



To the beat of your own drum: Cortical regularization of non-integer ratio rhythms toward metrical patterns

Benjamin A. Motz*, Molly A. Erickson, William P. Hetrick

Department of Psychological and Brain Sciences, Indiana University Bloomington, 1101 East 10th Street, Bloomington, IN 47405, United States

ARTICLE INFO

Article history:

Accepted 14 January 2013

Available online 28 February 2013

Keywords:

Rhythm

ERP

Time perception

Omission elicited potential (OEP)

Meter

ABSTRACT

Humans perceive a wide range of temporal patterns, including those rhythms that occur in music, speech, and movement; however, there are constraints on the rhythmic patterns that we can represent. Past research has shown that sequences in which sounds occur regularly at non-metrical locations in a repeating beat period (non-integer ratio subdivisions of the beat, e.g. sounds at 430 ms in a 1000 ms beat) are represented less accurately than sequences with metrical relationships, where events occur at even subdivisions of the beat (integer ratios, e.g. sounds at 500 ms in a 1000 ms beat). Why do non-integer ratio rhythms present cognitive challenges? An emerging theory is that non-integer ratio sequences are represented incorrectly, “regularized” in the direction of the nearest metrical pattern, and the present study sought evidence of such perceptual regularization toward integer ratio relationships. Participants listened to metrical and non-metrical rhythmic auditory sequences during electroencephalogram recording, and sounds were pseudorandomly omitted from the stimulus sequence. Cortical responses to these omissions (omission elicited potentials; OEPs) were used to estimate the timing of expectations for omitted sounds in integer ratio and non-integer ratio locations. OEP amplitude and onset latency measures indicated that expectations for non-integer ratio sequences are distorted toward the nearest metrical location in the rhythmic period. These top-down effects demonstrate metrical regularization in a purely perceptual context, and provide support for dynamical accounts of rhythm perception.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Since at least the 19th-century, it has been observed that when listening to repeating sounds, listeners tend to impose top-down structure onto the rhythmic experience:

In every case of rhythmical perception we group the isolated sensations of sound into a more or less complex system of ideas that are temporally arranged. (Meumann, 1894; translated in Ruckmich, 1913)

What is the structural form of this “complex system of ideas”? More recent research suggests that the mental representation of repeating sounds can be described as a metrical hierarchy (Deutsch, 1986; Fraisse, 1982; Lerdahl & Jackendoff, 1983; Povel, 1981; Povel & Essens, 1985). A metrical hierarchy typically consists of repeating intervals of equal duration, which are further subdivided into equal intervals. This critical notion of nested metrical levels has implications for the kinds of rhythmic structures that humans are able to effectively represent. Specifically, “metrical” rhythms in which sounds occur at equally-spaced subdivisions of

a repeating cycle (e.g. at 0 and 400 ms within a repeating 1200 ms cycle; 33% interval) have strong perceptual advantages compared to rhythms with unevenly-spaced sounds (e.g., at 0 and 516 ms within a repeating 1200 ms cycle), which we will call “non-metrical”. Given that we demonstrate preferential treatment of metrical patterns, the present study was conducted to determine how the brain represents non-metrical rhythms characterized by sounds occurring at time points that are not spaced evenly between beats.

Evidence abounds that metrical patterns are represented more accurately than non-metrical patterns. People are better at remembering (Deutsch, 1986; Palmer & Krumhansl, 1990), reproducing (Povel & Essens, 1985), synchronizing with (Patel, Iverson, Chen, & Repp, 2005), and detecting changes in (Grube & Griffiths, 2009; Jones & Yee, 1997; Large & Jones, 1999) sequences with sound events that occur at equal subdivisions of a repeating time period (also called integer ratio relationships or harmonics; e.g., 1:2 or 1:3), as opposed to unequal subdivisions (non-integer ratio relationships, e.g., 1:2.7). People are also better at making perceptual judgments about individual acoustic events when they are in the context of metrically organized rhythms (Jones, Moynihan, MacKenzie, & Puente, 2002), and metrical stress patterns in speech facilitate higher-order semantic processing (Rothermich, Schmidt-Kassow, & Kotz, 2012).

* Corresponding author. Fax: +1 812 855 4691.

E-mail address: bmotz@indiana.edu (B.A. Motz).

This privileged treatment of metrical relationships has been described in the theoretical framework of dynamical systems (Eck, 2002; Jones & Yee, 1997; Large & Jones, 1999; Treisman, Faulkner, Naish, & Brogan, 1990). Perception of periodically repeating stimuli engages neural oscillators that synchronize with the external pattern, and these oscillators endogenously deploy focused attention when the listener would expect to hear the next upcoming sound. When exposed to sound sequences that have multiple hierarchical levels (e.g., eighth notes occurring between beats in 4/4 time), internal oscillations at corresponding levels become coupled to one another to preserve evenly-spaced interval relationships between the metrical subdivisions of the rhythm (Large & Jones, 1999; Large & Kolen, 1995). In other words, top-down imposition of metrical levels on these patterns is the dynamic result of oscillators resonating at integer multiples of the duration between beats (Large & Snyder, 2009).

If expectations for upcoming sounds are deployed by a cortical system that self-organizes to represent evenly-spaced intervals, it is possible that our expectations during perception of non-metrical sequences will conform toward metrical relationships. This type of distortion toward metrical patterns, a phenomenon dubbed “regularization” (Grahn & Brett, 2007), has been well-documented with tasks involving synchronized movement (commonly, finger tapping) with external non-metrical patterns (e.g., Collier & Wright, 1995; Cummins & Port, 1998; Essens, 1986; Povel, 1981; Repp, London, & Keller, 2011). When asked to synchronize with a non-metrical sound sequence, participants’ movements tend to be biased toward the nearest metrical pattern. But in order to synchronize with a sound sequence, the listener must first be able to perceive the sound sequence, and some have suggested that movement biases may be caused by perceptual limitations (e.g., Bingham, 2004; Repp et al., 2011). If this is the case, and our perceptual apparatus is predisposed to represent metrical patterns, we may expect that regularization would occur in purely perceptual tasks, without any overt movement. For example, dynamical theories of beat induction might predict that, even if a sound regularly occurs at 53% the distance between beats (a non-integer ratio location), a listener’s expectations would be shifted in the direction of the nearest integer ratio subdivision of that rhythmic cycle (50%).

Using a discrimination task, Repp et al. (2011) found initial evidence that regularization toward metrical patterns occurs perceptually. Timing perturbations in non-integer ratio auditory rhythms (e.g., with a tone subdividing a continuing beat at 30% of the beat’s period) were more poorly detected when the perturbation was in the direction of a nearby integer ratio (e.g., shifted positively toward 33%), and more easily detected when the tone was shifted away from the integer ratio (e.g., shifted negatively, further from 33%). Given that discrimination performance is inversely proportionate to the similarity of two stimuli, these results indicated that perceived non-integer ratio rhythms are systematically represented more similarly to integer-ratio rhythms. The emerging theory that our perceptual representation is regularized, biased toward the nearest integer ratio subdivision, fits well with decades of experimental evidence of disadvantages with non-metrical patterns in reproduction tasks.

The present study was designed to test the theory that perceptual regularization occurs, by measuring the moment when participants expect a sound to occur in the context of metrical and non-metrical rhythms. In light of the difficulties associated with disentangling expectation of a sound from perception of that sound, an omission elicited potential (OEP; aka emitted potential or omitted stimulus potential) paradigm was used. The OEP is an endogenous electroencephalograph (EEG) component elicited by a perceptual gap when a stimulus had been anticipated. This cortical response includes a positive-going component that has

maximal amplitude at the top of the head, peaks about 350 ms after the expected stimulus would have onset (Besson, Ffita, Czternasty, & Kutas, 1997; Besson, Ffita, & Requin, 1994; Ruchkin & Sutton, 1973; Sutton, Tueting, Zubin, & John, 1967), and may be accompanied by other omission-elicited EEG components and spectral perturbations (e.g. Bendixen, Schröger, & Winkler, 2009; Fujioka, Trainor, Large, & Ross, 2009; Janata, 2001; Snyder & Large, 2005; Zanto, Large, Fuchs, & Kelso, 2005). Unlike most other event-related potentials (ERPs) that are evoked by an external stimulus, these omission responses are observed in the absence of any external signal, and as such, they must be anchored to the timing of internal cognitive predictions for upcoming sounds. Thus, by measuring the onset latency of the OEP to an omitted sound, one can measure approximately when the omitted sound was expected to occur (Hernández & Vogel-Sprott, 2010; Jongsma et al., 2005; Macar & Vidal, 2004).

In the present study, the timings of expectations for upcoming sounds are compared between three simple rhythmic patterns, characterized by the relative timings of stimuli that subdivide periodic beats (integer ratio, 1:3 or 33%; non-integer ratios, 43% and 53%). It was hypothesized that the onset of the OEP to the non-integer ratio rhythms would differ from the OEP onset to the integer ratio rhythm in specific ways. First, because the nearest simple harmonic of a 43% subdivision is 50% (or 1:2), the listener should expect that the sound at 43% will occur later in time than it actually does. As such, the onset of the OEP to an omission at 43% should occur later than the OEP to an omission at an integer ratio (33%). Similarly, the nearest simple harmonic of a 53% subdivision is also 50%; in this case, the listener should expect the onset of the sound to occur earlier than it actually does, and the onset of the OEP should be early with respect to both the integer ratio (33%) rhythm and the non-integer ratio 43% rhythm.

2. Method

2.1. Participants

Thirteen undergraduate students from Indiana University were recruited for their paid participation (5 males; 8 females). Participants ranged in age from 18 to 24 years ($M = 20.54$, $SD = 1.90$). Each participant’s number of years of formal musical training was used to create a median split, segregating participants with little or no musical training ($n = 6$; $M = 1.67$ years training, $SD = 2.07$) from those with more musical expertise ($n = 7$; $M = 10.79$ years training, $SD = 4.34$). All participants reported normal hearing and no history of hospitalization for psychological or neurological disorders.

2.2. Materials and procedure

Participants were seated comfortably in an electrically and acoustically shielded chamber for the duration of the experiment. Auditory stimuli were presented using the DOS-based Stim Program (Neuroscan), and delivered binaurally at 73 dB SPL through ABR Tubeheadphone insert earphones. Three types of acoustic sequences were constructed from white and pink noise bursts (equal amplitude; 50 ms duration with 10 ms linear ramps). In all three conditions, pink noise bursts occurred at regular intervals of 1 s, forming the main periodicity, the beat. In every other interval between pink noise bursts, a white noise burst subdivision was introduced. The three conditions differed in the relative onset time of the white noise burst within the cycle defined by pink noise bursts, either 33%, 43%, or 53% of the distance between the pink noise bursts (330 ms, 430 ms, or 530 ms from the preceding beat, respectively; see Fig. 1). White noise burst subdivisions were presented in every other pink noise interval (rather than each interval) for two

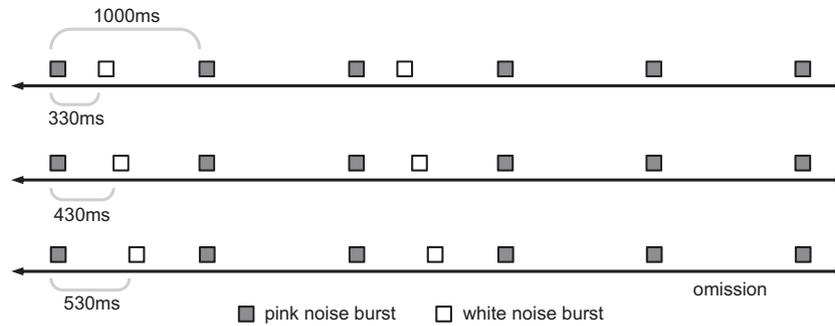


Fig. 1. Stimulus sequences for the three experimental conditions. Conditions differed in the timing of white noise burst subdivisions relative to periodic pink noise burst beats (all sounds are equal amplitude, 50 ms duration). White noise bursts occurred in every other interval. In the metrical condition, shown at the top of the figure, the white noise burst occurs at 33% the distance between beats (1:3). In the following conditions, the white noise subdivision occurs at non-integer ratio locations, 43% and 53% the distance between pink noise bursts. Example stimulus sequences provided as audio files in [Supplementary materials](#).

reasons: (1) so that the more regular pink noise beats would clearly define the main periodicity, rather than the less-frequent white noise burst subdivisions (regularization has been shown to require clear induction of periodicity; [Grahn & Brett, 2007](#)); and (2) so that during offline signal processing, the OEP to a white noise burst's omission could be dissociated from the auditory evoked potentials associated with the nearby pink noise bursts in a non-omission cycle (via subtraction; see *Signal Processing and Analysis*). Stimulus runs were created with 180 pink noise beat intervals, and participants listened to two runs of each rhythm. At pseudorandom intervals, 18–22 of the white noise subdivisions were omitted from the sequence (no fewer than 3 and no more than 6 standard white noises burst subdivisions between omissions). In order to keep their attention on the sequence, and to provide a behavioral measure of vigilance between conditions, participants were instructed to silently count the omissions, without making any movements, and report the total number of omissions at the end of each run. The mean target-to-target interval (TTI) was kept constant for all sequences, an average of 9 s between omissions. For each of the three conditions, exactly 40 omissions occurred between the two runs. The six runs were performed in the following order for all participants, with short breaks between each run: 53%, 43%, 33%, 33%, 43%, and 53%.

At no point in the experiment were participants exposed to a 50% rhythm, with white noise bursts evenly subdividing the beat into two equal intervals. Exposure to this metrical pattern might cause distortions of rhythmic expectations in non-integer ratio conditions due to memory artifacts or priming effects, as observed by [Desain and Honing \(2003\)](#). Thus, the 50% pattern was excluded to make the strongest case that any observed distortions toward this 50% location were due to endogenous biases.

2.3. Electroencephalogram recording

The electroencephalogram (EEG) was continuously recorded from 61 Ag–AgCl electrodes (10–20 cap system) at 1000 samples per second. EEG data were filtered online at a bandpass of 0.02 Hz and 300 Hz and amplified using an EPA Sensorium bio-amplifier set at a gain of 10 K. The horizontal electrooculogram (EOG) was recorded from electrode sites F9 and F10. Additional electrodes were placed on the superior and inferior orbits to record the vertical EOG. The reference electrode was placed on the tip of the nose. EEG data were recorded using the Neuroscan Acquire 4.1 software package, with impedances on all electrode sites kept below 10 k Ω .

2.4. Signal processing and analysis

The continuous EEG was segmented offline in windows of 600 ms, including a 100 ms pre-stimulus baseline. Following

baseline correction, epochs containing eyeblinks, signal drift, or other artifacts were excluded from further analysis, as identified by a conservative voltage cutoff (if the EEG signal of nine fronto-central electrodes [F3, FZ, F4, FC3, FCZ, FC4, C3, CZ, C4] exceeded $\pm 75 \mu\text{V}$) and visual inspection of the data. Each of the three rhythm conditions was averaged separately, time-locked to the moment that an omitted sound would have occurred. Since cortical responses to adjacent pink noise bursts partially overlapped these omission epochs, OEP epochs were corrected by subtracting the aligned average auditory evoked potentials to pink noise stimuli, using only those pink noise burst intervals without white noise subdivisions or omissions.

OEP amplitude was defined as the waveform's voltage difference between minimum and maximum values (trough-to-peak) as measured at CPz (where OEP amplitude was maximal), identified separately for each participant in each rhythm condition. The peak was the maximum voltage between 0 and 450 ms post-omission at CPz (OEP peaks are typically observed within 400 ms post-omission, [Ruchkin & Sutton, 1979](#)). The trough was identified as the minimum voltage value at CPz occurring prior to the peak. The time indices of peak and trough measured at CPz were also used to measure OEP amplitudes at other electrode sites in the same time frame. OEP onset latency was measured at the nearest moment preceding the peak when the OEP was equal to half the difference between trough and peak amplitude at CPz (see [Fig. 2](#)). Using pseudo-real EEG sweeps with known onset latencies, [Kiesel, Miller, Jolicœur, and Brisson \(2008\)](#) demonstrated that this

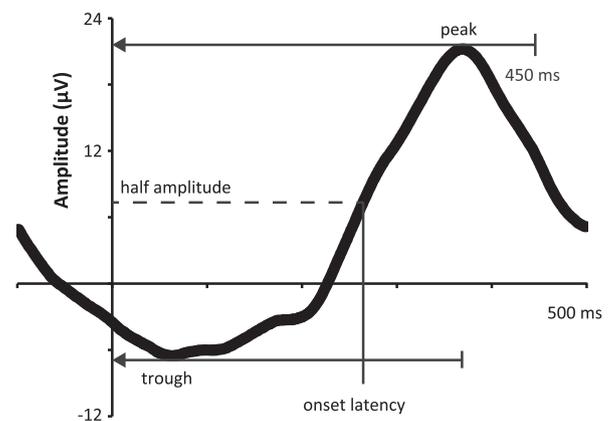


Fig. 2. Measurement a single subject's average OEP recorded at electrode site CPz, in the 33% condition. The peak is the maximum voltage between 0 and 450 ms post-omission, and the trough is the minimum voltage preceding the peak. Onset latency is the first time index preceding the peak where the amplitude is half the voltage difference between peak and trough (labeled as half amplitude).

50% relative criterion technique provides highly accurate latency estimates, significantly more accurate than simple peak latency measures (for a similar late endogenous ERP component, the P3b).

3. Results

3.1. Behavioral results

There were between 18 and 22 omissions per stimulus sequence. Across all runs and participants, the average absolute value of the difference between reported counts and the correct number of omissions was ± 0.65 ($SD = 0.67$), and 54% of reported counts were precisely correct. There was no significant difference between the number of overestimates and the number of underestimates of the actual number of omissions, $t(12) = -.857$, *ns*. Furthermore, a repeated-measures analysis of variance (ANOVA) found no significant differences in the accuracy of these reported counts between the three rhythm conditions $F(2,22) = 1.068$, *ns*, or between participants' levels of musical experience $F(1,11) = .012$, *ns*.

3.2. Omission elicited potential (OEP) results

Omissions in stimulus sequences elicited broad positive deflections in the EEG waveform, reaching maximal amplitude over centro-parietal midline scalp sites by about 400 ms after the omission. The average scalp distribution of this omission elicited potential, as observed at 100, 200, 300, and 400 ms post-omission, is shown in Fig. 3.

A three-way repeated-measures ANOVA (with Greenhouse–Geisser correction) was conducted on OEP amplitude, with the variables of rhythm condition (33/43/53%), musical experience (inexperienced/experienced), and different electrode sites (Fz/CPz/POz). There was a main effect of site, $F(2,22) = 23.307$, $p < .001$, $\eta_p^2 = .679$, with markedly higher amplitude at CPz, as was expected from the centro-parietal scalp distribution of the OEP. There was also a main effect of rhythm condition across electrode sites, $F(2,22) = 6.256$, $p = .022$, $\eta_p^2 = .363$, driven by larger OEP amplitude in the 33% and 53% conditions, relative to the 43% condition. There was no significant effect of musical experience on OEP amplitude, nor significant interactions between musical experience, condition, or site.

A three-way ANOVA at CPz also indicated significant effects of rhythm condition on OEP amplitude, $F(2,22) = 5.118$, $p = .023$, $\eta_p^2 = .318$ (see Fig. 4), again caused by increased average trough-to-peak amplitude in the 33% ($M = 17.8 \mu\text{V}$) and 53% ($M = 16.4 \mu\text{V}$) conditions compared with the 43% condition ($M = 13.2 \mu\text{V}$). Post-hoc pairwise comparisons found significant differences between both the 33% and 43% conditions ($p = 0.011$), and 53% and 43% conditions ($p = 0.015$), but not between 33% and 53%. The trough-to-peak amplitude measurements used in the present study are sensitive to both the minimum and maximum values of the ERP waveform, so these effects may be caused by increased negativity within 300 ms following the omission (which is believed to index orienting responses; see Jongsma

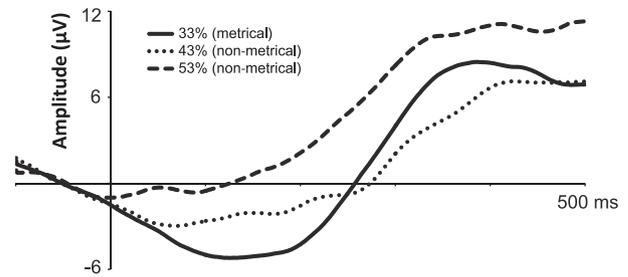


Fig. 4. Grand average OEP in the three rhythm conditions, recorded at CPz. Waveforms are time-locked to the moment a white noise burst would have onset, and baselined to the 100 ms pre-omission time-window.

et al., 2005), or by increased positivity in the later positive component of the OEP waveform. There is some interpretive difficulty teasing apart these two possibilities, given that the absolute amplitude of the 33% condition's OEP trace may have been disproportionately affected by our subtractive corrections of auditory evoked potentials associated with the pink noise bursts preceding the omission. Regardless, these results do suggest that cortical responses to an omitted subdivision vary systematically with the subdivisions' proximity to a metrical location. Omissions in a metrical sequence (33%) elicited OEPs with the largest amplitude, omissions in a closely-metrical sequence (53%, nearby 50%) elicited OEPs with moderate amplitude, and omissions in a markedly non-metrical sequence (43%, relatively distant from 50%) elicited OEPs with the smallest amplitude. To ensure that these amplitude differences were not caused by order effects or repetition of the 33% condition in the block sequence, OEP amplitude at CPz was compared between the first three stimulus blocks (53%, 43%, 33%) and the last three stimulus blocks (33%, 43%, 53%). There were no significant effects of sequence order, nor effects of musical experience on these amplitude measures at CPz, nor significant interactions between musical experience, sequence order, or condition.

There were clear differences in OEP onset latency, measured at site CPz, between rhythm conditions, $F(2,22) = 10.228$, $p = .001$, $\eta_p^2 = .482$. Positive deflections in the OEP waveform occurred earliest in the 53% (non-metrical) condition ($M = 239$ ms), next in the 33% (metrical) condition ($M = 278$ ms), and last in the 43% (non-metrical) condition ($M = 309$ ms), as shown in Fig. 5. These OEP onset latency differences are consistent with the general pattern of distortions predicted by metrical regularization. Using the OEP onset latency in the 33% condition as a reference, OEPs in the 53% condition occurred 39 ms earlier on average ($p = 0.021$), approximating the -30 ms shift to the 50% metrical subdivision between beats. Similarly, OEPs in the 43% condition were, on average, roughly 31 ms later than the 33% OEP ($p = 0.031$), in the direction of the $+70$ ms shift toward the 50% metrical location. To varying degrees, 11 out of 13 subjects showed evidence of this pattern of regularization, with earlier OEPs in the 53% condition and delayed OEPs in the 43% condition. There were no significant effects of musical experience or sequence order on OEP onset

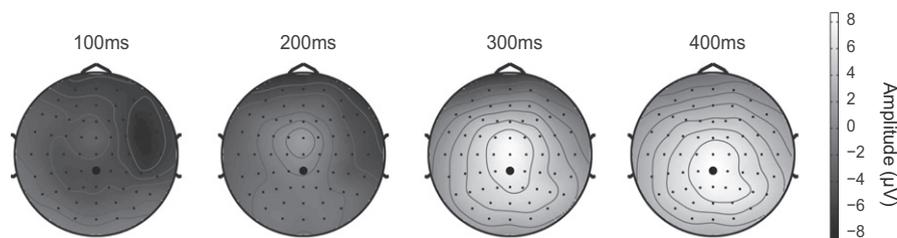


Fig. 3. Grand average scalp distribution from all subjects in all rhythm conditions at 100 ms, 200 ms, 300 ms, and 400 ms after an omitted white noise burst would have onset. The marker for electrode site CPz, where OEP amplitude is maximal, is shown in bold.

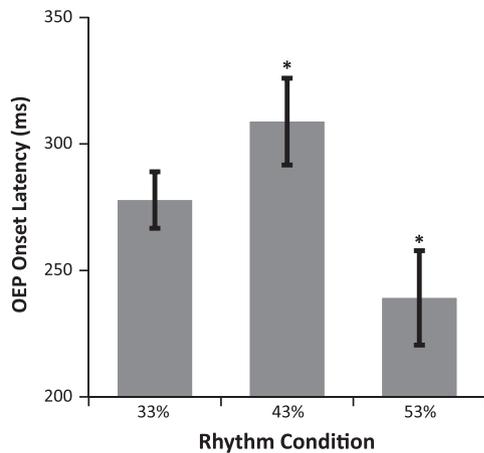


Fig. 5. OEP onset latency differences. Onset latency is measured using a relative criterion technique, at the moment the OEP waveform reaches 50% the distance between minimum and maximum amplitude. Error bars indicate 95% confidence interval. Asterisks mark significant differences from the 33% (integer ratio) condition.

latency, and similarly no significant interactions between condition, musical experience, or sequence order.

4. Discussion

The aim of the present study was to test for evidence of metrical regularization in cortical responses to rhythmic stimuli. The results revealed that listeners' covert reactions to omitted sounds at non-metrical locations in rhythmic patterns were measurably distorted in time, shifted in the direction of the nearest even subdivision of the rhythmic period.

The observed pattern of timing differences provides evidence of top-down regularization toward metrical rhythmic structures, a cognitive bias in the direction of the nearest integer ratio relationship. For both non-metrical conditions (subdivisions at 43% and 53% between beats), the nearest simple integer ratio location was 50%. As predicted, OEPs to omitted subdivisions at 43% were delayed in the direction of 50%, and similarly, OEPs to omitted subdivisions at 53% occurred earlier, compared with the timing of the OEP to a metrical omission (33%).

Similar distortions have been observed with rhythm production tasks¹, and the present results extend these findings, providing confirmatory evidence of integer ratio biases during rhythm perception. None of the presented auditory sequences contained subdivisions at 50%, so these timing distortions cannot be attributed to memory artifacts or priming effects from prior exposure to isochronous metrical patterns (e.g., Desain & Honing, 2003). The pattern of results is also inconsistent with the alternative hypothesis that expectations for non-integer ratio subdivisions are simply impoverished or delayed, since the OEP in the 53% condition occurred significantly earlier than the OEP to a metrical omission. Instead, it seems that cognitive mechanisms for rhythm perception are predisposed to distort perceived non-integer ratio rhythms in the direction of integer ratio representations.

4.1. Perceptual regularization toward metrical locations

Results from the present study support dynamical accounts of rhythm perception, and suggest that metrical patterns behave as

"attractor states" in our rhythmic repertoire. Theoretically, neural oscillators that encode time intervals between sounds (i.e., the distance between the pink noise beat and the white noise subdivision) are biased by the state of coupled oscillators representing higher-order intervals of the rhythmic sequence (i.e., the distance between pink noise beats). A preferred state of this network involves lower-order oscillators resonating at harmonics of higher-order oscillators (such integer ratio biases are unlikely without clear activation of a higher-order oscillator, as with temporally variable irregular rhythms; see Grahn & Brett, 2007). Once a clear periodicity has been established at the beat level, perception of individual sounds between beats are biased toward even subdivisions of the beat interval. In this way, integer ratio patterns, where the interval between beats is an integer-multiple of the subdivision interval, are represented accurately, and non-integer ratio patterns tend to be drawn toward the nearest metrical representation.

A critical aspect of this dynamical interpretation is that regularization is progressive, not absolute. Instead of absolute conformity to an isochronous pattern, a perceived non-integer ratio rhythm only partially approximates the nearest metrical pattern. Specifically, OEP onset latencies in the 43% condition were significantly delayed, but fell short of the full 70 ms delay required to reach the 50% metrical location. For this reason, we assert that perceived non-integer ratio rhythmic structures are not obligatorily perceived or represented as metrical patterns, instead they may be partially regularized. Our cognitive representation for a non-integer ratio rhythm is the product of a dynamical interaction between perceptual biases toward metrical patterns, and perceptual stimulation at non-metrical locations. As such, non-metrical rhythms are neither perceived as veridical non-integer ratio patterns, nor absolutely regularized to be perceived as strictly integer ratio patterns; instead they are partially regularized in the direction of the nearest metrical rhythm.

A practical example of this phenomenon occurs when listening to swing music. Performed swing music features unevenly-spaced eighth notes, commonly with ride cymbal taps played conspicuously later than the upbeat subdivision of the downbeat (in our terminology, at locations roughly 55–60% of the beat period; Benadon, 2006). While listening to this non-integer ratio swing rhythm, someone may expect upcoming sounds slightly closer toward the nearest integer ratio, but not necessarily at this nearest integer ratio. The present results suggest that such non-metrical rhythms are not distorted to be perceived as "deadpan," but we nevertheless expect sounds slightly earlier than they actually occur. Moreover, since our expectations slightly precede the moment when each sound will occur, the listener may even repeatedly "trigger" an OEP (or the listener has primed such processes to occur) with each delayed hit of the ride cymbal. The inability to precisely anticipate the timing of off-beat sounds may contribute to the swing rhythm's subjective aesthetic (Butterfield, 2011). Listeners do readily detect and appreciate the expressive qualities of temporal deviations in music (Chapin, Jantzen, Kelso, Steinberg, & Large, 2010; Grahn & Rowe, 2009), and furthermore, listeners have preconceived notions about whether upcoming beats should be accelerated or delayed (Repp, 1999). A controlled investigation of the musical context that may elicit regularization, both accelerated and delayed toward the nearest metrical location, is a potentially fruitful area for future research.

Given that regularization is a dynamic process, expectations for non-integer ratio sounds should be impaired when these subdivisions occur further from attractor states. There was evidence for such a qualitative difference between non-integer ratio conditions, as OEP amplitudes to omitted subdivisions were enhanced when the subdivisions were closer to simple metrical locations. The OEP in the 53% condition (near to 50%) was significantly larger than the OEP in the 43% condition (relatively distant from 50%). OEP

¹ In a synchronization task, Repp et al. (2011) observed non-integer ratios around 44% to be significantly distorted toward 33%, which may be a range effect, given repeated exposures and attempts to synchronize with patterns featuring subdivisions nearer to 33% in that experiment (see also Repp, London, & Keller, 2012). In addition to the present study, others have observed ratios around 43% distorted toward 50% (e.g., Snyder, Hannon, Large, & Christiansen, 2006).

amplitude has been shown to increase when the omitted stimulus is more predictable (Besson et al., 1994; Janata, 2001; Ruchkin, Sutton, & Tueting, 1975), suggesting that expectancy is improved for more metrical rhythmic patterns. This finding is consistent with an abundance of evidence for processing advantages with metrical rhythms. The fact that the OEP in the 53% condition was significantly larger than the 43% condition also precludes the possibility that our results are confounded by the omission's distance from the preceding pink noise burst, since such effects would predict higher amplitude when the omission is nearer the preceding beat (Penney, 2004). But in spite of the observed amplitude differences, OEPs to omissions in both non-integer ratio rhythms were still robust, and while the strength of our expectations for sounds at non-integer ratio locations may be depleted relative to metrical expectations, reduced predictability cannot fully account for the present finding that our expectations are systematically regularized in time in the direction of integer ratio relationships.

Despite our focus on two simple integer ratios in the present study (1:2 and 1:3), these are not the only conceivable attractor states in the beat period. For instance, the 43% condition is very close a 2:5 metrical subdivision (40%); why not regularize the 43% white noise burst toward this nearer metrical location? Evidence suggests that, despite the metricality of a 2:5 integer ratio pattern (or any complex metrical pattern), simpler subdivisions are cognitively preferred rhythmic forms for Western listeners. Many studies have shown that rhythmic patterns with fewer subdivisions (e.g., 1:2 and 1:3) are better represented than those with more complex subdivisions (Cummins & Port, 1998; Essens & Povel, 1985; Fraisse, 1982; Jones & Yee, 1997; Desain & Honing, 2003; Grahn & Brett, 2007; Hannon & Trehub, 2005). These preferences for simpler subdivisions may be due to intrinsic limitations of time perception. In the present experiment, the white noise burst at 430 ms in the 1000 ms beat interval could conceivably be accommodated by segmenting the beat period into five equal 200 ms intervals (regularizing the white noise subdivision toward 400 ms, 40%), but these small intervals may be too short to effectively represent (maximum sensitivity for isochronous patterns is between 300 and 900 ms; Drake & Botte, 1993). Fewer and longer subdivisions in a metrical pattern, such as 1:2 and 1:3, may yield improved representations even despite the inaccuracies associated with regularization. However, these biases for simpler subdivisions are not entirely caused by tempo sensitivity; they also emerge via experience (experience may even override any effects of tempo sensitivity; Hannon, Soley, & Ullal, 2012). Complex musical meters that, to a Western ear, are difficult to perceive may be processed more easily by someone acculturated (or even briefly exposed during infancy) to that metrical pattern in a different musical tradition (e.g., Bulgarian or Macedonian; Hannon & Trehub, 2005; see also Hannon, Soley, & Levine, 2011).

4.2. Regularization in perception and production tasks

The present study extends the findings of Repp et al. (2011), and provides strong confirmatory evidence of metrical regularization in a purely perceptual paradigm, without overt movement or discrimination judgments, instead using cortical responses as temporal indices of expectations. Similar biases toward metrical patterns also apply to the motoric production of rhythmic actions; individual gestures in rhythmic behaviors (finger taps, steps in dance, stressed syllables in speech, etc.) tend to conform toward simple even subdivisions of higher-order coordinated actions (Collier & Wright, 1995; Cummins & Port, 1998; Essens, 1986; Repp et al., 2011; Sakai et al., 1999). For example, in the case of rhythmic speech, if you were to say the phrase “beg for a dime” repetitively, you would progressively tend to place the stressed syllable “dime,” at the moment equal to 50% or 33% the length of the repeating per-

iod, even when impelled to place the word “dime” at a non-integer ratio location marked by a co-occurring tone pattern (Cummins & Port, 1998). Producing a rhythmic behavior (such as repeating a phrase, tapping, or dancing) dynamically constrains and distorts the timing of individual gesture production (Saltzman & Byrd, 2000). However, given that rhythmic movement also creates a rhythmic perceptual experience, it may be the case that these behavioral findings are the result of motor coupling to perceptually-induced rhythmic representations (Repp et al., 2011; Wilson, Collins, & Bingham, 2005; Wilson, Snapp-Childs, & Bingham, 2010). The tendency to regularize toward metrical patterns in both perception and production tasks suggests that there are interdependent, if not common, substrates for rhythm perception and action, a hypothesis receiving favorable treatment from diverse approaches (e.g., Bingham, 2004; Chen, Penhune, & Zatorre, 2008; Grahn & Brett, 2007; Large, 2000; Overy & Molnar-Szakacs, 2009).

4.3. Specificity

On the surface, theoretical accounts of similar perceptual treatment of both metrical and non-metrical patterns may seem inconsistent with functional brain imaging findings. At least three studies have observed distinct patterns of brain activity in functional magnetic resonance imaging (fMRI) scans during perception of metrical and non-metrical rhythmic sequences (Bengtsson et al., 2009; Grahn & Brett, 2007; Sakai et al., 1999). If, as the present data suggest, we distort non-integer ratio sequences toward the nearest metrical structure, should not perception of these two patterns evoke common neural substrates? Not necessarily. Even if there were a common mechanism for induction of integer ratio and non-integer ratio rhythms, we should expect broad functional differences during engagement with these two types of sequences resulting from non-metrical challenges specific to the experimental task, anticipation errors in the non-metrical condition, reinforcement of well-timed expectations in the metrical condition, or from the process of regularization itself; such differences are not necessarily indicative of distinct representational mechanisms for metrical and non-metrical rhythms. Nevertheless, we do not assume that all mental temporal processing functions are isolated to a single cortical mechanism, and some of these activation differences between metrical and non-metrical conditions may reflect strategic attempts to represent non-metrical conditions using a cognitive timekeeper other than dynamical oscillators.

The present findings of rhythmic expectations becoming regularized in the direction of metrical patterns may be taken as evidence supporting oscillator theories of temporal processing, but should not be interpreted as evidence for the exclusivity of oscillator mechanisms for the perception of time intervals. Some temporal processing functions may occur intrinsically, without any dedicated clock mechanism (Karmarkar & Buonomano, 2007; c.f. Ivry & Schlerf, 2008). Moreover, alternative theories of dedicated clocks, such as pulse-accumulator models (e.g., Ivry & Hazeltine, 1995), also have explanatory power in time perception tasks, and recent theorists have proposed that multiple dedicated temporal processing mechanisms might coexist in our cognitive repertoire (Grahn & McAuley, 2009; Grube, Cooper, Chinnery, & Griffiths, 2010; Ivry & Richardson, 2002; Lewis & Miall, 2003; Snyder, Pasinski, & McAuley, 2011). Central to this compromise is the idea that distinct timing mechanisms (e.g., beat-based oscillators and interval-based accumulators) might be activated by listeners in different contexts, and that timing ability arises from selective engagement of the appropriate clock mechanism, depending on the task. The coexistence of distinct mechanisms makes intuitive sense: oscillator strategies are better-equipped to represent naturalistic temporal structures such as those involved in speech and

movement perception, while interval-based accumulator strategies are better-equipped to represent discrete time intervals that are not embedded in regular metrical temporal patterns (Grondin, 2010). Along these lines, some researchers have suggested that representation of non-integer ratio rhythms might be improved for individuals who explicitly encode inter-stimulus-intervals in rhythmic sequences using an interval-based cognitive mechanism (Grahn & Brett, 2007; Sakai et al., 1999).

5. Conclusion

The results of the present study demonstrate that individuals employ effective cortical representations of non-integer ratio sequences, and that these representations are reliably temporally distorted, regularized in the direction of the nearest metrical structure. When presented with the challenges of a non-integer ratio rhythmic pattern, the listener may have slightly diminished expectations, but nevertheless, their expectations will be shifted in the direction of the nearest idealized rhythm.

Acknowledgments

The authors are grateful to Becky Gerber, who never skipped a beat with data collection. The authors also thank Bruno H. Repp and two anonymous reviewers for their insightful comments on earlier versions of this manuscript. This work was supported in part by NIMH grant R01 MH074983-01 to W.H., a NSF Graduate Research Fellowship to M.E., and supplementary internal funding from the Indiana University Department of Psychological and Brain Sciences.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bandc.2013.01.005>.

References

- Benadon, F. (2006). Slicing the beat: Jazz eighth-notes as expressive microrhythm. *Ethnomusicology*, 50, 73–98.
- Bendixen, A., Schröger, E., & Winkler, I. (2009). I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory system. *Journal of Neuroscience*, 29, 8447–8451.
- Bengtsson, S. L., Ullén, F., Ehrsson, H. H., Hashimoto, T., Kito, T., Naito, E., et al. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex*, 45, 62–71.
- Besson, M., Faita, F., Czernasty, C., & Kutas, M. (1997). What's in a pause: Event-related potential analysis of temporal disruptions in written and spoken sentences. *Biological Psychology*, 46, 3–23.
- Besson, M., Faita, F., & Requin, J. (1994). Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neuroscience Letters*, 168, 101–105.
- Bingham, G. P. (2004). A perceptually driven dynamical model of bimanual rhythmic movement (and phase perception). *Ecological Psychology*, 16, 45–53.
- Butterfield, M. W. (2011). Why do jazz musicians swing their eighth notes? *Music Theory Spectrum*, 33, 3–26.
- Chapin, H., Jantzen, K., Kelso, J. A. S., Steinberg, F., & Large, E. (2010). Dynamic emotional and neural responses to music depend on performance expression and listener experience. *PLoS One*, 5, e13812.
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, 18, 2844–2854.
- Collier, G. L., & Wright, C. E. (1995). Temporal rescaling of simple and complex ratios in rhythmic tapping. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 602–627.
- Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics*, 26, 145–171.
- Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, 32, 341–365.
- Deutsch, D. (1986). Recognition of durations embedded in temporal patterns. *Perception and Psychophysics*, 39, 179–186.
- Drake, C., & Botte, M. C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception and Psychophysics*, 54, 277–286.
- Eck, D. (2002). Finding downbeats with a relaxation oscillator. *Psychological Research Psychologische Forschung*, 66, 18–25.
- Essens, P. J. (1986). Hierarchical organization of temporal patterns. *Perception and Psychophysics*, 40, 69–73.
- Essens, P. J., & Povel, D.-J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception and Psychophysics*, 37, 1–7.
- Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). New York: Academic Press.
- Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*, 1169, 89–92.
- Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, 19, 893–906.
- Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage*, 47, 1894–1903.
- Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, 29, 7540–7548.
- Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, 72, 561–582.
- Grube, M., Cooper, F. E., Chinnery, P. F., & Griffiths, T. D. (2010). Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 11597–11601.
- Grube, M., & Griffiths, T. D. (2009). Metricity-enhanced temporal encoding and the subjective perception of rhythmic sequences. *Cortex*, 45, 72–79.
- Hannon, E. E., Soley, G., & Levine, R. S. (2011). Constraints on infants' musical rhythm perception: Effects of interval ratio complexity and enculturation. *Developmental Science*, 14, 865–872.
- Hannon, E. E., Soley, G., & Ullal, S. (2012). Familiarity overrides complexity in rhythm perception: A cross cultural comparison of American and Turkish listeners. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 543–548.
- Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, 16, 48–55.
- Hernández, O. H., & Vogel-Sprott, M. (2010). Reaction time and brain waves in omitted stimulus tasks: A multisensory study. *Journal of Psychophysiology*, 24, 1–6.
- Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 3–18.
- Ivry, R. B., & Richardson, T. C. (2002). Temporal control and coordination: The multiple timer model. *Brain Cognition*, 48, 117–132.
- Ivry, R. B., & Schlerf, R. B. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences*, 12, 273–280.
- Janata, P. (2001). Brain electrical activity evoked by mental formation of auditory expectations and images. *Brain Topography*, 13, 169–193.
- Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, 13, 313–319.
- Jones, M. R., & Yee, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 693–709.
- Jongsma, M. L., Eichele, T., Quian, Q. R., Jenks, K. M., Desain, P., Honing, H., et al. (2005). Expectancy effects on omission evoked potentials in musicians and non-musicians. *Psychophysiology*, 42, 191–201.
- Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: Encoding time in neural network states. *Neuron*, 53, 427–438.
- Kiesel, A., Miller, J., Jolicœur, P., & Brisson, B. (2008). Measurement of ERP latency differences: A comparison of single-participant and jackknife-based scoring methods. *Psychophysiology*, 45, 250–274.
- Large, E. W. (2000). On synchronizing movements to music. *Human Movement Science*, 19, 527–566.
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119–159.
- Large, E. W., & Kolen, J. F. (1995). Resonance and the perception of musical meter. *Connection Science*, 6, 177–208.
- Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, 1169, 89–92.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Lewis, P. A., & Miall, R. C. (2003). Overview: An image of human neural timing. In W. H. Meck (Ed.), *Functional and neural mechanisms of interval timing* (pp. 515–532). Boca Raton, FL: CRC Press.
- Macar, F., & Vidal, F. (2004). Event-related potentials as indices of time processing: A review. *Journal of Psychophysiology*, 18, 89–104.
- Meumann, E. (1894). Untersuchungen zur psychologie und ästhetik des rhythmus [Studies on the psychology and aesthetics of rhythm]. *Philosophical Studies*, 10, 249–322.
- Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception*, 26, 489–504.
- Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 728–741.
- Patel, A. D., Iverson, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricity and modality on synchronization with a beat. *Experimental Brain Research*, 163, 226–238.

- Penney, T. B. (2004). Electrophysiological correlates of interval timing in the stop-reaction-time task. *Cognitive Brain Research*, 21, 234–249.
- Povel, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 3–18.
- Povel, D.-J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, 2, 411–440.
- Repp, B. H. (1999). Detecting deviations from metronomic timing in music: Effects of perceptual structure on the mental timekeeper. *Perception and Psychophysics*, 61, 529–548.
- Repp, B. H., London, J., & Keller, P. E. (2011). Perception–production relationships and phase correction in synchronization with two-interval rhythms. *Psychological Research Psychologische Forschung*, 75, 227–242.
- 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, 205–223.
- Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm’s gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia*, 50, 232–244.
- Ruchkin, D., & Sutton, S. (1973). Visual evoked and emitted potentials and stimulus significance. *Bulletin of the Psychonomic Society*, 2, 144–146.
- Ruchkin, D., & Sutton, S. (1979). Latency characteristics and trial-by-trial variations of emitted cerebral potentials. In J. E. Desmedt (Ed.), *Cognitive components in cerebral event related potentials and selective attention* (pp. 106–118). Basel: Karger.
- Ruchkin, D., Sutton, S., & Tueting, P. (1975). Emitted and evoked P300 potentials and variation in stimulus probability. *Psychophysiology*, 18, 271–282.
- Ruckmich, C. A. (1913). The role of kinaesthesia in the perception of rhythm. *American Journal of Psychology*, 24, 305–359.
- Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada, T., Iwata, N. K., et al. (1999). Neural representation of a rhythm depends on its interval ratio. *Journal of Neuroscience*, 19, 10074–10081.
- Saltzman, E. L., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, 19, 499–526.
- Snyder, J. S., Hannon, E. E., Large, E. W., & Christiansen, M. H. (2006). Synchronization and continuation tapping to complex meters. *Music Perception*, 24, 135–146.
- Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, 24, 117–126.
- Snyder, J. S., Pasinski, A. C., & McAuley, J. D. (2011). Listening strategy for auditory rhythms modulates neural correlates of expectancy and cognitive processing. *Psychophysiology*, 48, 198–207.
- Sutton, S., Tueting, P., Zubin, J., & John, E. R. (1967). Information delivery and the sensory evoked potential. *Science*, 155, 1436–1439.
- Treisman, M., Faulkner, A., Naish, P. L. N., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception*, 19, 705–743.
- Wilson, A. D., Collins, D. R., & Bingham, G. P. (2005). Perceptual coupling in rhythmic movement coordination: Stable perception leads to stable action. *Experimental Brain Research*, 164, 517–528.
- Wilson, A. D., Snapp-Childs, W., & Bingham, G. P. (2010). Perceptual learning immediately yields new stable motor coordination. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 1508–1514.
- Zanto, T. P., Large, E. W., Fuchs, A., & Kelso, J. A. S. (2005). Gamma-band responses to perturbed auditory sequences: Evidence for synchronization of perceptual processes. *Music Perception*, 22, 535–552.