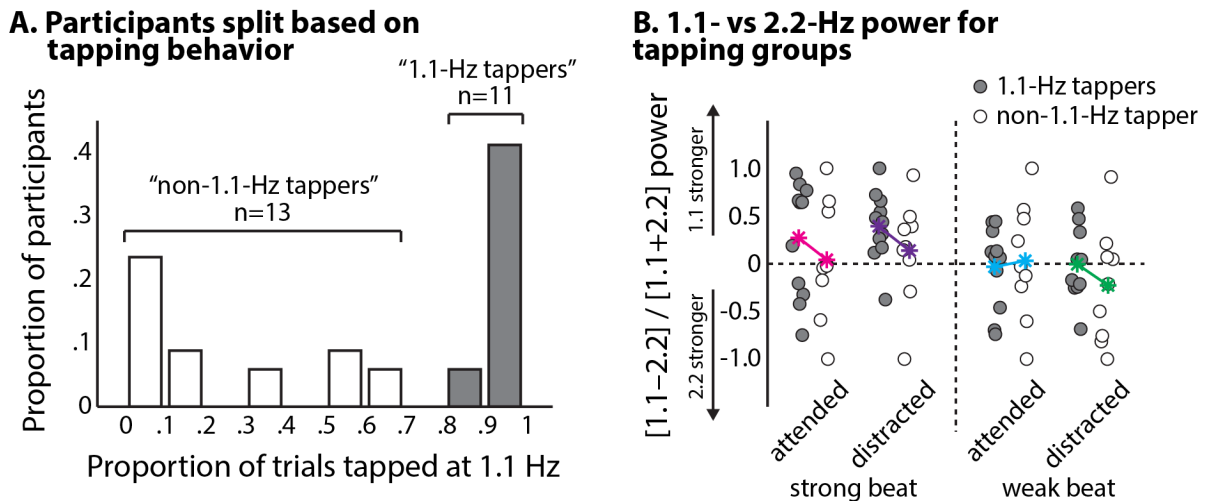
Figure 4. A) Histogram showing the proportion of participants that tapped different proportions of trials at 1.1 Hz. A subset of n=11 participants tapped >70% of trials at 1.1 Hz, and so were deemed "1.1-Hz tappers". The rest (n=13) were categorized as "non-1.1-Hz tappers". **B)** Ratio of 1.1- to 2.2-Hz power plotted as difference-over-sum; +1 means 1.1-Hz power was stronger, and –1 means 2.2-Hz power was stronger. Gray dots show data for 1.1-Hz tappers and white dots for 2.2-Hz tappers. Colored asterisks show the mean for each column of data.

Discussion

The current study examined frequency-domain EEG responses to non-repeating strong- and weak-beat rhythms while attention was either directed toward the rhythms or diverted to a concurrent visual task. We examined spectral power at beat-related frequencies that were confirmed to be relevant by an offline sensorimotor synchronization task. Attention to the rhythms enhanced the steady-state evoked response at beat-relevant frequencies, but, critically, only for rhythms that elicited a strong sense of beat.

Sensorimotor Synchronization

The auditory stimuli were constructed to possess periodicity at 1.1 Hz, 2.2 Hz, and 4.4 Hz. We predicted that participants would tap the beat at either 1.1 Hz or 2.2 Hz, but not 4.4 Hz because 4.4 Hz is faster than people generally perceive or tap the beat (Parncutt, 1994; Repp, 2005a). In line with these predictions, participants tapped most often at rates of 1.1 Hz and 2.2 Hz (Fig. 2a) during the sensorimotor synchronization task, validating the choice of frequencies to examine in the neural data. We created weak-beat control rhythms by altering the timing of some, but not all, of the tones making up strong-beat rhythms. Specifically, we did not alter the timing of tones in the first and third positions of each four-position bar but jittered the timing of the intervening tones in the second and fourth positions. Our goal was to disrupt beat perception while leaving intact temporal regularity at the 1.1 Hz and 2.2 Hz beat rates. In line with this goal, we found that participants tapped at the same rates for the two rhythm types but tapped more variably and less accurately to weak-beat than to strong-beat rhythms. This is consistent with previous work showing that perception of an underlying periodicity is disrupted by inserting tones at random intervals between tones that are regularly spaced (Demany & Semal, 2002), and confirms that the perceived beat was indeed stronger for strong-beat than weak-beat rhythms, as intended (Grahn & Brett, 2007; Henry et al., 2017; Repp, 2005b; Repp & Su, 2013).

Attention affects steady-state potentials evoked by strong-beat rhythms, but not weak-beat rhythms

The key experimental question was whether the effect of attention on steady-state evoked potentials was generalized, and thus similar for strong- and weak-beat rhythms, or if diverting attention from auditory stimuli reduced the strength of the perceived beat and thus specifically affected responses to strong-beat rhythms. The results are consistent with the latter hypothesis.

Spectral power in the EEG was enhanced at 2.2 Hz and 4.4 Hz when participants attended to the strong-beat rhythms relative to when they were distracted by a visual task. The power of steady-state responses to weak-beat rhythms did not depend on attention. Thus, distraction specifically reduced the strength of the neural response to strong-beat rhythms.

One potential shortcoming of the current design was that participants performed a visual task during distracted blocks but did not perform any task and were moreover free to close their eyes during the attended blocks (Gilmore & Russo, 2021; Nozaradan et al., 2012). Critically, we observed attention-related effects on SSEPs for strong-beat rhythms *only* (as indexed by statistical interactions between Attention and Rhythm Type at 2.2 Hz and 4.4 Hz), and we argue that confounding effects related to the absence of a task in the attended blocks should have been equally present for strong- and weak-beat rhythms. Nonetheless, both performance of a visual task as well as closure of the eyes affect brain activity in the alpha frequency band (~7–13 Hz) over parietal and occipital electrodes (Mulholland, 1972). Thus, we made efforts to rule out that these EEG-spectral differences may have influenced power at beat-related frequencies. First, our analysis of spectral power at beat-related frequencies involved a normalization procedure, where we subtracted the median power across 16 neighboring frequency bins (8 smaller, 8 larger) from the power at the frequency of interest (Nozaradan et al., 2012). The goal of this normalization was to eliminate as much as possible any global effects (from e.g., eye closure or task performance) on the frequency spectrum as a whole that may have influenced power at beat-related frequencies (Donoghue et al. 2020). Second, we determined that there was no difference in how often participants closed their eyes during the strong- vs. weak-beat conditions. Finally, we included alpha activity in our statistical models testing the effects of attention on beat-related spectral power. Again, we did not find any evidence that modulations of alpha power related to task performance or eye closure influenced our results. Thus, we feel confident that increased 2.2- and 4.4-Hz spectral power

in response to strong-beat rhythms during attention compared to distraction was not attributable to differences in alpha power that arise in the EEG spectrum because participants needed to keep their eyes open during the visual task (under distraction) but not during the attended blocks.

A potentially puzzling finding (or lack thereof) was that we observed effects of attention on SSEPs to strong-beat rhythms at 2.2 Hz and 4.4 Hz, but not 1.1 Hz. Although all are considered "beat-related" frequencies, most participants actually tapped at 1.1 Hz when asked to tap along with the beat of the stimuli (Fig. 2a). Thus, we may have expected that 1.1 Hz, as the most common behaviorally selected beat frequency, should have been most sensitive to show the neural effects of attention on beat perception. To try to better understand the null effects at 1.1 Hz, we attempted to test for effects of attention on steady-state evoked potentials only for those participants that tapped consistently at 1.1 Hz. However, we still did not observe attention effects on 1.1-Hz power, though we note that this analysis was underpowered from the start. It is notable that several previous studies found something similar to what we observed here, that condition effects that would theoretically be related to beat perception were stronger at twice (the "harmonic" of) the beat rate (Harding et al., 2019; Kaneshiro et al., (2020); Lenc et al., 2018; Tierney & Kraus, 2014). Although there may indeed be a fundamental reason why attention did not modulate spectral energy at 1.1 Hz (the rate at which most participants tended to tap) in the current study, or why previous studies also observed strongest effects at the first harmonic of the beat, the reason may also be more trivial. For example, the relative influence of 1/f noise is greater at lower frequencies, and so will contaminate our estimates of 1.1-Hz power more than estimates of 2.2-Hz power or 4.4-Hz power. This is a fundamental hurdle for studies examining steady-state evoked potentials at rates corresponding to the perceived beat of musical stimuli, as the range of rates in which beat perception is possible falls squarely into a frequency range where 1/f noise is strongest in the EEG spectrum (Ding & Simon, 2009).

Overall, these findings can potentially be interpreted in one of two ways. First, it could be the case that attention modulates beat perception per se; that is, directing attention away from the auditory rhythms reduces perceived beat strength, which manifests as a reduction in steady-state power at beat-related frequencies. In this case, steady-state power would indeed be an index of perceived beat strength, and our results would support the position of previous studies advocating for steady-state evoked potentials as a direct measure of beat perception (Gilmore & Russo, 2017; Lenc et al., 2018; Nozaradan et al., 2011, 2012b, Nave et al., 2022). Alternatively, it could instead be that beat perception persists whether or not an individual is attending to the auditory rhythm. This possibility is supported by MMN studies demonstrating differences in neural responses to deviants occurring on vs. off the beat, even when listeners are instructed to ignore the auditory rhythms and attend to, for example, a silent subtitled movie (Bouwer et al., 2014; Bouwer et al., 2016; Ladinig et al., 2009). It is possible, however that, despite listeners' experiencing similarly strong beat percepts with and without attention, inattention itself reduces steady-state power. This means that attention would modulate the correspondence between what is perceived and how the brain responds to a stimulus. Although we consider this unlikely because of the specificity of the effect to beat-related frequencies only in the context of listening to strong-beat rhythms, we recognize that the design of the current study does not allow us to distinguish between these two possibilities. Moreover, given failures to relate steady-state power to perception on an individual-by-individual or trial-by-trial basis (see section: *Steady-state evoked responses did not relate to tapping behavior on an individual-by-individual basis*) despite attempts to do so, it is critical to better understand the correspondence between steady-state power and perception, and how this relationship might be modulated by factors such as attention or enculturation. Regardless of which interpretation is more accurate, we nonetheless consider it important to proceed cautiously with paradigms measuring steady-state evoked potentials as an index of beat perception when attention cannot

be ensured, such as with infants (Flaten et al., 2022) or non-human animals (Honing et al., 2018). For now, our approach would be the following: if steady-state evoked potentials suggest the *presence* of beat perception in infants or non-human animals, the results can likely be interpreted as such. However, the *absence* of beat perception abilities based on steady-state results cannot be conclusive, as a relative weakness of steady-state power could mean that an organism did not or cannot perceive a beat, but could also indicate that the organism was simply not paying attention to the stimulus rhythm (Wilson & Cook, 2016).

One potential way out of this conundrum would be to look to previous literature for another technique that might complement or validate steady-state evoked potentials as an index of beat perception, and thereby allow us to tease apart contributions of beat perception and attention to steady-state evoked potentials. Previous work has focused on the MMN, an EEG event-related potential that occurs in response to violations of expectations (Bouwer & Honing, 2015). When a beat is perceived, violations of expectations elicit a stronger MMN when an unexpected deviant stimulus (or unexpected omission of a stimulus event) occurs on the beat compared to off the beat. Critically, this on- vs off-beat difference in MMN responses has been shown to be of similar magnitude when listeners are paying attention to the auditory rhythm as when they are distracted by a silent subtitled movie (Bouwer et al., 2014; Bouwer et al., 2016; Ladinig et al., 2009). These studies indicate that beat perception is indeed "pre-attentive", and would here suggest that attention influences SSEP power, but not beat perception per se. However, all studies to our knowledge that have reported this finding have used rhythm stimuli where the beat is communicated by a rich acoustic accent structure: either layered drum patterns consisting of, for example, kick, snare, and high-hat samples, or rhythms with alternating loud and quiet events, so-called intensity accents, that make the beat easier to perceive. In the current study, rhythms did not possess any acoustic accents, but beat perception arose based on the temporal structure of the rhythm (Bouwer et al., 2018), and required active construction on the

part of the listener (Bouwer & Honing, 2015; Grahn, 2012; Grahn & Rowe, 2009, 2013; Lenc et al., 2020). Listeners, especially those without musical training, find it more difficult to perceive a beat in the types of rhythms we used here compared to those with acoustic accents (Grahn & Rowe, 2009, 2013; Bouwer et al., 2018), and it is unclear the extent to which attention is required when a listener needs to actively construct a beat from temporal information alone. Moreover, one recent study examined steady-state evoked potentials and MMN responses in parallel in a paradigm where infants were exposed to audiovisual rhythms with ambiguous metrical structure, but were primed to interpret the rhythm as having either a duple or triple metrical structure (Flaten et al., 2022). Although MMN results suggested that infants did perceive the primed beat as intended, SSEP results did not show this, meaning that the neural measures dissociated even though care was taken to ensure that infants paid attention to the rhythm stimuli. Currently, the link between MMN and steady-state evoked potential measures of beat perception specifically for rhythms for which listeners must actively construct a beat is unclear, and more work is required to understand how the measures might complement each or validate each other so that we can move forward theoretically.

Importantly, the current findings do not invalidate work using (higher-frequency) steady-state evoked potentials as an objective audiometric assessment for infants (e.g., Cone-Wesson et al., 2002). Moreover, these results can be interpreted in the context of those in the more classical frequency-tagging literature, in which steady-state evoked potentials have similarly been used to examine intermodal attention (Keitel et al., 2011; Saupe et al., 2009). Steady-state evoked potentials to concurrently presented auditory and visual stimuli are enhanced for the attended relative to the unattended modality, which increases salience of both task-relevant and irrelevant-characteristics of a stimulus (Müller et al., 1998). However, in classical frequency tagging work, enhancements of task-irrelevant stimulus features provide evidence that attention acts as a global gain mechanism (e.g., enhancement of 40 Hz power in

the EEG signal when a 40-Hz amplitude modulation is applied to a speech stimulus, as in Keitel et al., 2011). In contrast, in the current study, attention enhanced responses at beat-related frequencies *only* for strong-beat rhythms, even though the same frequencies (i.e., 2.2 Hz and 4.4 Hz) were present in the stimulus for weak-beat rhythms. The lack of attentional enhancement in the weak-beat condition suggests that the increases in the strong-beat condition are not simply the result of a global gain, leading to an enhanced overall response to the attended stimulus (Keitel et al., 2011). Rather, the attention-related increases in the beat condition are related to the enhancement of beat perception, resulting in enhanced neural power at beat-related frequencies.

This divergence between attentional effects in the current study and in classical frequency-tagging work may potentially be explained by differences between the two approaches that are not generally discussed. Classical frequency tagging involves modulating a stimulus at a rate that is irrelevant to the question of interest and assessing changes in the neural response to that frequency as an index of another process, such as attention. When using steady-state evoked potentials to investigate beat perception, however, power at the beat-related frequencies is used directly to infer something about the perception of those frequencies themselves. This difference suggests that interpretations about beat perception based on comparisons to classical frequency-tagging work should be made cautiously.

Non-repeating rhythms elicit steady-state evoked potentials at beat frequencies

One goal of the current study was to provide a proof-of-principle that steady-state evoked potentials related to beat perception would be observed with non-repeating rhythms. To date, most applications of this technique have either been applied to highly repetitive stimuli where the same few bars repeat over and over again (Cirelli et al., 2016; Nozaradan et al., 2016, 2012), or to music (Doelling & Poeppel, 2015; Tierney & Kraus, 2014). We consider the current study

an important step, as robust beat perception occurs in response to non-repeating rhythms, but little is known about the relationship between steady-state evoked potentials and beat perception with rhythms that do not repeat. One notable exception is a study where non-repeating rhythms were employed specifically to gradually change beat strength over the course of the rhythm (Lenc et al., 2020). However, here we were interested in assessing attentional effects on non-repeating strong- and weak-beat rhythms whose perceived beat strength was intended to stay constant over the course of the stimulus. A benefit to using rhythms like those in the current study is that they can be more tightly controlled in terms of acoustic features than music can, but still enable decoupling of the perception of a beat from perception of rhythm.

Using non-repeating rhythms raises an interesting analytical issue, namely, how to define beat-related frequencies in the stimulus or EEG spectrum. When using repeating stimuli, it is often easy to categorize beat-related frequencies by visually inspecting the frequency spectrum of the stimulus envelopes (Cirelli et al., 2016; Lenc et al., 2018; Nozaradan et al., 2012b, 2016). Indeed, many previous studies have compared peaks in neural power to peaks in stimulus power. However, here we chose to use an alternative method of identifying beat-related frequency because non-repeating rhythms have peaks that are generally less prominent than repeating rhythms. To illustrate this, we plot in Figure 5 a frequency spectrum for a repeating rhythm (Fig. 5a) and for a non-repeating rhythm (Fig. 5b). The non-repeating rhythm was one of the 36-bar experimental strong-beat stimuli from the current study, and the repeating rhythm was made by repeating 4 bars of that same rhythm 9 times. Thus, the beat rate (i.e., the timing of the beat within the rhythms) is the same both sequences. It is obvious that repeating the rhythm gives rise to a cleaner spectrum (Smith, 1999), but more importantly, what would be defined as a beat-related frequency in the two spectra (based on e.g., a Z-scoring approach; Nozaradan et al., 2012b) would be quite different for the two stimuli. That is, the peaks in the

spectrum for the repeating rhythm at the expected beat-related frequencies (1.1, 2.2, 4.4 Hz; pink arrows) are clear, but, with the exception of 4.4 Hz, barely stand out as peaks against the rest of the power spectrum. Thus, predicting differences in percept based on differences in stimulus power is not straightforward. Moreover, approaches that compare stimulus peaks to EEG peaks, or compare beat-related and non-beat-related peaks, may potentially draw very different conclusions about the strength of the perceived beat in repeating rhythms compared to non-repeating rhythms, based only on the stimulus spectrum (Henry et al., 2017). Here, we circumvented the issue of equating stimulus power to beat strength by 1) informing our choice of beat-related frequencies with behavioral tapping data (Lenc et al., 2018), and 2) comparing neural responses to the same rhythm (therefore identical spectra) under two different conditions (attended vs. distracted).
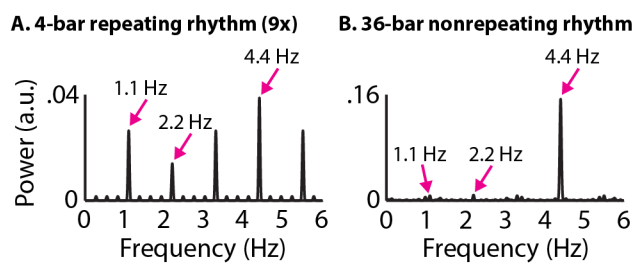

Figure 5: Power spectra for the envelopes of a (A) repeating rhythm (4 bars repeated 9 times, as is commonly used in other studies) and a (B) non-repeating 36-bar rhythm non-repeating rhythm (used in the current study). The repeating stimulus comprised nine repetitions of a four-bar segment taken from the non-repeating rhythm. Pink arrows mark beat-related frequencies 1.1 Hz, 2.2 Hz, and 4.4 Hz.

Steady-state evoked responses did not relate to tapping behavior on an individual-by-individual basis

Although the current experiment was not necessarily designed to relate individual differences in steady-state evoked potentials to participants' individual perception (tapping behavior), we nonetheless tested whether there were neural differences between individuals that tapped consistently at one potential beat rate (1.1 Hz) and those that tapped at another rate (2.2 Hz) or switched their tapping rate from trial to trial. For example, one might predict that individuals that tapped at 1.1 Hz would have more power at 1.1 Hz than 2.2 Hz, and vice versa. However,

there were no significant relationships between tapping behavior and EEG results. One previous study found a significant relationship between neural measures of beat perception (a difference score comparing steady-state responses to strong-beat vs. weak-beat rhythms) and a model-based "prediction index" (Nozaradan et al., 2016). Moreover, several studies have shown group differences in steady-state power between musicians and nonmusicians (Cirelli et al., 2016, Doelling & Poeppel, 2015). However, we are not aware of work linking the magnitude of beat-related steady-state evoked potentials directly to either the strength of a perceived beat, assessed behaviorally, or the tendency to perceive a beat at a specific metrical level. If steady-state evoked potentials are to be considered as a proxy for beat perception, we consider it an important endeavor for future work to convincingly relate brain to behavior using designs that are specifically tuned for this purpose.

Conclusions

Steady-state evoked potentials at beat-related frequencies were enhanced when rhythms were attended to, but only if the rhythms gave rise to strong a sense of beat. Steady-state evoked potentials to weak-beat rhythms were unaffected by attention. Enhancement was observed even though the rhythms were non-repeating. Taken together, these findings support the idea that steady-state evoked potentials may relate to beat perception for non-repeating rhythms. However, this relationship may be affected by attention, may not be observed for lower beat-related frequencies (e.g., 1.1 Hz), and does not appear to consistently relate to individual differences in beat-rate percept.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Bouwer, F. L., Burgoyne, J. A., Odijk, D., Honing, H., & Grahn, J. A. (2018). What makes a rhythm complex? The influence of musical training and accent type on beat perception. *PLoS ONE*, *13*(1), 1–26. https://doi.org/10.1371/journal.pone.0190322

Bouwer, F. L., Werner C. M., Knetemann, M., & Honing, H. (2016). Disentangling beat perception from sequential learning and examining the influence of attention and musical abilities on ERP responses to rhythm. *Neuropsychologia, 85*, 80-90. https://doi.org/10.1016/j.neuropsychologia.2016.02.018.

Bouwer, F. L., & Honing, H. (2015). Temporal attending and prediction influence the perception of metrical rhythm: evidence from reaction times and ERPs. *Frontiers in Psychology*, *6*(July), 1–14. https://doi.org/10.3389/fpsyg.2015.01094

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), 1–9. https://doi.org/10.1371/journal.pone.0097467

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Cirelli, L. K., Spinelli, C., Nozaradan, S., & Trainor, L. J. (2016). Measuring Neural Entrainment to Beat and Meter in Infants: Effects of Music Background. *Frontiers in Neuroscience*, *10*(May), 1–11. https://doi.org/10.3389/fnins.2016.00229

Cone-Wesson, B., Rickards, F., Poulis, C., Parker, J., Tan, L., & Pollard, J. (2002). The auditory steady-state response: Clinical observations and applications in infants and children. *Journal of the American Academy of Audiology*, *13*(5), 270–282.

de Fleurian, R., Blackwell, T., Ben-Tal, O., & Müllensiefen, D. (2017). Information-theoretic measures predict the human judgment of rhythm complexity. *Cognitive Science*, *41*(3), 800-813.

Ding, N. & Simon, J. Z., (2009). Neural Representations of Complex Temporal Modulations
in the human auditory cortex. *Journal of Neurophysiology, 102*(5), 2731-2743. doi:
10.1152/jn.00523.2009

Demany, L., & Semal, C. (2002). Limits of rhythm perception. *The Quarterly Journal of
Experimental Psychology*, *55A*(2), 643–657. https://doi.org/10.1080/0272498014300040

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: functional roles
and interpretations. *Frontiers in Human Neuroscience*, *8*(May), 1–7.
https://doi.org/10.3389/fnhum.2014.00311

Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B., & Poeppel, D. (2019). An
oscillator model better predicts cortical entrainment to music. *Proceedings of the
National Academy of Sciences*, 201816414. https://doi.org/10.1073/pnas.1816414116

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by
expertise. *Proceedings of the National Academy of Sciences*, 201508431.
https://doi.org/10.1073/pnas.1508431112

Flaten, E., Marshall, S. A., Dittrich, A., & Trainor, L. J. (2022). Evidence for top-down metre
perception in infancy as shown by primed neural responses to an ambiguous rhythm.
European Journal of Neuroscience, 55(8), 2003-2023. https://doi.org/10.1111/ejn.15671

Gilmore, S. A., & Russo F. A., (2021). Neural and behavioural evidence for vibrotactile beat
perception and bimodal enhancement. Journal of Cognitive Neuroscience, 33(4), 635-
650. doi: 10.1162/jocn_a_01673.

Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms.
*Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation
Cérébrale*, *220*(1), 51–61. https://doi.org/10.1007/s00221-012-3114-8

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain.
Journal of Cognitive Neuroscience, *19*(5), 893–906.

https://doi.org/10.1162/jocn.2007.19.5.893

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(23), 7540–7548. https://doi.org/10.1523/JNEUROSCI.2018-08.2009

Grahn, J. A., & Rowe, J. B. (2013). Finding and feeling the musical beat: Striatal dissociations between detection and prediction of regularity. *Cerebral Cortex*, *23*(4), 913–921. https://doi.org/10.1093/cercor/bhs083

Hagmann, C. E., & Cook, R. G. (2010). Testing meter, rhythm, and tempo discriminations in pigeons. *Behavioural Processes*, *85*(2), 99–110. https://doi.org/10.1016/j.beproc.2010.06.015

Harding E. E., Sammler D, Henry M. J., Large E. W., Kotz S. A., (2019). Cortical tracking of rhythm in music and speech. *Neuroimage*, *185*, 96-101. doi: 10.1016/j.neuroimage.2018.10.037

Henry, M. J., Herrmann, B. B., & Grahn, J. A. (2017). What can we learn about beat perception by comparing brain signals and stimulus envelopes? *PloS One, 12*(2), 1–30. https://doi.org/10.1371/journal.pone.0172454

Honing, H., Bouwer, F.L., Háden, G.P. (2014). Perceiving Temporal Regularity in Music: The Role of Auditory Event-Related Potentials (ERPs) in Probing Beat Perception. In: Merchant, H., de Lafuente, V. (eds) Neurobiology of Interval Timing. Advances in Experimental Medicine and Biology, vol 829. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-1782-2_16

Honing, H., Bouwer, F. L., Prado, L., & Merchant, H. (2018). Rhesus Monkeys (Macaca mulatta) Sense Isochrony in Rhythm, but Not the Beat: Additional Support for the Gradual Audiomotor Evolution Hypothesis. *Frontiers in Neuroscience*, *12*(July), 1–15.

https://doi.org/10.3389/fnins.2018.00475

Honing, H., Ladinig, O., Háden, G. P., & Winkler, I. (2009). Is beat induction innate or learned? Probing emergent meter perception in adults and newborns using event-related brain potentials. *Annals of the New York Academy of Sciences*, *1169*(Illc), 93–96. https://doi.org/10.1111/j.1749-6632.2009.04761.x

Honing, H., Merchant, H., Háden, G. P., Prado, L., & Bartolo, R. (2012). Rhesus Monkeys (Macaca mulatta) Detect Rhythmic Groups in Music, but Not the Beat. *PLoS ONE*, *7*(12), e51369. https://doi.org/10.1371/journal.pone.0051369

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences*, *1169*, 58–73. https://doi.org/10.1111/j.1749-6632.2009.04579.x

Kaneshiro, B., Nguyen, D. T., Norcia, A. M., Dmochowski, J. P., & Berger, J., (2020). Natual music evokes correlated EEG responses reflecting temporal structure and beat. *NeuroImage*, *214*, 116559. doi:10.1016/j.neuroimage.2020.116559

Keitel, C., Schröger, E., Saupe, K., & Müller, M. M. (2011). Sustained selective intermodal attention modulates processing of language-like stimuli. *Experimental Brain Research*, *213*(2–3), 321–327. https://doi.org/10.1007/s00221-011-2667-2

Ladinig, O., Honing, H., Haden, G., & Winkler, I. (2009). Probing attentive and preattentive emergent meter in adult listeners without extensive music training. *Music Perception*, *26*(4), 377–387. https://doi.org/10.1525

Ladinig, O., Honing, H., Háden, G., & Winkler, I. (2011). "Probing attentive and pre-attentive emergent meter in adult listeners without extensive music training": Erratum. *Music Perception, 28*(4), 444. https://doi.org/10.1525/mp.2011.28.4.444

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), 201801421. https://doi.org/10.1073/pnas.1801421115

Lenc, T., Keller, P. E., Varlet, M., & Nozaradan, S. (2020). Neural and Behavioral Evidence for Frequency-Selective Context Effects in Rhythm Processing in Humans, Cerebral Cortex Communications, 1(1), 1-15. doi: 10.1093/texcom/tgaa037

McAuley, J. D. (2010). Tempo and Rhythm. In M. Riess Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (Vol. 36, pp. 165–199). Springer New York. https://doi.org/10.1007/978-1-4419-6114-3_6

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

Merker, B. (2014). Groove or swing as distributed rhythmic consonance: Introducing the groove matrix. *Frontiers in Human Neuroscience*, *8*(June), 1–4. https://doi.org/10.3389/fnhum.2014.00454

Mulholland, T. B. (1972). Occipital alpha revisited. *Psychological Bulletin, 78*(3), 176–182. https://doi.org/10.1037/h0032962

Müller, M. M., Picton, T. W., Valdes-Sosa, P., Riera, J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (1998). Effects of spatial selective attention on the steady-state visual evoked potential in the 20–28 Hz range. *Cognitive Brain Research*, *6*(4), 249–261. https://doi.org/10.1016/S0926-6410(97)00036-0

Nave K. M., Hannon E. E., Snyder J. S., (2022). Steady-state evoked potentials of subjective beat perception in musical rhythms. *Psychophysiology, 59*(2):e13963. doi: 10.1111/psyp.13963.

Nozaradan, S., Peretz, I., & Keller, P. E. (2016). Individual Differences in Rhythmic Cortical Entrainment Correlate with Predictive Behavior in Sensorimotor Synchronization. *Scientific Reports*, *6*(August 2015), 20612. https://doi.org/10.1038/srep20612

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the Neuronal

Entrainment to Beat and Meter. *Journal of Neuroscience*, *31*(28), 10234–10240.

https://doi.org/10.1523/JNEUROSCI.0411-11.2011

Nozaradan, S., Peretz, I., & Mouraux, A. (2012a). Steady-state evoked potentials as an index

of multisensory temporal binding. *NeuroImage*, *60*(1), 21–28.

https://doi.org/10.1016/j.neuroimage.2011.11.065

Nozaradan, S., Peretz, I., & Mouraux, A. (2012b). Selective neuronal entrainment to the beat

and meter embedded in a musical rhythm. *The Journal of Neuroscience : The Official*

Journal of the Society for Neuroscience, *32*(49), 17572–17581.

https://doi.org/10.1523/JNEUROSCI.3203-12.2012

Parncutt, R. (1994). A Perceptual Model of Pulse Salience and Metrical Accent in Musical

Rhythms. *Music Perception: An Interdisciplinary Journal*, *11*(4), 409–464.

Pelli, D. (1997). The VideoToolbox software for visual psychophysics: transforming numbers

into movies*. *Spatial Vision*, *10*(4), 437–442.

Povel, D.-J., & Essens, P. (1985). Perception of Temporal Patterns. *Music Perception*, *2*(4),

411–440.

Prögler, J. a. (1995). Searching for swing: Participatory discrepancies in the jazz rhythm

section. *Ethnomusicology*, *39*(1), 21–54. https://doi.org/10.2307/852199

Regan, D. (1977). Steady-state evoked potentials. *Journal of the Optical Society of America*,

67(11), 1475. https://doi.org/10.1364/JOSA.67.001475

Repp, B. H. (2005a). Rate Limits of On-Beat and Off-Beat Tapping With Simple Auditory

Rhythms. *Music Perception: An Interdisciplinary Journal*, *23*(2), 165–188.

Repp, B. H. (2005b). Sensorimotor synchronization: A review of the tapping literature.

Psychonomic Bulletin & Review, *12*(6), 969–992. https://doi.org/10.3758/BF03206433

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research

(2006-2012). *Psychonomic Bulletin & Review*, *20*(3), 403–452.

https://doi.org/10.3758/s13423-012-0371-2

Rouse, A. A., Cook, P. F., Large, E. W., & Reichmuth, C. (2016). Beat keeping in a sea lion

as coupled oscillation: Implications for comparative understanding of human rhythm.

Frontiers in Neuroscience, *10*(JUN), 1–12. https://doi.org/10.3389/fnins.2016.00257

Saupe, K., Widmann, A., Bendixen, A., Müller, M. M., & Schröger, E. (2009). Effects of

intermodal attention on the auditory steady-state response and the event-related

potential. *Psychophysiology*, *46*(2), 321–327. https://doi.org/10.1111/j.1469-

8986.2008.00765.x

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments

of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18.

https://doi.org/10.1016/j.tins.2008.09.012

Smith, S. W. (1999). *The scientist and engineer's guide to digital signal processing* (2nd ed.).

California Technical Publishing.

Su, Y.-H., & Pöppel, E., (2012). Body movement enhances the extraction of temporal

structures in auditory sequences, *Psychological Research, 76,* 373-382.

https://doi.org/10.1007/s00426-011-0346-3

Tierney, A., & Kraus, N. (2013). Neural responses to sounds presented on and off the beat of

ecologically valid music. *Frontiers in Systems Neuroscience, 7*(May), 1-7. doi:

10.3389/fnsys.2013.00014

Tierney, A., & Kraus, N. (2014). Neural Entrainment to the Rhythmic Structure of Music.

Journal of Cognitive Neuroscience, *X*(Y), 1–9. https://doi.org/10.1162/jocn_a_00704

van Noorden, L., & Moelants, D. (1999). Resonance in the Perception of Musical Pulse.

Journal of New Music Research, *28*(1), 43–66.

Wilson, M., & Cook P. F., (2016). Rhythmic entrainment: why humans want to, fireflies can't

help it, pet birds try, and sea lions have to be bribed. *Psychonomic Bulletin & Review, 23*(6), 1647-1659. doi: 10.3758/s13423-016-1013-x

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(7), 2468–2471. https://doi.org/10.1073/pnas.0809035106

Figure 3. **Stimulus and task schematics. A) Strong-beat and weak-beat rhythms.** Short segments (2 bars) of an example strong- and weak-beat stimulus pair. Top shows two bars of a strong-beat rhythm, bottom shows two bars of a weak-beat rhythm, the latter of which was created by alternately shorting or lengthening the IOI preceding any tones occurring at the second or fourth positions by one of five jitter amounts. Arrows indicate which tones were adjusted and whether IOIs were shortened or lengthened. **B. Beat-rate spectral power**. Frequency spectra for beat (purple) and non-beat (green) rhythms. Spectral power for strong-beat and weak-beat rhythms was similar at 1.1 Hz, higher at 2.2-Hz for weak-beat rhythms, and higher at 4.4-Hz for strong-beat rhythms; note that inset bar graphs each have independent y-axis scales to emphasize condition differences. **C. Task schematic for attended and distracted conditions.** Rhythms were presented in two conditions: attended (left) and distracted (right). In the attended condition, listeners either fixated a cross or closed their eyes while listening attentively to the strong-beat and weak-beat rhythms. In the distracted condition, listeners performed a rapid serial visual presentations (RSVP) task in which they counted how many lower-case x's occurred in each trial while strong-beat and weak-beat rhythms were presented.

Figure 4: A. Histograms showing proportion of tapping trials in which particular frequencies were tapped for strong-beat (top, purple) and weak-beat (bottom, green) rhythms. The most commonly tapped frequencies were 1.1 Hz and 2.2 Hz, the intended "beat frequencies". **B.** Comparisons of tapping variability (coefficient of variation, CoV) and tapping accuracy (coefficient of asynchrony, CoA) for strong-beat (purple) and weak-beat (green) trials; single-participant data are shown as individual gray circles, and bar graphs show the mean across

participants. Stars indicate statistically significant differences between strong-beat and weak-beat rhythms.

Figure 3: A. Steady-state evoked potentials to beat and non-beat rhythms. Normalized EEG power spectra to strong-beat (left) and weak-beat (right) rhythms for attended (magenta, cyan) and distracted (purple, green) trials. **B. Attention influences beat-rate steady-state evoked potentials.** Bar plots show comparisons of power at beat (1.1 Hz, 2.2 Hz) and stimulus presentation (4.4 Hz) frequencies to strong-beat (magenta, purple) and weak-beat (cyan, green) rhythms in the attended and distracted conditions stimuli. Topographies show distributions of normalized power at beat-related frequencies (1.1 Hz, 2.2 Hz, 4.4 Hz).

Figure 4. A) Histogram showing the proportion of participants that tapped different proportions of trials at 1.1 Hz. A subset of n=11 participants tapped >70% of trials at 1.1 Hz, and so were deemed "1.1-Hz tappers". The rest (n=13) were categorized as "non-1.1-Hz tappers". **B)** Ratio of 1.1- to 2.2-Hz power plotted as difference-over-sum; +1 means 1.1-Hz power was stronger, and –1 means 2.2-Hz power was stronger. Gray dots show data for 1.1-Hz tappers and white dots for 2.2-Hz tappers. Colored asterisks show the mean for each column of data.

Figure 5: Power spectra for the envelopes of a (A) repeating rhythm (4 bars repeated 9 times, as is commonly used in other studies) and a (B) non-repeating 36-bar rhythm non-repeating rhythm (used in the current study). The repeating stimulus comprised nine repetitions of a four-bar segment taken from the non-repeating rhythm. Pink arrows mark beat-related frequencies 1.1 Hz, 2.2 Hz, and 4.4 Hz.