Title

Mapping the dynamic allocation of temporal attention in musical patterns.

Permalink

Journal

ISSN

Authors

Hurley, Brian K
Fink, Lauren K
Janata, Petr

Publication Date

DOI

Peer reviewed

Running Head: MAPPING DYNAMIC TEMPORAL ATTENTION

Mapping the dynamic allocation of temporal attention in musical patterns

Brian K. Hurley

University of California, Davis

Lauren K. Fink

University of California, Davis

Petr Janata

University of California, Davis

Author Note

Brian K. Hurley, Department of Psychology and Center for Mind and Brain, University of California, Davis; Lauren K. Fink, Neuroscience Graduate Group and Center for Mind and Brain, University of California, Davis; Petr Janata, Department of Psychology and Center for Mind and Brain, University of California, Davis.

Correspondence concerning this article should be addressed to Petr Janata, Center for Mind and Brain, University of California, Davis, 267 Cousteau Place, Davis 95618. E-mail: pjanata@ucdavis.edu

Word Count: 12,136

MAPPING DYNAMIC TEMPORAL ATTENTION

Abstract

Many environmental sounds, such as music or speech, are patterned in time. Dynamic attending theory (DAT), and supporting empirical evidence, suggests that a stimulus' temporal structure serves to orient attention to specific moments in time. One instantiation of this theory posits that attention synchronizes to the temporal structure of a stimulus in an oscillatory fashion, with optimal perception at salient time points or oscillation peaks. We examined whether a model consisting of damped linear oscillators succeeds at predicting temporal attention behavior in rhythmic multi-instrumental music. We conducted three experiments in which we mapped listeners' perceptual sensitivity by estimating detection thresholds for intensity deviants embedded at multiple time points within a stimulus pattern. We compared participants' thresholds for detecting intensity changes at various time points with the modeled salience prediction at each of those time points. Across all experiments, results showed that the resonator model predicted listener thresholds, such that listeners were more sensitive to probes at time points corresponding to greater model-predicted salience. This effect held for both intensity increment and decrement probes and for metrically simple and complex stimuli. Moreover, the resonator model explained the data better than predictions based on canonical metric hierarchy or auditory scene density. Our results offer new insight into the temporal orienting of attention in complex auditory scenes using a parsimonious computational model for predicting attentional dynamics.

Keywords: dynamic attending, oscillator, resonator, rhythm, adaptive threshold, psychophysics, model

MAPPING DYNAMIC TEMPORAL ATTENTION

Public Significance Statement

Sounds in the environment, such as music or speech, typically have a rhythmic temporal structure. Previous research suggests that listeners do not distribute their attention to all moments in time equally, but rather focus their attention on certain points in time. In this study, we used a computational model to predict when listeners were most likely to be attending within rhythmic music patterns. We tested our predictions by asking listeners to detect subtle changes at various time points throughout several patterns. In comparing the listeners' data to our model predictions, we showed that it is possible to compute a fine-grained prediction of how humans orient their attention in time.

MAPPING DYNAMIC TEMPORAL ATTENTION

Mapping the dynamic allocation of temporal attention in musical patterns

Periodic structure exists in many parts of the acoustic environment, including music, speech, and non-human animal calls. To efficiently process such stimuli, a listener must parse the timing of acoustic events and selectively allocate attention to the times at which salient events are expected to occur. The prevailing account of how this is accomplished posits an underlying oscillatory process, whereby a sinusoidal fluctuation of attention synchronizes to the period and phase of external rhythms, with peaks in the attentional oscillation corresponding to maximal attentional focus (Jones & Boltz, 1989; Large & Jones, 1999). In this paper, we ask whether damped linear oscillators can reliably predict temporal attention in rhythmic musical sequences.

Evidence suggests that the temporal structure of rhythmic stimuli implicitly orients attention to moments in time. Dynamic attending theory (DAT) postulates that the processing of rhythmic sequences entails synchronization between "attunement rhythms" (Jones & Boltz, 1989) or "attending rhythms," modeled as oscillators that represent endogenous, periodic fluctuations in neural resources (Large & Jones, 1999), and external stimulus rhythms. According to DAT, attending rhythms become entrained to a regularly timed stimulus rhythm (e.g., the beat in music or the syllable pattern in speech) such that a stimulus event is most salient when its onset timing is in synchrony with the established entraining rhythm.

Computational models of DAT by Large and colleagues formalize rhythmically driven oscillatory attention using nonlinear dynamical systems, in which oscillators entrain to external stimulus periodicities and affect inter-oscillator activity by entraining to one another (Large & Jones, 1999; Large & Kolen, 1994; Large & Palmer, 2002; Large & Snyder, 2009; Large,

Herrera, & Velasco, 2015). Alternatively, DAT can be modeled using systems of linear oscillators. Linear oscillators are driven by periodic stimulation at their tuned periodicities, scale linearly in their responses to stimulus amplitude, and, when arranged in banks of oscillators tuned to multiple periodicities, are sensitive to hierarchically nested periodicity structures (Tomic & Janata, 2008). Both linear and nonlinear oscillators share filtering properties in that they excite maximally to stimulus periodicities that match their tuning. However, nonlinear oscillators are considerably more complex in their behavior, as they respond nonlinearly to stimulus amplitude, oscillate spontaneously, and exhibit higher-order resonance, none of which linear oscillators do (Large, 2010). The computational components – such as parameters for bifurcation, nonlinear dampening, and higher-order terms – that enable these nonlinear behaviors also contribute to the increased complexity. As such, if a linear oscillator model succeeded at predicting temporal attention behavior, it would offer a more computationally parsimonious instantiation of DAT compared to those of nonlinear models.

In the current study, we assess the precision with which a model built upon banks of reson filters – a type of linear oscillator – predicts dynamic attending behavior with musical rhythms. Our model differs from similar linear models of rhythm and meter perception in (1) its use of reson filters, which are well suited for extracting multiple periodicities of an acoustic signal, rather than comb filters (Klapuri, Eronen, & Astola, 2006; Scheirer, 1998) or Gaussian filters (Todd, 1994), and (2) its objective to estimate the continuous, multi-level periodicity structure of acoustic signals, rather than just the beat or the tempo (e.g., Parncutt, 1994; Scheirer, 1998). We direct readers to Tomic and Janata (2008) for a comprehensive review of related linear rhythm and meter models.

Several studies support DAT, showing that the detection and judgment of stimulus events embedded within a rhythmic context are enhanced when the stimulus rhythm comprises isochronous inter-onset intervals (IOIs) rather than irregular IOIs, and when targets are presented in phase rather than out of phase with the entraining rhythm. This effect is observed for judgments on multiple stimulus dimensions, including discriminations of interval duration, event duration, and pitch (Barnes & Jones, 2000; Jones, Moynihan, MacKenzie, & Puente, 2002; Large & Jones, 1999; McAuley & Fromboluti, 2014), although recent work questions the robustness of DAT effects in pitch comparison paradigms (Bauer et al., 2015). Evidence shows that rhythm also drives attentional dynamics in speech processing. For instance, phoneme detection is faster when the syllable containing the target phoneme is aligned with a preceding rhythmic pattern (Cason & Schön, 2012; Pitt & Samuel, 1990). Furthermore, entrainment to rhythms also influences perception across modalities, such that visual target detection is enhanced when the target is presented in synchrony with a concurrently playing auditory entraining rhythm (Bolger, Trost, & Schön, 2013; Miller, Carlson, & McAuley, 2013).

Behavioral evidence of rhythmically guided attention also converges with neurophysiological evidence. Numerous studies report modulation of attention-related components of the event-related potential (ERP) as a function of rhythmic timing. For example, the P300 is enhanced when participants detect deviant stimuli embedded in isochronous rather than randomly timed sequences (Otterbein, Abel, Heinemann, Kaiser, & Schmidt-Kassow, 2012; Schmidt-Kassow, Schubotz, & Kotz, 2009; Schwartze, Rothermich, Schmidt-Kassow, & Kotz, 2011) and N100 modulations are observed when an auditory target appears in phase with a preceding isochronous sequence (e.g., Lange, 2009, 2010; Sanabria & Correa, 2013). Studies of

MAPPING DYNAMIC TEMPORAL ATTENTION

EEG and MEG spectra provide additional insights into rhythmically driven perceptual enhancements. For instance, the frequency spectra of auditory steady-state evoked potentials from individuals listening to rhythmic patterns reveal increased neural activity at beat- and meter-related periodicity frequencies (Nozaradan, Peretz, Missal, & Mouraux, 2011; Nozaradan, Peretz, & Mouraux, 2012). Furthermore, fluctuations of induced gamma-band (20-60 Hz) oscillatory activity coincide with and even precede expected tone onsets in rhythmic patterns (Fujioka, Trainor, Large, & Ross, 2009; Snyder & Large, 2005; Zanto, Large, Fuchs, & Kelso, 2005; Zanto, Snyder, & Large, 2006). Recent work in monkey physiology suggests that, in the presence of rhythmic stimuli, attention selectively entrains low-frequency (e.g., 1-4 Hz) neuronal oscillations to the stimulus such that sensory response gain is amplified at rhythmically expected time points (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010). Moreover, cross-frequency coupling of attention-related neuronal oscillations allows for response excitation across multiple timescales (Schroeder & Lakatos, 2009). These observations support the presence of an oscillatory pattern in temporal attention to rhythms and provide a systems-level brain mechanism for DAT.

Given its rhythmic make-up, most Western music contains multiple, harmonically related periodicities that combine to create an accent structure and elicit an endogenously generated temporal hierarchy, termed *meter* (London, 2012). The most salient temporal level, or periodicity, within this hierarchy is called the beat. Patterns of strong (accented) and weak (unaccented) beats, as well as patterns of rhythmic event onsets between beats, give rise to additional periodicities at integer multiples and subdivisions of the beat period. By probing for goodness-of-fit ratings for probe tones which occurred on various metrical levels for a range of

metrical structures, Palmer and Krumhansl (1990) showed that listeners regarded tones as more expected when they occurred at higher points in the theoretical metric hierarchy (e.g., the first beat in a four-beat sequence).

DAT posits that frequency-tuned attending rhythms entrain to the various periodicities present in a rhythmic stimulus, and that the alignment of peaks across simultaneously entrained, harmonically related attending rhythms dynamically bias attention to moments of high metrical salience (Jones & Boltz, 1989; Large & Jones, 1999; Large & Palmer, 2002). Studies of rhythmic attending using metrical rhythmic stimuli appear to support this prediction. For instance, Repp (2010) had participants detect changes in intensity and duration of notes within metrically structured melodies. Participants detected these deviants most accurately at metrically strong positions relative to metrically weak positions. Similarly, when participants perform visual and auditory target detection tasks while exposed to an auditory metrical entraining stimulus, response times are generally faster when targets occur at metrically strong positions than at metrically weak positions (Bolger, Coull, & Schön, 2014; Bolger et al., 2013). Additionally, metric timing elicits attention-related modulations in the ERP, such that the detection of targets at metrically strong positions in auditory rhythms evokes greater amplitude and decreased latency in the P300 response (Abecasis, 2005; Brochard, Abecasis, Potter, Ragot, & Drake, 2003; Potter, Fenwick, Abecasis, & Brochard, 2009).

Most of the research on temporal attending has utilized isochronous rhythms with accenting used to impart hierarchical structure across multiple timespans. However, musical rhythm contains much richer temporal structure than simply isochrony. Moreover, music often contains rhythmically salient acoustic events that do not appear on strong beats, such as

MAPPING DYNAMIC TEMPORAL ATTENTION

syncopations, in which salient events occur on weak beats or subdivisions of the beat. Although previous work has investigated the mental representation of more complexly structured rhythms (e.g. Large & Palmer, 2002; Essens & Povel, 1985), few have addressed the dynamic orienting of attention in such rhythms. In the current study, we examine temporal attention within stimuli that reflect ethologically relevant temporal patterning.

We investigate rhythmic processing by combining a computational model of temporal attention with psychophysical mapping of attention allocation in musical auditory scenes. We examine the extent to which, in comparison to music-theoretic models of metric organization from which estimates of temporal salience might be derived, the parametric outputs of our resonator model of attention are able to predict perceptual thresholds at different locations within a metric structure. We approached these questions using a deviance detection paradigm in which transient intensity changes were embedded in percussion patterns possessing rhythmic complexity found in natural music. We estimated the temporal orienting of attention to specific moments in time by measuring deviance-detection thresholds at multiple time points in each stimulus pattern. We then assessed the resonator model's salience predictions by comparing the measured thresholds to model output at the corresponding time points. We hypothesized that participants would be more sensitive to deviants presented at moments of higher model-predicted salience, which we interpreted as evidence for greater stimulus-driven attentional focus.

MAPPING DYNAMIC TEMPORAL ATTENTION

Experiment 1

Method

Participants. Twenty-eight undergraduate students (13 females) aged 18 to 43 ($M =$ 21.7) at the University of California, Davis, participated in exchange for course credit. Twelve participants reported having at least three years of musical instrument or vocal training. That group of participants had 7.8 years of musical training on average. The experiments reported in this paper were not designed to test differences between musicians and non-musicians. Because such comparisons would therefore be under-powered, we report descriptive musicianship statistics for all experiments but do not include musicianship as a variable in our analyses. For all experiments reported in this paper, participants provided informed consent in accordance with a protocol approved by the Institutional Review Board of the University of California, Davis.

Stimuli and apparatus. We used five stimuli from a previous, unpublished experiment in our lab; those stimuli were created using a custom audio sample sequencer written in Max/MSP (Cycling '74). Each stimulus was a three-timbre percussion pattern comprising conga, shaker, and snap instrument timbres from a Proteus 2000 sound module (E-mu Systems, Scotts Valley, CA). In all stimuli but one, the conga was a mid-conga (i.e. with an intermediate pitch height). In one stimulus (complex2) the conga timbre was a high-pitched conga. As these stimuli were selected from a previous, unrelated experiment, this timbre difference was not by design for the current study. Moreover, as we are primarily concerned in this paper with the aggregate temporal structure of the stimulus patterns rather than with the individual timbres, we were not concerned about the timbral difference between complex2 and the other stimuli. Nevertheless, for reference, we calculated each instrument timbre's fundamental frequency (f0), spectral

centroid (a measure of brightness), attack time, and intensity using the Music Information

Retrieval Toolbox (Lartillot, Toiviainen, & Eerola, 2008) in MATLAB (Mathworks, Natick,

MA). The mid-conga sample had an f0 of 250 Hz, a spectral centroid of 599 Hz, an attack time

of 27 ms, and an average rms intensity of 5,789. The high conga had an f0 of 800 Hz, a spectral

centroid of 3,751 Hz, an attack time of 16 ms, and an average rms intensity of 2,743. The shaker

had an f0 of 255 Hz, a spectral centroid of 10,466 Hz, an attack time of 45 ms, and an rms

intensity of 1,864. The snap had an f0 of 734 Hz, a spectral centroid of 4,493 Hz, an attack time

of 37 ms, and an rms intensity of 2,506.

Stimulus patterns were arranged as single measures in 4/4 meter at 107 beats per minute

(BPM) and were presented as continuously repeating loops. As each loop played, intensity

increases could occur at any of four time points, each associated with varying levels of predicted

perceptual salience (Figure 1; see supplemental materials for audio examples of stimuli), as

assessed using Tomic & Janata's (2008) resonator model (see the *Resonator model* section).

These intensity increases occurred globally across all voices of the rhythmic pattern and lasted

200 ms, with an on and offset ramp of 1ms. A given stimulus loop iteration could contain only

one randomly selected intensity increment probe. After a probe was presented, the probability of

a probe in the following loop iteration was 0. After a standard (i.e. non-deviant) loop iteration,

the probability of a probe on the subsequent loop iteration was .80. The first author developed

Max/MSP software to play stimuli, dynamically apply increment probes, implement a dynamic

threshold algorithm (see *Adaptive threshold* section), and collect responses. Stimuli were played

at a comfortable listening level through two Tannoy Reveal 601p studio monitor speakers

MAPPING DYNAMIC TEMPORAL ATTENTION

(Tannoy Ltd, Coatbridge, UK), with each speaker situated approximately one meter away from participants. The experiment took place in a sound-attenuating chamber.

Resonator model. We assessed stimulus periodicity structure by processing four continuous repetitions of each of the stimulus loops through a resonator model (Tomic & Janata, 2008) (Figure 2). A key feature of this model is its use of *reson* filters – damped linear oscillators that are sensitive to periodically recurring onsets – to model the periodicity structure of an input signal. The model simulates the filtering, resonance, and transduction mechanisms of the peripheral auditory system, as implemented in the IPEM toolbox (Leman, Lesaffre, & Tanghe, 2001), which decomposes the input signal into 40 critical bands of auditory nerve firing rate codes. It then estimates onset patterns within each critical band and sums every eight adjacent channels, resulting in five bands. Onset patterns in each of the five bands are passed through a bank of 99 reson filters tuned from 0.25 to 10 Hz, resulting in five bands of reson filter output. A windowed root-mean-square (rms) calculation is applied to each reson filter's output and is treated as a metric of energy within each of the 99 resonator frequencies. The five bands of rms are averaged to create an average periodicity surface (APS), a spectrogram of energy within periodicity frequencies over time. The APS is then averaged across time to produce a mean periodicity profile (MPP). In a representation similar to a frequency amplitude spectrum, the MPP reveals the prominent periodicity frequencies found in the input signal (Figure 2B). We encourage readers to examine Tomic and Janata's (2008) paper for a detailed description of the model.

Model metrics for selecting probe timing. We combined information from stimulus MPPs with resonator output to generate a single time series of estimated salience dynamics, which we used to select probe times of greater and lesser modeled salience. First, we found each

stimulus' most prominent periodicities by identifying the peaks of the MPP that were greater

than 5% of the MPP's amplitude range. Next, for each of the five bands of reson filter output, we

assessed resonator activity solely at the peak periodicities from the MPP (Figure 2). Finally, for

each time point within the rhythmic pattern, we calculated the mean peak-periodicity resonator

amplitude across all five bands, which yielded a single time series for each stimulus (Figure 2A;

also depicted in Figure 1). We used this time course as our estimate of perceptual salience.

Adaptive Threshold. At the beginning of a stimulus block, increment probes (intensity

increments) at each of the four time points started with a 10 dB SPL amplitude increase,

representing a "best guess" of participants' difference limens. Throughout the remainder of each

stimulus block, the magnitudes of increments were adjusted by an adaptive threshold procedure,

Zippy Estimation by Sequential Testing (ZEST), which has been demonstrated to converge on

threshold estimates with speed, accuracy, and reliability (King-Smith, Grigsby, Vingrys, Benes,

& Supowit, 1994; Marvit, Florentine, & Buus, 2003) in multiple probe conditions simultaneously

(Navarro Cebrian & Janata, 2010a, 2010b). We implemented the ZEST algorithm with a set of

custom MAX/MSP patches. Under ZEST, an *a priori* probability density function (p.d.f.) is

specified to reflect an assumed probability distribution of a listener's threshold. The mean of the

initial p.d.f. determines the difference level of the first increment probe (here 10 dB SPL). The

p.d.f for each probe was calculated based on whether the participant detected the probe or not,

within a1000ms window following probe onset. After each observation, Bayes' Theorem is used

to generate a new p.d.f. given the listener's preceding responses, and the mean of each newly

calculated p.d.f. determines the difference level of the next increment. The mean of the p.d.f.

calculated after the final observation constitutes the final threshold estimate. We employed a

MAPPING DYNAMIC TEMPORAL ATTENTION

stopping criterion of 20 observations for each probed temporal position. Separate thresholds were tracked concurrently for each of the four stimulus probe positions. A stimulus block terminated once thresholds converged for all four probe positions.

Procedure. Participants were instructed to listen to the looping rhythmic patterns and to press the space bar on a computer keyboard when they heard brief increases in volume. Participants had until 1000 ms after each intensity deviant to respond. However, participants were unaware of the 1000 ms response window and were instead instructed to respond as quickly and as accurately as possible following each perceived increment. They first performed a training session in which they detected intensity increments with a one-minute, three-timbre percussion loop that was similar to the experiment stimuli (see online supplemental materials for an audio sample of the training stimulus). Next, participants performed the experiment phase of the increment detection task across five blocks, each block containing a continuous loop that was randomly selected from the five stimulus patterns (Figure 1). Stimulus blocks lasted an average of 6.49 minutes, and participants took breaks of self-selected durations between blocks. The total time of the experiment was approximately 32 minutes, plus the participants' individually determined break lengths. The session ended with questionnaires in which participants reported their levels of musical training and the specific nature of their musical experience. Questionnaire forms were administered using Ensemble (Tomic & Janata, 2007), a web interface for behavioral experiments.

Results and Discussion

We tested our hypothesis that resonator level predicts increment detection thresholds using a linear mixed-effects model (LMM). We used an LMM for our within-subjects design due

MAPPING DYNAMIC TEMPORAL ATTENTION

to its abilities to accommodate interdependence among repeated observations and its ability to

model participants as sources of random variation (Pinheiro & Bates, 2009). We fit the following

LMM to the data using the nlme package (Pinheiro, Bates, DebRoy, Sarkar, & the R

Development Core Team, 2013) in R (R Core Team, 2013):

$$\text{Threshold}_{ij} = \beta_0 + \beta_1 \text{Resonator}_i + \beta_2 \log_{10}(\text{Block}_i) + b_i + \varepsilon_{ij}$$

Threshold_{ij} is the dependent variable measured at resonator level j within participant i. β_0 is the

fixed-effect intercept, and β_1 and β_2 are fixed effects of resonator level and $\log_{10}$(block number),

respectively. b_i is the random effect of threshold intercept for participant i, and ε_{ij} represents

residual error. Model parameters were estimated using the Restricted Maximum Likelihood

(REML) method. Effect sizes of fixed effects were evaluated using Cohen's f^2, a standardized

measure of an independent variable's effect size in the context of a multivariate model (Cohen,

1988). We calculated f^2 effect sizes following the guidelines of Selya, Rose, Dierker, Hedeker, &

Mermelstein (2012) for mixed-effects multiple regression models. We did not perform an *a*

priori power analysis, as this was our first full data set from this paradigm beyond small pilot

samples.

   Resonator level significantly predicted increment detection thresholds [ß = -2.67, *SE* =

0.32, $t(530) = -8.48$, $p < 0.001$, $f^2 = 0.13$], such that higher resonator amplitude corresponded to

lower thresholds (Figure 3A). Furthermore, thresholds increased across blocks in a logarithmic

pattern, such that they increased more over the first half of the experiment and plateaued in the

last half of the experiment. Thus, the $\log_{10}$ of block number significantly predicted increment

detection thresholds [ß = 4.79, *SE* = 0.65, $t(530) = 7.41$, $p < 0.001$, $f^2 = 0.10$].

MAPPING DYNAMIC TEMPORAL ATTENTION

The results suggest that output from Tomic and Janata's (2008) resonator model is a suitable predictor of salience dynamics in complex, multi-timbral rhythmic patterns. Although the aggregate data show considerable variation in detection thresholds (Figure 3A), the negative relationship between resonator level and increment detection threshold was visible within most participants' data (Figure 4). Participant fatigue likely contributed to variability in task performance, as indicated by the positive relationship between block number and threshold. Moreover, many subjects stated after the experiment that they had difficulty sustaining attention throughout the session.

Beyond these observations, one question is whether the enhanced processing that results from rhythmically driven temporal attention occurs similarly across deviance types. Results of previous studies have suggested that perception of increments vs. decrements in texture (Huron 1990, 1992) and IOI duration (Repp 1998, 1999) are asymmetric. With regard to the present study it is possible, on the one hand, that deviance detection is facilitated at salient time points, regardless of the type of deviant that is probed. Alternatively, to the extent that intensity increments are perceptually analogous to intensity accents, it is possible that increments may alter a listener's perception of metric structure. Intensity increments at metrically salient time points may also coincide with a listener's expectation of accentuation, and may thus be registered as metrically appropriate, and thus non-deviant accents. If amplitude increments do interfere with perceived metric structure, one control for this would be to test deviance detection for intensity decrements. Thus, in Experiment 2 we tested decrement detection using the same paradigm as in Experiment 1.

MAPPING DYNAMIC TEMPORAL ATTENTION

Experiment 2

As noted above, it is possible that transient intensity increments, like those used in Experiment 1, perturb a listener's mental representation of metric structure by altering the perceived accent pattern. In Experiment 2, we addressed these concerns by modifying the paradigm from Experiment 1, such that participants detected intensity decrements rather than increments. We hypothesized that, similarly to Experiment 1, participants would be more sensitive to intensity decrements – thus showing lower absolute-value thresholds – at time points of greater modeled salience.

Method

Thirty-eight UC Davis undergraduate students, age 18 to 24 ($M = 20.7$), participated in exchange for course credit. Thirteen participants had at least three years of musical training, with a mean of 7.2 years of training among that group. For all experiments reported in this paper, no individual participated in more than one experiment.

Stimulus rhythm patterns and the temporal position of probes were identical to those in Experiment 1 (Figure 1). However, participants detected intensity decrements rather than increments for both the training stimuli and the main multi-timbre detection task. Decrement magnitudes were controlled adaptively using the threshold procedures described in the *Adaptive Threshold* section of Experiment 1, in this case with a starting amplitude decrease of -10 dB SPL. As participants detected reductions in intensity, thresholds were estimated as negative values. However, in our analyses we present thresholds as absolute values to facilitate comparison across experiments. All other apparatus items and procedures were identical to Experiment 1.

MAPPING DYNAMIC TEMPORAL ATTENTION

Results and Discussion

We fit a mixed-effects model to the data with absolute-value threshold as the dependent

measure; resonator level and $\log_{10}$(block number) as fixed-effect predictors; and participant

threshold intercept as a random effect. As in Experiment 1, higher resonator levels significantly

predicted lower absolute-value thresholds [ß = -5.83, *SE* = 0.35, *t* (720) = -16.86, *p* < 0.001, f^2 =

0.39] (Figure 3B). Similarly to Experiment 1, $\log_{10}$(block number) predicted thresholds, such that

absolute-value thresholds increased logarithmically as blocks progressed [ß = 3.49, *SE* = 0.67,

t(720) = 5.19, *p* < 0.001, f^2 = 0.04].

Experiment 2 results replicated those from Experiment 1, indicating that resonator model

output predicts attentional dynamics in rhythmic patterns across probe types. These observations

counter the possibility that Experiment 1 effects were increment-specific, or that increment

probes fundamentally perturb meter perception. Detecting decrements appeared to be more

difficult than detecting increments, as suggested by higher average absolute-value thresholds for

many stimulus probe positions (Figure 3). This is unsurprising given evidence that increasing

intensity in auditory objects is more salient than decreasing intensity (Neuhoff, 1998; Patterson,

1994; Phillips, Hall, & Boehnke, 2002). Moreover, stimulus intensity decrements that followed

large decrement thresholds (e.g., -18 dB SPL) were likely perceived as momentary silence or

removed acoustic events. Evidence from Cervantes Constantino, Pinggera, Paranamana,

Kashino, & Chait (2012) and Huron (1990; 1992) indicates that removed events within an

auditory scene are more difficult to detect than inserted events. Lastly, the present results further

suggest that the duration and number of blocks contribute to participant fatigue, as manifested by

progressively poorer performance over time. In the following experiments, we addressed this

MAPPING DYNAMIC TEMPORAL ATTENTION

issue in an effort to stabilize participants' performance over the course of a given experimental

session.

Experiment 3

In Experiment 3, we repeated the increment and decrement detection tasks from

Experiments 1 and 2 with several critical methodological changes. First, we sought to understand

the inter-individual variability in deviance detection thresholds observed for the multi-timbral

loops. We expected that participants would differ in their ability to perform intensity deviance

detection under optimal conditions. Thus, we obtained participants' baseline deviance detection

thresholds in isochronous single-instrument contexts, as such contexts isolated deviance

detection from temporal and auditory scene complexities present in the multi-timbral stimuli.

Second, given the evidence in Experiments 1 and 2 that task duration negatively impacted

participants' performance, we sought to reduce within-participant variability and task fatigue by

adjusting the threshold estimation procedures to shorten stimulus blocks.

We also tested whether individuals' aptitudes in an aspect of musical temporal processing

– namely beat detection – helped explain inter-participant variability in the multi-timbral

detection task. Although our task did not involve judgments about the music's beat, we suspected

that differences in participants' abilities to identify the beat (a repeating isochronous interval) in

excerpts of standard recorded music might correlate with differences in their abilities to track the

timing of overall rhythmic structure, which in turn could affect performance in our detection

task. As such, we measured beat-processing thresholds in a custom, adaptive threshold

modification of the Beat Alignment Test (BAT; Iversen & Patel, 2008), an assessment designed

MAPPING DYNAMIC TEMPORAL ATTENTION

to detect individual differences and impairments in musical beat perception. We also compared

task performance and the resonator model's predictive ability between metrically complex

stimuli used in Experiments 1 and 2 and new stimuli that had a simpler metric structure.

Method

   Participants.   We collected data on two samples of undergraduate students at UC Davis

in exchange for course credit. Twenty-five participants aged 18 to 27 (M = 20.8) performed the

increment detection task. Fifteen of these participants had at least three years of musical training

(mean years of training within these participants = 8.9) A separate group of 29 participants aged

18 to 23 (M = 20.5) performed the decrement detection task. Thirteen of these participants were

musically trained for at least three years (mean years of training within these participants = 8.2).

All participants reported having normal hearing.

   Adaptive threshold.   Threshold estimation in Experiments 1 and 2 employed a twenty-

observation stopping criterion. Alternatively, ZEST can be set to terminate dynamically once a

critical p.d.f. variance is reached. This approach better exploits ZEST's threshold convergence

speed, rather than stopping after a fixed number of observations. Thus, in Experiment 3 we

utilized a dynamic stopping criterion to reduce the duration of stimulus blocks. Following pilot

testing, we considered a p.d.f. standard deviation (SD) of 1.10 dB SPL to indicate threshold

convergence. Once this stopping criterion was met for a given probe position, the mean of the

convergent p.d.f. served as the participant's final threshold estimate for that probe and

increments were no longer presented at that position. The block terminated once thresholds

converged for all four probe positions. The threshold for a probe position terminated after 20

observations if it had not converged by then. Pilot testing of the decrement detection task

revealed that, when subjects began the task with no responses to targets, thresholds converged within three observations due to a lack of response variability and thus a hastened narrowing of the p.d.f. We circumvented this premature convergence in the decrement task by requiring a minimum of seven observations before each threshold could converge. All other aspects of threshold estimation were identical to the previous experiments.

   Stimuli and Procedures. Stimuli consisted of two metrically complex and two metrically simple stimuli that looped at a tempo of 107 BPM. The stimuli from Experiments 1 and 2 named "complex 4" and "complex 5" were the complex stimuli. We created two additional relatively simple stimuli using the same composition methods described in Experiment 1 (Figure 1). Firstly, the simple stimuli differed from the complex stimuli in that IOIs between stimulus pattern events were never shorter than an eighth note (1/2 beat), whereas the complex stimuli contained IOIs equal to a sixteenth note (1/4 beat). Secondly, the underlying beat was emphasized in the simple stimuli by ensuring that acoustic events occupied every beat and ½ beat position, whereas some complex stimuli lacked acoustic events on beat positions or salient beat subdivisions. Thirdly, in both simple and complex stimuli the lowest voice (conga) contained the same number of onsets (four), while the two higher voices contained eight events in the complex stimuli and only four in the simple. We acknowledge the potential interaction between instrument timbre and metric salience level. Hence, whereas in the complex stimuli there is an unequal distribution of event onsets across voices, in the simple stimuli, all instruments have four event onsets. We chose four events because we thought it important to keep the number of conga events the same across conditions due to the importance of low frequency information in driving groove perception (Stupacher, Hove, & Janata, 2016), movement induction (Burger, Thompson, Luck, Saarikallio, & Toiviainen, 2012), and vestibular

MAPPING DYNAMIC TEMPORAL ATTENTION

activity (Todd, 2001). Ultimately, we view our simple stimuli as simple because of their reduced and aligned rhythmic activity, which should result in more easily perceived metric structure.

Experiment procedures were identical to those in Experiments 1 and 2, with the exception that the target intensity changes lasted approximately 140 ms, whereas intensity changes in the previous experiments lasted approximately 200 ms. The revised intensity deviant duration matched the duration of a sixteenth note interval, the smallest IOI present in the stimuli.

Prior to the deviance detection task with the multi-timbral stimuli, we assessed participants' baseline deviance detection thresholds using single-timbre, isochronous sequences. The initial baseline task also doubled as a training session for participants. Participants heard the three single-timbre isochronous sequences (snap, shaker, or conga) in random order. During the task, the single instrument sample repeated at a constant IOI of 561 ms (107 bpm). Transient increments or decrements (depending on the participant's group) were presented following the same rules as in the main experiment. Each isochronous sequence continued playing until the participant's threshold for that timbre adaptively converged. A single threshold was estimated for each timbre, and the mean of those thresholds served as participants' baseline deviance detection threshold.

Following the multi-timbral deviance detection task, participants performed a custom, adaptive threshold modification of the perceptual portion of the BAT (*Part 3: Perceptual judgment of the beat*; Iversen & Patel, 2008). This portion of the BAT presents participants with a beeping metronome superimposed on excerpts of commercially recorded music and tests participants' ability to detect perturbations in the phase and period of the metronome with respect to the phase and period of the music's beat. The metronome was an isochronous 1000-Hz pure tone with a rise and fall time of 5 ms and a duration of 100 ms. See Table 1 in Iversen and Patel

(2008) for a complete list of stimuli used in the BAT. In our modified version, we estimated participants' thresholds for detecting phase and period metronome perturbations following the procedures described in *Adaptive threshold*. For each participant, a single threshold was calculated for each perturbation type (phase and period) by presenting the BAT stimuli in random order and adaptively adjusting the metronome perturbations on each trial using the ZEST algorithm until the threshold converged. Given evidence that detecting phase perturbations in the BAT is more difficult than detecting period perturbations (Iversen & Patel, 2008), as well our pilot observations that confirmed this, the period block began the threshold algorithm with a 10% perturbation and the phase block began with a 30% perturbation. Both tests had a catch-trial (no perturbation) probability of .20, and thresholds were set to converge when the SD of the ZEST p.d.f. reached 1.15%. In other words, as with the main task in the present experiment, which also used a similar dynamic threshold-stopping criterion, the number of trials varied among participants depending on the trial-to-trial consistency of their task performance. We implemented our adaptive-threshold version of the BAT using the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) in MATLAB[1].

Results

We fit mixed-effects models to the increment and decrement data sets with threshold as the dependent measure (absolute-value threshold for the decrement data). We included resonator level, isochronous baseline threshold, BAT (tempo and phase test) thresholds, and stimulus complexity as fixed effects. Participants' threshold intercepts were modeled as random effects.

A separate mixed-effects model on the single-timbre isochronous thresholds [fixed effect: instrument (mid-conga, shaker, or snap); random intercept: participant] indicated no difference in

MAPPING DYNAMIC TEMPORAL ATTENTION

baseline increment thresholds among the isochronous instruments, $F(2, 46) = 1.20$, $p = 0.310$, f^2 = 0.03. Isochronous decrement thresholds did differ significantly among the instrument timbres, $F(2, 56) = 7.18$, $p = 0.002$, $f^2 = 0.23$. A *post hoc* comparison (*p*-values corrected for false discovery rate; Benjamini & Hochberg, 1995) indicated that decrement thresholds were higher for the snap than those for the shaker [$t(56) = 3.78$, $SE = 0.44$, $p = 0.001$], but no other timbre comparisons differed significantly. However, we intended these single-timbre isochronous thresholds to be baseline metrics of each participant's increment or decrement thresholds. As such, to restrict the number of parameters in the mixed-effects models for the experiment data, the isochronous baseline model parameter referred to in the previous paragraph and reported in the results below is the mean of each participant's single-timbre isochronous thresholds.

Lastly, we used Welch's t-tests to compare thresholds and threshold variability between experiments.

Increment detection. Compared to the average duration of blocks in Experiments 1 and 2 (M = 6.49 minutes), the length of task blocks was substantially reduced (M = 3.61 minutes) under our dynamic threshold convergence procedures. Whereas Experiments 1 and 2 required 20 observations for thresholds to converge, the average number of observations required for thresholds to converge in the current increment detection experiment was approximately 11.

As observed in Experiments 1 and 2, higher resonator levels significantly predicted lower increment detection thresholds [ß = -5.68, $SE = 0.61$, $t(356) = -9.27$, $p < 0.001$, $f^2 = 0.30$] (Figure 5A). Performance on both the tempo and phase portions of the BAT failed to predict thresholds [tempo: ß = 0.09, $SE = 0.08$, $t(16) = 1.14$, $p = 0.272$, $f^2 = 0.00$; phase: ß = -0.06, $SE = 0.05$, $t(16) = -1.22$, $p = 0.242$, $f^2 = 0.01$]. Baseline isochronous thresholds also did not predict thresholds in

MAPPING DYNAMIC TEMPORAL ATTENTION

the multi-timbre task [ß = 0.08, SE = 0.29, $t(22)$ = 0.28, p = 0.780, f^2 = 0.02], nor did block

number [ß = 0.15, SE = 0.13, $t(356)$ = 1.15, p = 0.252, f^2 = 0.00]. The $\log_{10}$ of block number was

also not significant [ß = 0.94, SE = 0.62, $t(356)$ = 1.50, p = 0.134, f^2 = 0.01]. Thresholds for

complex stimuli were significantly lower than those for simple stimuli, but only slightly so [ß = -

0.94, SE = 0.45, $t(356)$ = -2.11, p = 0.036, f^2 = 0.01]. Moreover, the relationship between

resonator level and thresholds differed between complex and simple stimuli such that larger

reson values were more closely related to smaller thresholds (i.e. the negative resonator ~

threshold slope was steeper) for complex than for simple stimuli [ß = 3.23, SE = 0.71, $t(356)$ =

4.58, p < 0.001, f^2 = 0.06] (Figure 6A). On average, the thresholds for Experiment 3 (M = 8.82)

were lower than those of Experiment 1 (M = 10.43), but a two-tailed Welch's t-test indicated that

this difference only approached significance [$t(39.22)$ = -1.94, p = 0.059]. Increment thresholds

were significantly less variable in Experiment 3 (mean subject SD = 3.01) than in Experiment 1

(mean subject SD = 4.01) [$t(50.62)$ = -3.25, p = 0.002].

Decrement detection.   Mean block duration (3.07 minutes) was again considerably

reduced compared to Experiments 1 and 2 (M = 6.49 minutes). On average, thresholds required

approximately 9 observations to converge.

Higher resonator level predicted lower absolute-value decrement detection thresholds [ß

= -8.83, SE = 1.53, $t(434)$ = -5.79, p < 0.001, f^2 = 0.08] (Figure 5B). Similarly to the increment

data, BAT performance did not predict thresholds [tempo: ß = 0.08, SE = 0.05, $t(22)$ = 1.75, p =

0.095, f^2 = 0.00; phase: ß = 0.02, SE = 0.02, $t(22)$ = 0.80, p = 0.432, f^2 = 0.00]. Participants'

isochronous baseline thresholds, however, were positively correlated with participants'

experiment thresholds [ß = 0.55, SE = 0.23, $t(27)$ = 2.34, p = 0.027, f^2 = 0.01] (Figure 6C).

Additionally, baseline thresholds interacted with resonator activity, such that the degree to which reson amplitude explained participants' thresholds correlated positively with their baseline isochronous thresholds [ß = 0.61, *SE* = 0.28, *t*(434) = 2.20, *p* = 0.028, f^2 = 0.01] (Figure 6B). Unlike the increment group, the number of elapsed blocks was also associated with increased thresholds [ß = 0.70, *SE* = 0.17, *t*(434) = 4.03, *p* = 0.001, f^2 = 0.04]. Stimulus complexity significantly predicted task performance such that simpler stimuli led to lower thresholds [ß = -1.81, *SE* = 0.61, *t*(434) = -2.98, *p* = 0.003, f^2 = 0.02] (Figure 6A). However, complexity did not interact with resonator level [ß = 1.04, *SE* = 0.96, *t*(434) = 1.09, *p* = 0.276, f^2 = 0.00]. Average decrement thresholds did not differ significantly between Experiments 2 and 3 [*t*(63.95) = 1.10, *p* = 0.275], and neither did decrement threshold variability [*t*(60.29) = -1.48, *p* = 0.144].

Discussion. As Experiment 3 replicates the results of Experiments 1 and 2, even while controlling for stimulus- and participant-level covariates, we demonstrate that our resonator model reliably predicts temporal attention in rhythmic musical patterns. Participants' increment detection thresholds were less variable and somewhat lower in this experiment than in Experiment 1, and unlike Experiment 1, participants' increment thresholds did not fluctuate throughout the experiment. This is likely due to our changes in the threshold convergence criterion, which shortened block duration to less than half that of Experiments 1 and 2. As such, our procedural adjustments appear to have improved the precision with which we could estimate participants' thresholds. However, average decrement detection thresholds and their variability did not change between Experiments 2 and 3. That block number predicted a growth in decrement but not increment thresholds over time suggests that decrement detection in our task imposed an increased cognitive load that accumulated over time and was not ameliorated by

shortening the task. Such asymmetry between increment and decrement detection is not surprising given previous research suggesting that detecting decrements is more difficult than detecting increments (Cervantes Constantino et al., 2012; Neuhoff, 1998; Patterson, 1994; Phillips et al., 2002).

Manipulating stimulus complexity and testing for baseline deviance detection thresholds helped to explain stimulus- and participant-level variations in the data. Our results show that stimuli with simpler metric complexity elicited lower thresholds on average, and that, in the case of increment detection, resonator activity was more tightly associated with task performance for metrically complex stimuli. The latter observation may reflect that individuals rely more heavily on dynamic attentional cues of rhythmic structure when temporal structure is more complex and when parsing event timing is more difficult. We discuss individual differences more extensively in the General Discussion.

Model Comparisons

The resonator model that we used to predict temporal attention is agnostic to principles of music theory. Given the almost ubiquitous use of a music-theoretic construct of a metric hierarchy, in which there are binary and ternary subdivisions of spans of time, for defining events and moments of interest in studies of rhythm perception and rhythmic dynamic attending (e.g., Abecasis et al., 2009; Bolger et al., 2014; Bouwer, Van Zuijen, & Honig, 2014; Ladinig, Honing, Háden, & Winkler, 2009; Palmer & Krumhansl, 1990; Snyder & Large, 2005), one may ask whether a simple metric hierarchy model explains our data as well as or better than our resonator model. To answer this question, we modeled a metric hierarchy's predictions of the

threshold data from each experiment. We implemented the metrical hierarchy model by coding temporal locations within each stimulus using saliency values of 1 (least salient) to 4 (most salient) according to music-theoretic predictions in a 4/4 meter [see "theoretic predictions" in the 4/4 stimulus in Figure 1 of Palmer and Krumhansl (1990) for an example schematic]. This coding scheme, in both Palmer and Krumhansl (1990) and therefore our paper, is based on the work of Lerdahl and Jackendoff (1983).

Another possible explanation of our experimental results is that stimulus events were perhaps more salient when multiple instruments sounded simultaneously, leading to lower thresholds at those corresponding times. This would imply that increased stimulus energy, rather than periodicity structure, predicted temporal saliency. We tested this "auditory scene density" alternative by coding each deviant time point with the number of instrument events playing concurrently.

Results and Discussion

We tested each of the three alternatives (resonator, metric hierarchy, and auditory scene density) with mixed-effects models, using threshold as the dependent variable, the given model's predicted saliency of probe times as the fixed effect, and random-effect intercepts for each participant. We fit a separate set of these three models to the increment detection data (Experiment 1 and the increment sample from Experiment 3) and the decrement detection data (Experiment 2 and the decrement sample from Experiment 3). In order to assess whether the minor adjustments in threshold procedures between experiments led to significant differences in threshold outcomes or affected marginal music-model predictor effects, we also included a fixed-effect parameter for experiment number (Experiment 1 versus 3 for increment thresholds and

MAPPING DYNAMIC TEMPORAL ATTENTION

Experiment 2 versus 3 for decrement thresholds) in each mixed-effects model. We assessed the

relative goodness of model fit using the Akaike Information Criterion (AIC) (Akaike, 1974) and

Bayes' Information Criterion (BIC) (Schwarz, 1978). We calculated effect sizes of fixed-effects

using Cohen's f^2.

As Table 1 indicates, our resonator model yielded the lowest AIC and BIC values for

both increment and decrement threshold data, indicating that the resonator model is preferred

above the metric hierarchy and scene density models, given these common metrics of model fit.

Furthermore, although resonator, metric hierarchy, and auditory scene density models were all

significant predictors of participant thresholds, the effect size was markedly larger for the reson

model than for the alternatives. However, we note that no broadly accepted significance test

exists for comparing non-nested mixed-effects models (i.e. each model contains a different fixed-

effect term). As such, we caution that a strong claim of model-fit superiority would require

further testing of alternative models. Nevertheless, these results suggest that the resonator model

better explains variance in participants' thresholds than a common music-theoretic model of

metric saliency or by the momentary density of the acoustic scene. Moreover, these observations

indicate the utility of predicting rhythmic attention with a model driven directly by the temporal

structure of individual stimuli rather than one that employs a canonical saliency map, i.e. metric

hierarchy, that is somewhat removed from the specific temporal dynamics of individual stimuli.

General Discussion

This study assessed the ability of resonator output from Tomic and Janata's (2008) model

to predict temporal attention in rhythmic musical patterns. We tested this using a psychophysical

procedure for mapping temporal attention, in which intensity deviants were presented at times

MAPPING DYNAMIC TEMPORAL ATTENTION

corresponding to various levels of modeled saliency and separate detection thresholds were measured for each probed position. As we hypothesized, our resonator model significantly predicted task performance, whereby participants were more sensitive to probes at moments of higher resonator-predicted salience. This result was highly reliable, as the effect of resonator prediction replicated across four samples of participants, both increment and decrement probes, and stimuli with varying degrees of rhythmic complexity.

Performance on our task varied markedly across participants, as Figure 6A indicates. However, assessments of rhythmic aptitude and psychophysical metrics were not reliable predictors of this inter-individual variability. Viewed from the perspective of DAT, if an attending rhythm oscillates with the phase and period of the isochronous sequence, participants always received deviant probes at times of maximal attentional focus, with observed thresholds in a relatively narrow range of approximately 2-8 dB SPL. The question therefore becomes whether intensity deviance detection under optimal temporal attention conditions is predictive of intensity deviance detection in a more complex musical scene in which multi-timbral stimuli encompass multiple periodicities, and deviant probes are presented at multiple levels of oscillator-modeled temporal saliency. Overall, we found that in the case of the multi-timbral patterns, the range of thresholds was much broader (approximately 2-20 dB SPL) and the average thresholds were larger, indicating a cost of musical scene complexity. Surprisingly, however, individuals' thresholds in the isochronous condition were only related to thresholds in the musical, multi-timbral scenes for the more challenging intensity decrement case and unrelated in the less-challenging intensity increment conditions.[2] However, it was the case that those individuals who performed best in the isochronous condition also showed the greatest sensitivity to probe position location (magnitude of aggregated oscillator output) when intensity

decrements had to be detected. In other words, these participants exhibited the strongest effects

of graded temporal attention, meaning that the difference in their thresholds between moments of

greater predicted salience and moments of lesser predicted salience were larger than for those

participants who exhibited larger thresholds in the isochronous task. Such temporal position

effects are difficult to explain in terms of differences in sensitivity to loudness changes alone.

It is notable that individuals' BAT performance did not correlate with performance on

our task. We surmised that perhaps the correlation between BAT performance and performance

on our intensity detection task differed for earlier and later portions of an experimental session,

given the differences we observed in intensity deviance detection as blocks progressed.

However, separate analyses on data from the first block and the final block of the intensity

deviant task still did not indicate any associations between performance on our task and

performance on the BAT (data not shown). The more likely explanation for the lack of

correlation may be that, despite the superficial similarity of both tasks pertaining to temporal

perception, our intensity deviant task and the BAT rely on distinct sets of psychological

processes. To perform the BAT, participants must solve the auditory scene analysis problem of

comparing timing judgments about an isochronous stream (the metronome) that is clearly

segregated from the musical piece (Bregman, 1990), with timing judgments about the percept of

the beat that is extracted from a multi-instrumental aggregate. This type of comparison likely

requires participants to engage in what Keller and colleagues term *prioritized integrative*

attending, a dual task in which the listener divides attention between tracking the metronome

stream and grouping together the elements of the aggregate musical scene (Keller & Burnham,

2005; Uhlig, Fairhurst, & Keller, 2013). Conversely, aside from the obvious difference between

detecting intensity changes and detecting timing perturbations, the demands of our task are quite

different from those of BAT. Although both tasks present complex musical auditory scenes,

stream segregation and reintegration of timing percepts of two distinct components of the

auditory scene are not required for our task because we modulate the entire scene at critical time

points. Still, in our study participants might selectively bias attention to some instruments over

others or may switch attention back and forth among the various instruments. Future

modifications of our task should probe individual instrument streams within various attentional

contexts to examine interactions between selective attention to auditory scene objects and DAT.

Overall, our results corroborate previous research in support of DAT's central tenet, that

the attentional process underlying rhythm processing follows an oscillatory pattern (e.g., Bolger

et al., 2013; Jones & Boltz, 1989; Jones et al., 2002; e.g., Large & Jones, 1999; McAuley &

Fromboluti, 2014; Miller et al., 2013). Unlike the oscillatory models developed by Large and

colleagues (Large & Jones, 1999; Large & Kolen, 1994; Large & Palmer, 2002; Large & Snyder,

2009), which take a nonlinear dynamical systems approach, the resonator model is relatively

computationally simple in that it consists of a system of linear oscillators [see Tomic and Janata

(2008) for computational details]. The model does not rely on inhibition or preset integer-related

tunings between layers of oscillators to model metric relationships; it is a purely stimulus-driven,

feed-forward approach to predicting salient periodicities and attentional fluctuations over time.

The model is also unique in its estimation of a continuous temporal salience map, which allows

for probing moment-to-moment attentional dynamics. We note, however, that the resonator

model in its current form should not be regarded as a complete model of either musical rhythm

perception or temporal attention mechanisms. Nevertheless, the current study sheds light on the

ability of a feedforward oscillator model with minimal assumptions to predict temporal attending

behavior. The model thus serves as a benchmark against which other computational models of DAT can be compared when using a similar behavioral paradigm.

Our approach for testing the loci of dynamic attention differs from the majority of previous research using music, most of which assumes that metrically strong locations are more salient than metrically weak locations. Instead, the model we used derives temporal salience predictions directly from a stimulus' periodicity structure without regard to music theoretic notions, such as strong and weak beats. Strong beats may often be highly salient; parallels clearly exist between predictions of the metric hierarchy model and those of the oscillator model (Tomic & Janata, 2008). Similarly, there is a tendency for multiple instruments to play at metrically salient moments, i.e. for the distribution of note densities to be temporally clustered. Nevertheless, we argue that the model should perform better when faced with rhythmic stimuli in which salience dynamics do not closely follow auditory scene density, a strong- and weak-beat accent structure (e.g., syncopated rhythms), or common Western metric structure. Although our results support that prediction, a conclusive statement would require directly comparing model performance on stimuli manipulated to contain or be free of syncopation, polyrhythmic structure, or some other deviation from strong- and weak-beat metrical structure. Additionally, future studies might consider using more complex music-theoretic models of meter, such as those of Hasty (1997), Temperley (2001, 2007) or London (2012), for comparison with oscillator models.

In the current study we developed a novel, musically oriented psychophysical task for assessing temporally graded perceptual sensitivity. Ethologically relevant acoustic stimuli, such as music and speech, often present as continuous streams. As our paradigm maps attention across multiple temporal locations within continuously playing stimuli, it more closely resembles a realistic auditory environment than paradigms that present short stimulus patterns across many

discrete trials. Moreover, unlike typical threshold procedures that require many observations for each threshold, our integration of a Bayesian threshold framework allows for estimating thresholds with many fewer observations. This is desirable for research with musical stimuli, in which too many repetitions, when probing multiple loci, can become unpleasant and even exhausting, as suggested by participant feedback and worsening performance for the longer sessions of Experiments 1 and 2. Thus we expect that our paradigm will be a useful research tool for experimenters who wish to investigate auditory processing and related phenomena with increased validity.

In conclusion, the present study offers a unique combination of attention mapping and computational modeling to investigate the temporal orienting of attention in rhythmic musical patterns. We present evidence that the loci of dynamic, rhythmic attention can be predicted from the output of reson filters. Our results are consistent with DAT and offer a new, alternative method for modeling the temporal orienting of attention.

References

Abecasis, D. (2005). Differential brain response to metrical accents in isochronous auditory

sequences. *Music Perception: An Interdisciplinary Journal*, *22*(3), 549–562.

https://doi.org/10.1525/mp.2005.22.3.549

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on

Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, *41*(3),

254–311. https://doi.org/10.1006/cogp.2000.0738

Bauer, A.-K. R., Jaeger, M., Thorne, J. D., Bendixen, A., & Debener, S. (2015). The auditory

dynamic attending theory revisited: A closer look at the pitch comparison task. *Brain

Research*, *1626*, 198–210. https://doi.org/10.1016/j.brainres.2015.04.032

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and

powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B

(Methodological)*, *57*(1), 289–300.

Bolger, D., Coull, J. T., & Schön, D. (2014). Metrical rhythm implicitly orients attention in time

as indexed by improved target detection and left inferior parietal activation. *Journal of

Cognitive Neuroscience*, *26*(3), 593–605. https://doi.org/10.1162/jocn_a_00511

Bolger, D., Trost, W., & Schön, D. (2013). Rhythm implicitly affects temporal orienting of

attention across modalities. *Acta Psychologica*, *142*(2), 238–244.

https://doi.org/10.1016/j.actpsy.2012.11.012

Bouwer, F. L., Van Zuijen, T. L., & Honing, H. (2014). Beat processing is pre-attentive for

metrically simple rhythms with clear accents: An ERP study. *PLoS ONE*, *9*(5), e97467.

https://doi.org/10.1371/journal.pone.0097467

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*.

Cambridge, MA: MIT Press.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our

internal clock: Direct brain evidence of subjective accents in isochronous sequences.

Psychological Science, *14*(4), 362–366.

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S., & Toiviainen, P. (2012). *Music moves

us: Beat-related musical features influence regularity of music-induced movement.* Paper

presented at the 12th International Conference in Music Perception and Cognition and the

8th Triennial Conference of the European Society for the Cognitive Sciences for Music,

Thessaloniki, Greece.

Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of

speech. *Neuropsychologia*, *50*(11), 2652–2658.

https://doi.org/10.1016/j.neuropsychologia.2012.07.018

Cervantes Constantino, F., Pinggera, L., Paranamana, S., Kashino, M., & Chait, M. (2012).

Detection of appearing and disappearing objects in complex acoustic scenes. *PLoS ONE,

7*(9), e46167. https://doi.org/10.1371/journal.pone.0046167

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY:

Routledge.

Essens, P. J., & Povel, D.-J. (1985). Metrical and nonmetrical representations of temporal

    patterns. *Perception & Psychophysics*, *37*(1), 1–7.

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*(1), 89–92. https://doi.org/10.1111/j.1749-6632.2009.04779.x

Hasty, C. (1997). *Meter as rhythm*. New York, NY: Oxford University Press.

Hurley, B. K., Martens, P. A., & Janata, P. (2014). Spontaneous sensorimotor coupling with

    multipart music. *Journal of Experimental Psychology: Human Perception and*

    Performance, *40*(4), 1679–1696. https://doi.org/10.1037/a0037154

Huron, D.B. (1990). Increment/decrement asymmetries in polyphonic sonorities. *Music*

    Perception, 7(4), 385-393.

Huron, D.B. (1992). The ramp archetype and the maintenance of passive auditory attention.

    Music Perception, 10(1), 83-91. DOI: 10.2307/40285540

Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge,

    MA: MIT Press.

Iversen, J. R., & Patel, A. D. (2008). The beat alignment test (BAT): Surveying beat processing

    abilities in the general population. In K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima,

    & M. Tsuzaki (Eds.), *Proceedings of the 10th International Conference on Music*

    Perception and Cognition (ICMPC 10) (pp. 465–468). Sapporo, Japan: Adelaide: Causal

    Productions.

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the

    psychology of the groove. *Journal of Experimental Psychology: General*, *141*(1), 54–75.

    https://doi.org/10.1037/a0024208

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459–491.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, *13*(4), 313–319.

Keller, P. E., & Burnham, D. K. (2005). Musical meter in attention to multipart rhythm. *Music Perception: An Interdisciplinary Journal*, *22*(4), 629–661. https://doi.org/10.1525/mp.2005.22.4.629

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research*, *34*(7), 885–912. https://doi.org/10.1016/0042-6989(94)90039-6

Klapuri, A. P., Eronen, A. J., & Astola, J. T. (2006). Analysis of the meter of acoustic musical signals. *IEEE Transactions on Audio, Speech, and Language Processing*, *14*(1), 342–355.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(14), 1–16.

Ladinig, O., Honing, H., Háden, G., & Winkler, I. (2009). Probing attentive and preattentive emergent meter in adult listeners without extensive music training. *Music Perception: An Interdisciplinary Journal*, *26*(4), 377–386. https://doi.org/10.1525/mp.2009.26.4.377

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, *320*(5872), 110–113. https://doi.org/10.1126/science.1154735

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for

time and pitch. *Brain and Cognition*, *69*(1), 127–137.

https://doi.org/10.1016/j.bandc.2008.06.004

Lange, K. (2010). Can a regular context induce temporal orienting to a target sound?

International Journal of Psychophysiology, *78*(3), 231–238.

https://doi.org/10.1016/j.ijpsycho.2010.08.003

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying

events. *Psychological Review*, *106*(1), 119–159.

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in

musical rhythm. Frontiers in Systems Neuroscience, 9, 159.

doi:10.3389/fnsys.2015.00159

Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection

Science*, *6*(2–3), 177–208.

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*,

26(1), 1–37.

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New

York Academy of Sciences*, *1169*(1), 46–57. https://doi.org/10.1111/j.1749-

6632.2009.04550.x

Large, E. W. (2010). Neurodynamics of music. In M. R. Jones, R. R. Fay, & A. N. Popper

(Eds.), *Music Perception* (pp. 201–231). New York, NY: Springer New York.

https://doi.org/10.1007/978-1-4419-6114-3_7

Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A Matlab toolbox for music information retrieval. In *Data Analysis, Machine Learning and Applications* (pp. 261–268). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-78246-9_31

Leman, M., Lesaffre, M., & Tanghe, K. (2001). An introduction to the IPEM toolbox for perception-based music analysis. In *Conference Program and Abstracts of SMPC 2001 Kingston, August 9-11, 2001*. Kingston, Ontario.

London, J. (2012). *Hearing in time: Psychological aspects of musical meter* (2nd edition). New York, NY: Oxford University Press.

Marvit, P., Florentine, M., & Buus, S. (2003). A comparison of psychophysical procedures for level-discrimination thresholds. *The Journal of the Acoustical Society of America*, *113*(6), 3348. https://doi.org/10.1121/1.1570445

McAuley, J. D., & Fromboluti, E. K. (2014). Attentional entrainment and perceived event duration. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1658), 20130401. https://doi.org/10.1098/rstb.2013.0401

Miller, J. E., Carlson, L. A., & McAuley, J. D. (2013). When what you hear influences when you see: Listening to an auditory rhythm influences the temporal allocation of visual attention. *Psychological Science*, *24*(1), 11–18. https://doi.org/10.1177/0956797612446707

Navarro Cebrian, A., & Janata, P. (2010a). Electrophysiological correlates of accurate mental image formation in auditory perception and imagery tasks. *Brain Research*, *1342*, 39–54. https://doi.org/10.1016/j.brainres.2010.04.026

Navarro Cebrian, A., & Janata, P. (2010b). Influences of multiple memory systems on auditory

mental image acuity. *The Journal of the Acoustical Society of America*, *127*(5), 3189–

3202. https://doi.org/10.1121/1.3372729

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*(6698), 123–124.

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. (2012). Selective neuronal entrainment to the beat and

meter embedded in a musical rhythm. *Journal of Neuroscience*, *32*(49), 17572–17581.

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

Otterbein, S., Abel, C., Heinemann, L. V., Kaiser, J., & Schmidt-Kassow, M. (2012). P3b

reflects periodicity in linguistic sequences. *PLoS ONE*, *7*(12), e51419.

https://doi.org/10.1371/journal.pone.0051419

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of

Experimental Psychology: Human Perception and Performance*, *16*(4), 728.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical

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

https://doi.org/10.2307/40285633

Patterson, R. D. (1994). The sound of a sinusoid: Spectral models. *The Journal of the Acoustical

Society of America*, *96*(3), 1409–1418. https://doi.org/10.1121/1.410285

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

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

Phillips, D. P., Hall, S. E., & Boehnke, S. E. (2002). Central auditory onset responses, and temporal asymmetries in auditory perception. *Hearing Research, 167*(1), 192–205.

Pinheiro, J. C., & Bates, D. (2009). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer-Verlag.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & the R Development Core Team. (2013). nlme: Linear and nonlinear mixed effects models (Version 3.1-113).

Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology. Human Perception and Performance, 16*(3), 564–573.

Potter, D. D., Fenwick, M., Abecasis, D., & Brochard, R. (2009). Perceiving rhythm where none exists: Event-related potential (ERP) correlates of subjective accenting. *Cortex, 45*(1), 103–109. https://doi.org/10.1016/j.cortex.2008.01.004

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/

Repp, B.H. (1998). The detectability of local deviations from a typical expressive timing pattern. *Music Perception, 15*(3), 265-289.

Repp, B.H. (1999). Detecting deviations from metronomic timing in music: Effects of perceptual structure on the mental timekeeper. *Perception & Psychophysics 61*(3), 529-548.

Repp, B. H. (2010). Do metrical accents create illusory phenomenal accents? *Attention, Perception & Psychophysics, 72*(5), 1390–1403. https://doi.org/10.3758/APP.72.5.1390

Sanabria, D., & Correa, Á. (2013). Electrophysiological evidence of temporal preparation driven by rhythms in audition. *Biological Psychology, 92*(2), 98–105. https://doi.org/10.1016/j.biopsycho.2012.11.012

MAPPING DYNAMIC TEMPORAL ATTENTION

Scheirer, E. D. (1998). Tempo and beat analysis of acoustic musical signals. *The Journal of the Acoustical Society of America*, *103*(1), 588–601.

Schmidt-Kassow, M., Schubotz, R. I., & Kotz, S. A. (2009). Attention and entrainment: P3b varies as a function of temporal predictability: *NeuroReport*, *20*(1), 31–36. https://doi.org/10.1097/WNR.0b013e32831b4287

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

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Current Opinion in Neurobiology*, *20*(2), 172–176. https://doi.org/10.1016/j.conb.2010.02.010

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biological Psychology*, *87*(1), 146–151. https://doi.org/10.1016/j.biopsycho.2011.02.021

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f^2, a measure of local effect size, from PROC MIXED. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00111

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, *24*(1), 117–126. https://doi.org/10.1016/j.cogbrainres.2004.12.014

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

Temperley, D. (2004). *The cognition of basic musical structures*. Cambridge, MA: MIT Press.

Todd, N. P. M. (1994). The auditory "Primal Sketch": A multiscale model of rhythmic grouping. *Journal of New Music Research*, *23*(1), 25–70. https://doi.org/10.1080/09298219408570647

Todd, N. (2001). Evidence for a behavioral significance of saccular acoustic sensitivity in humans. *The Journal of the Acoustical Society of America*, 110(1), 380-390. doi:10.1121/1.1373662

Tomic, S. T., & Janata, P. (2007). Ensemble: A web-based system for psychology survey and experiment management. *Behavior Research Methods*, *39*(3), 635–650.

Tomic, S. T., & Janata, P. (2008). Beyond the beat: Modeling metric structure in music and performance. *The Journal of the Acoustical Society of America*, *124*(6), 4024–4041. https://doi.org/10.1121/1.3006382

Uhlig, M., Fairhurst, M. T., & Keller, P. E. (2013). The importance of integration and top-down salience when listening to complex multi-part musical stimuli. *NeuroImage*, *77*, 52–61. https://doi.org/10.1016/j.neuroimage.2013.03.051

Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01111

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014).

    Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, *9*(4), e94446.

    https://doi.org/10.1371/journal.pone.0094446

Zanto, T. P., Large, E. W., Fuchs, A., & Kelso, J. S. (2005). Gamma-band responses to perturbed

    auditory sequences: evidence for synchronization of perceptual processes. *Music*

    Perception, *22*(3), 531–547.

Zanto, T. P., Snyder, J. S., & Large, E. W. (2006). Neural correlates of rhythmic expectancy.

    Advances in Cognitive Psychology, *2*(2–3), 221–231.

MAPPING DYNAMIC TEMPORAL ATTENTION

Footnotes

[1] The MAX patches for the dynamic attention mapping experiments and the MATLAB scripts for the adaptive version of the Beat Alignment Test are available on GitHub at https://github.com/janatalab/attmap.git and https://github.com/janatalab/adaptbat.git, respectively.

[2] It is possible that threshold variability for the isochronous stimuli is best explained by variability in basic sensory thresholds across participants, though we did not perform audiometric assessments to test this possibility.

MAPPING DYNAMIC TEMPORAL ATTENTION

Table 1

Alternative Predictors of Detection Thresholds, As Assessed by Linear Mixed-Effects Models

Threshold Type	Predictor		β	SE	*df*	f^2	AIC	BIC
Increment	Resonator		-2.81**	0.23	910	0.16	5315.94	5335.425
		Exp	-0.80	0.87	51	0.02		
	MH		-0.65**	0.11	910	0.04	5420.779	5440.264
		Exp	-1.40	0.87	51	0.01		
	Density		-1.39**	0.20	910	0.05	5409.263	5428.747
		Exp	-1.86*	0.87	51	0.01		
Decrement	Resonator		-5.70**	0.27	1160	0.38	7328.729	7349.182
		Exp	2.54**	0.64	65	0.05		
	MH		-1.49**	0.13	1160	0.11	7569.699	7590.152
		Exp	1.19	0.63	65	0.00		
	Density		-1.85**	0.25	1160	0.05	7633.993	7654.446
		Exp	0.35	0.63	65	0.00		

MAPPING DYNAMIC TEMPORAL ATTENTION

Note. Data were pooled across experiments by threshold type. Model estimates were obtained using linear mixed-effects models to regress fixed effects of temporal/stimulus model type and experiment number on threshold; participant intercept was included as a random effect. Threshold type "increment" = increment detection thresholds from Experiments 1 and 3. Threshold type "decrement" = decrement detection thresholds from Experiments 2 and 3. MH = metric hierarchy. Exp = experiment number (effect of Experiment 3 vs 1 on increment thresholds and Experiment 3 vs 2 on decrement thresholds when the reson, MH, or density were included as predictors; indention indicates the main predictor type exp was included with). SE = standard error. f^2 = Cohen's f^2 for effect size. AIC = Akaike's Information Criterion. BIC = Bayesian Information Criterion. For both AIC and BIC, a lower value indicates a more preferred model. * = $p < .05$. ** $p < .001$.

MAPPING DYNAMIC TEMPORAL ATTENTION
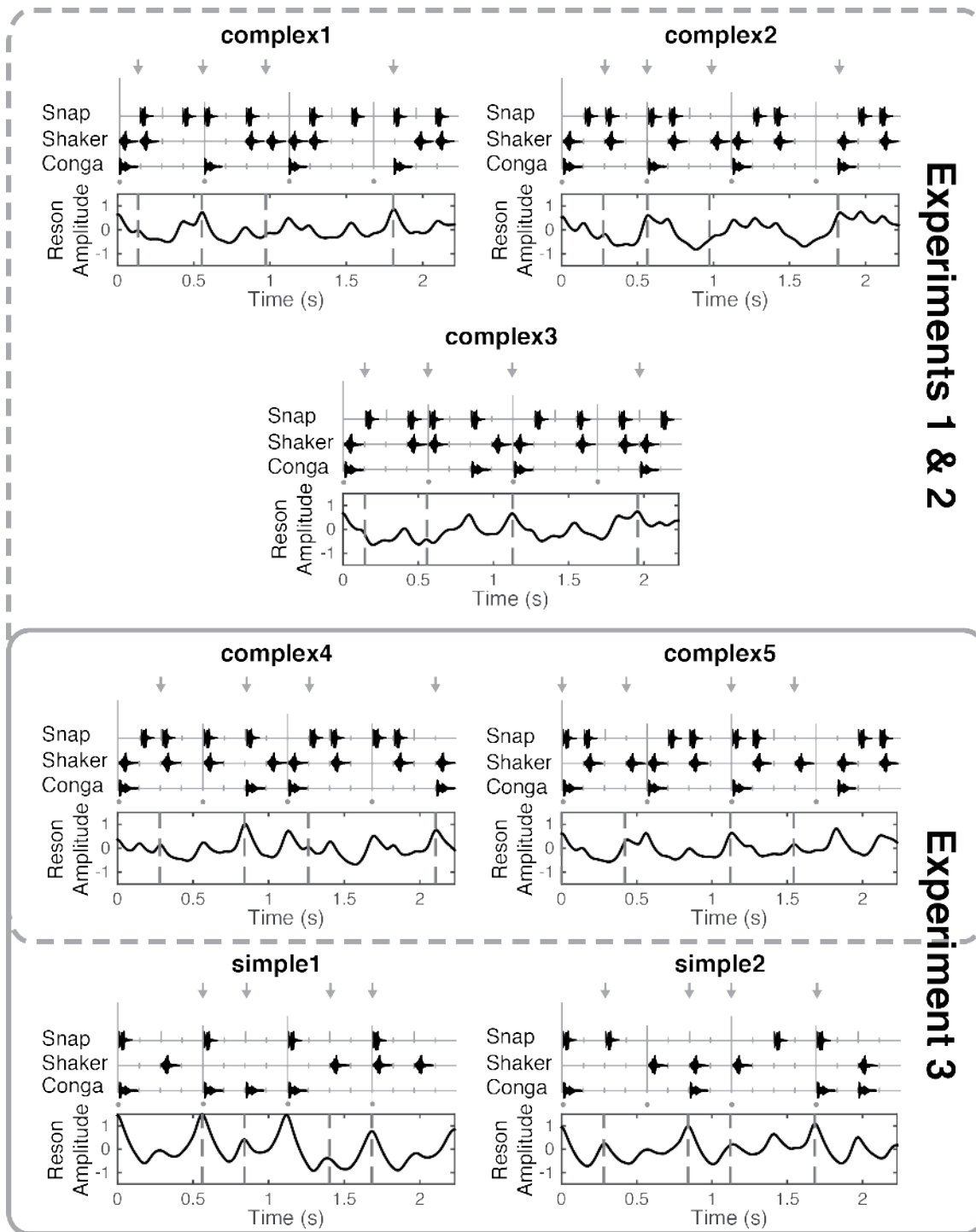
Figure 1. Stimulus patterns and corresponding reson-predicted salience time courses. Stimuli

were played as loops at 107 beats per minute. Tick marks underlying the stimulus patterns

correspond to time in sixteenth note steps (~140-ms intervals). Vertical tick length on the top

MAPPING DYNAMIC TEMPORAL ATTENTION

row of each pattern is scaled to relative theoretical salience levels under a canonical metric

hierarchy. Dots below patterns denote putative beat locations. The time series below each

stimulus pattern depicts the mean output from reson filters that are driven most strongly by the

stimulus' periodicities, as implemented with Tomic and Janata's (2008) resonator model.

Vertical dashed bars and downward pointing arrows denote the locations probed with changes in

intensity.

MAPPING DYNAMIC TEMPORAL ATTENTION
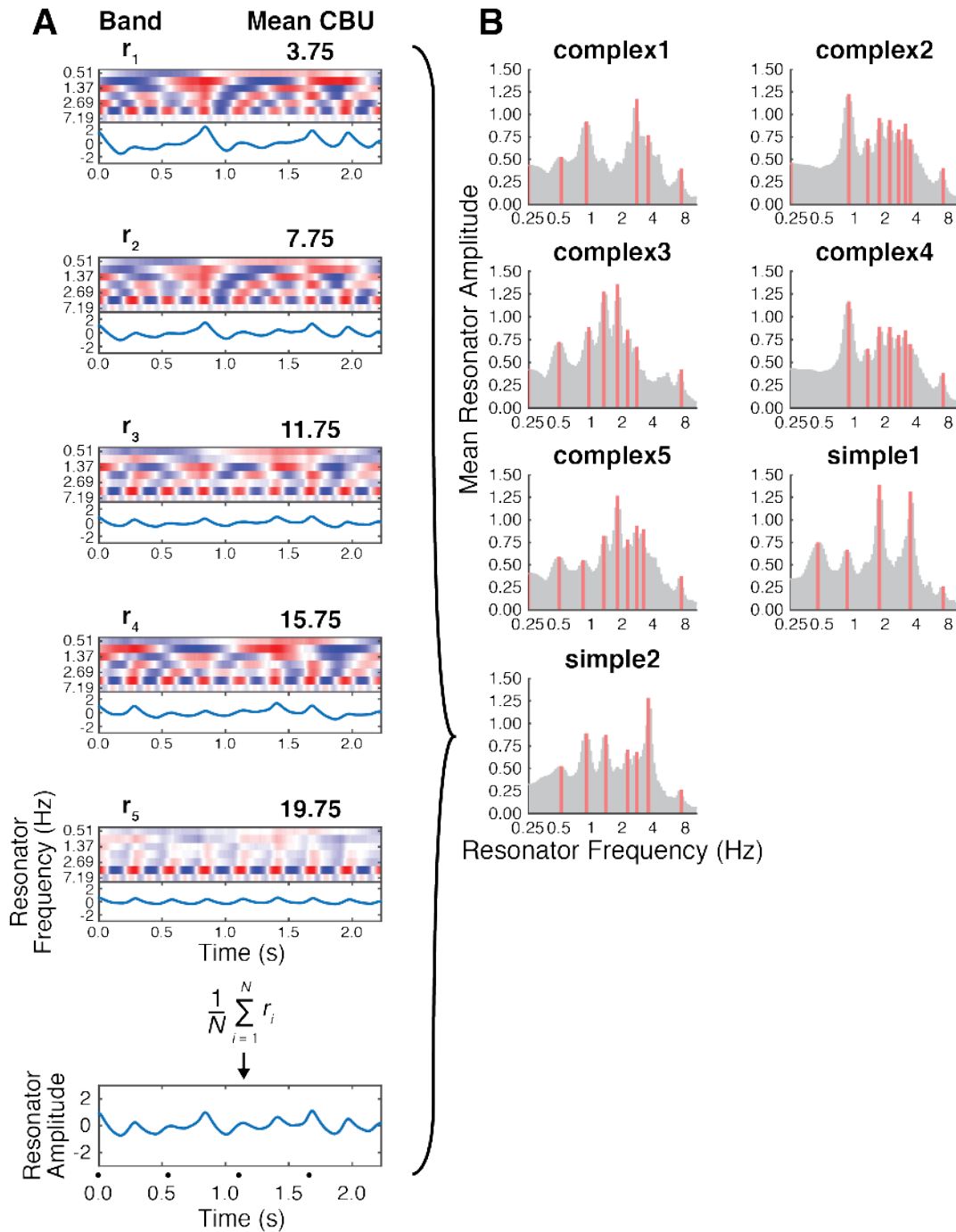
Figure 2. Model metrics for predicting temporal salience. (A) Output from reson filter banks (one filter bank for each of five spectral frequency bands) for the *simple2* stimulus. Values range from negative (blue) to positive (red). r_i = reson band index. CBU = critical band unit. Each band

MAPPING DYNAMIC TEMPORAL ATTENTION

depicts output from reson filters that correspond to the stimulus' prominent periodicities. The

time series below each band is the point-wise mean amplitude across all peak-periodicity reson

filters. The average of these r_n time series creates the peak-periodicity reson time series (bottom),

which we use to predict salience dynamics. For reference, dots below the abscissa of the bottom

time series denote the putative locations of beats. (B) Mean Periodicity Profiles (MPPs) for each

stimulus used in this paper. MPPs depict prominent periodicity frequencies by averaging the

energy (RMS) of reson filter output across time for 99 reson filters between 0.25 and 10 Hz. Red

bars denote peaks in the MPP, which are used to index peak-periodicity reson filters, as

illustrated in panel A for stimulus *simple2*.

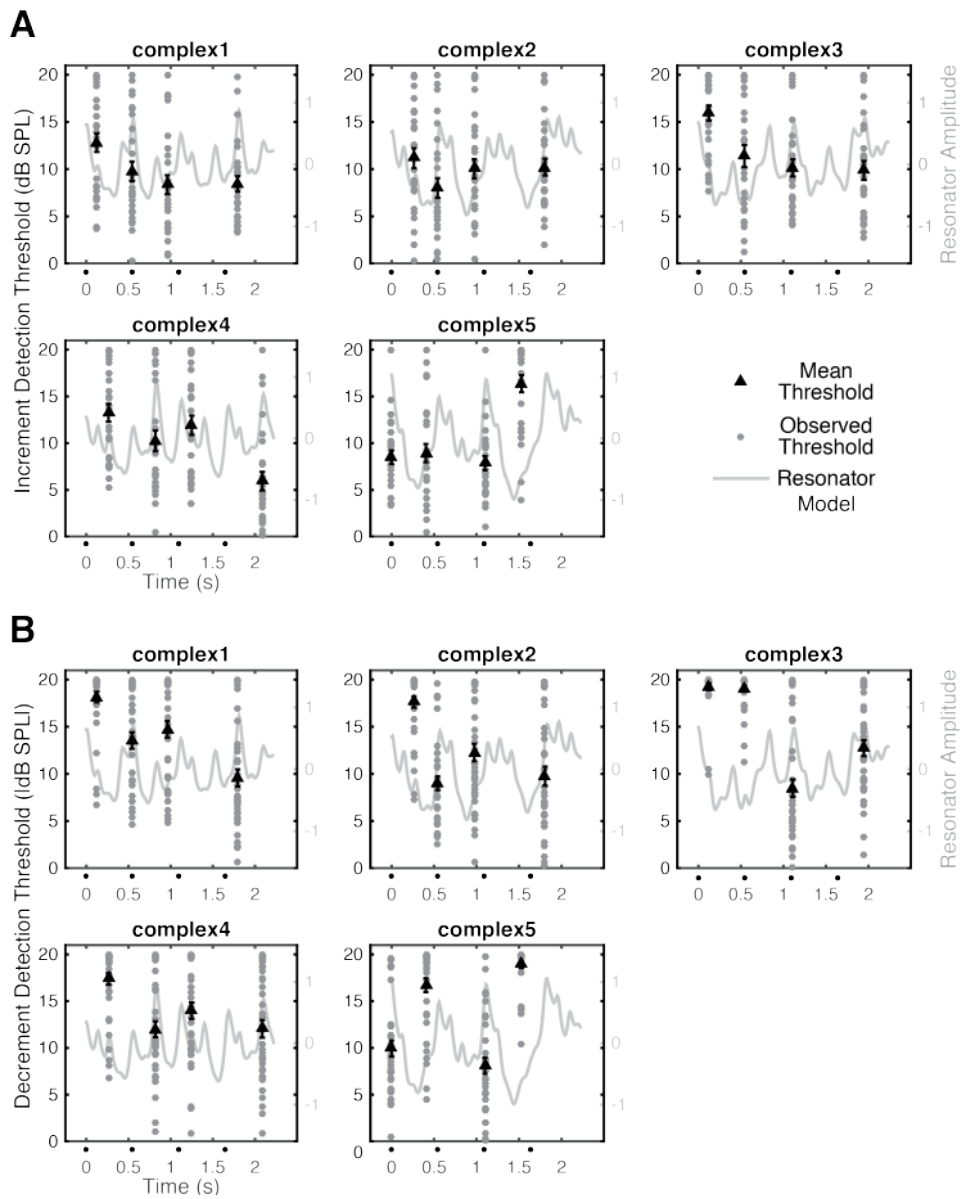MAPPING DYNAMIC TEMPORAL ATTENTION

Figure 3. Thresholds for detecting (A) intensity increments in Experiment 1 and (B) intensity

decrements in Experiment 2 as a function of probe position in the stimulus and resonator-

modeled salience. Thresholds are aligned on the horizontal axis with the corresponding probe

times. Circles = individual participant thresholds, triangles = mean threshold for a probe

position, error bars = ± standard error of the mean, and grey time series = mean output from

resonators driven most strongly by peak periodicities. For reference, black dots below the x-axes

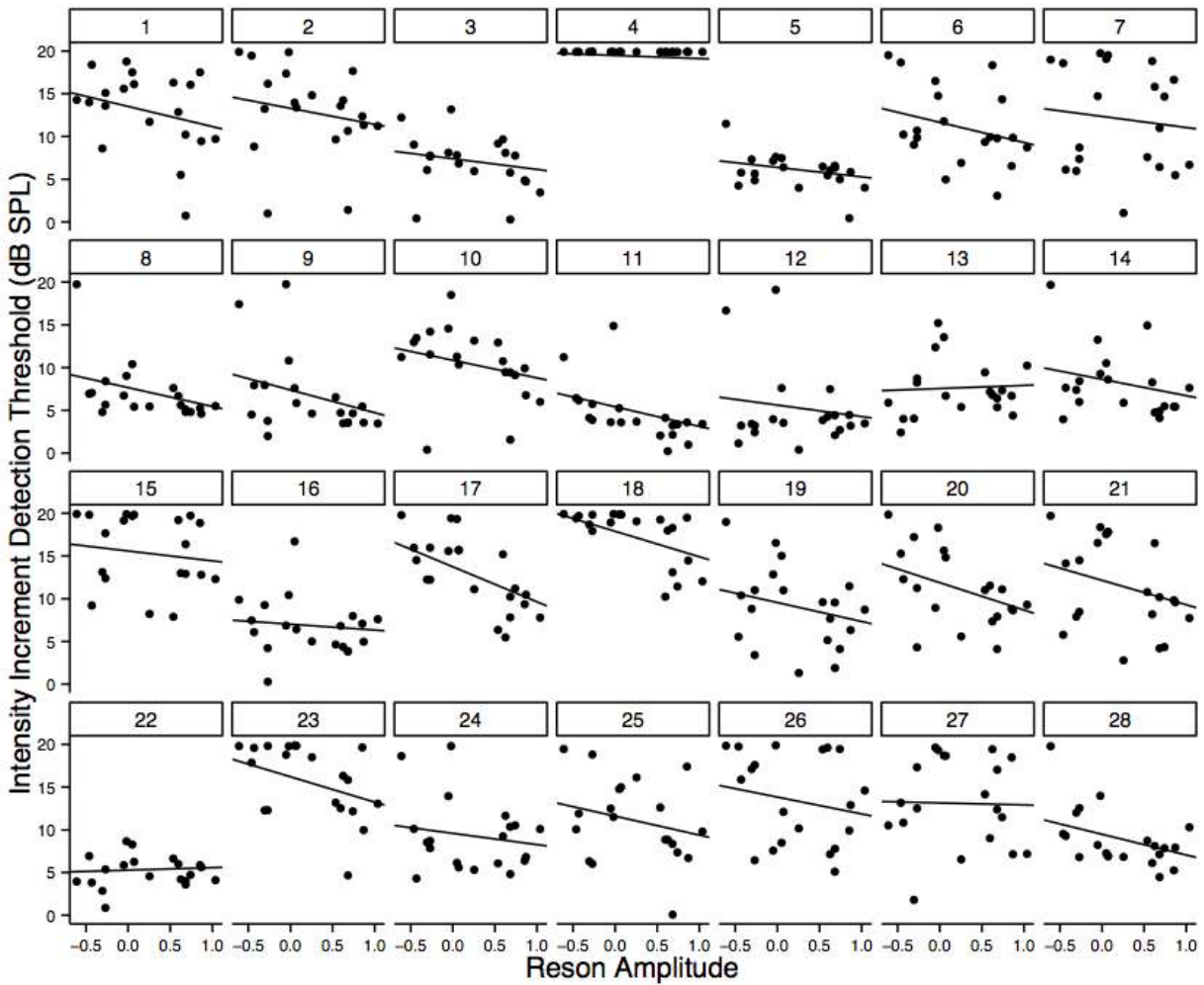MAPPING DYNAMIC TEMPORAL ATTENTION

denote putative beat locations.

Figure 4. Increment detection thresholds from Experiment 1 as a function of reson-filter output amplitude. Each panel is an individual participant's data. Lines reflect each subject's slope and intercept from a random-effects model.
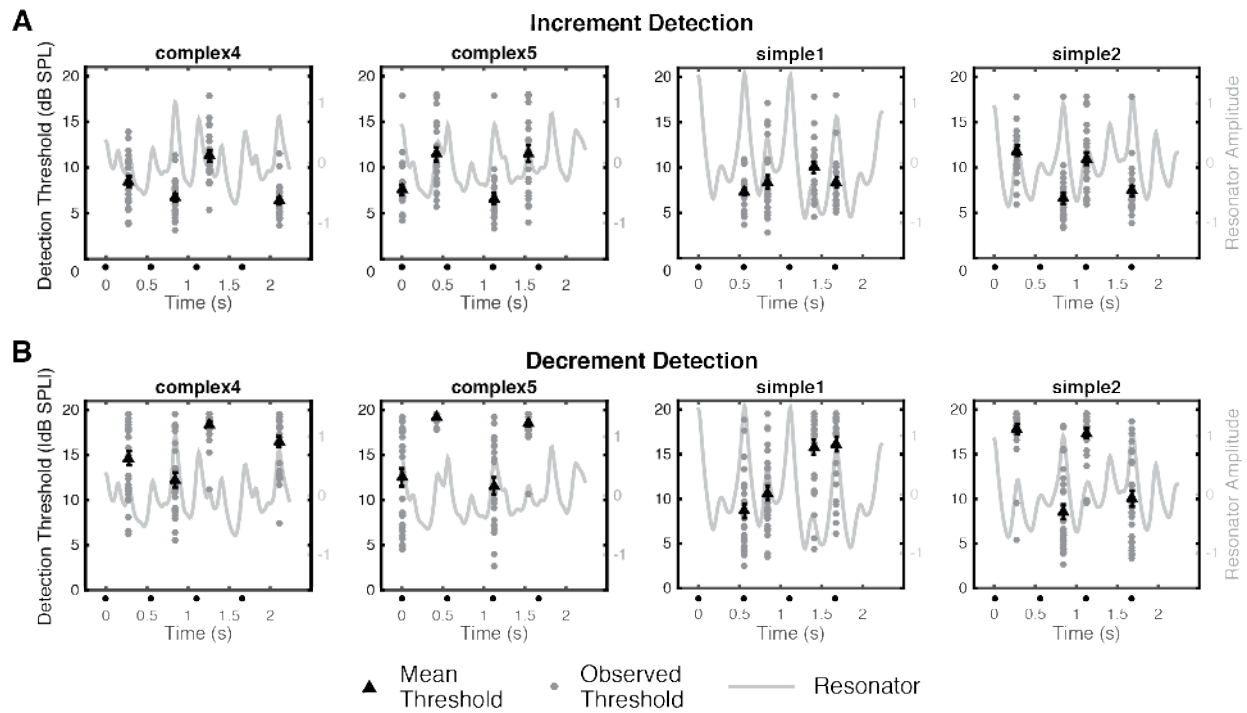
MAPPING DYNAMIC TEMPORAL ATTENTION

Figure 5. (A) Increment and (B) decrement detection thresholds from Experiment 3 as a function of probe position in the stimulus, resonator prediction, and stimulus complexity. Error bars = ± standard error of the mean. Black dots below the x-axes denote putative beat locations.

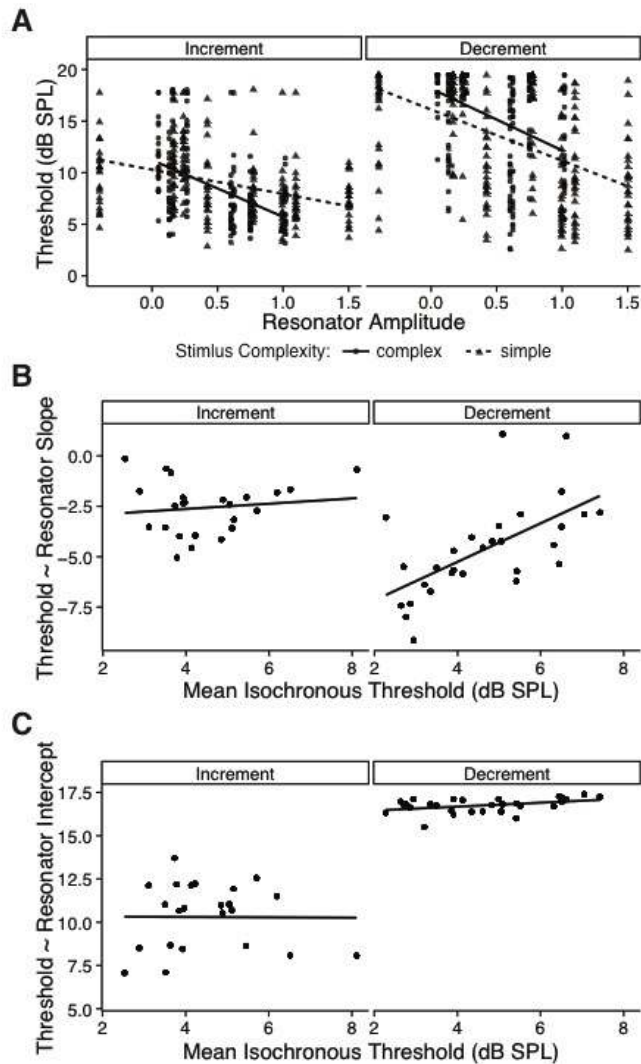MAPPING DYNAMIC TEMPORAL ATTENTION

Figure 6. Stimulus complexity as a moderator of the resonator model's prediction of temporal attending behavior. (A) Increment and decrement discrimination thresholds as a function resonator level for simple and complex stimuli. (B) Degree of threshold ~ resonator association (expressed as slope) in experiment stimuli as a function of intensity deviance discrimination in isochronous stimuli for each participant. (C) Threshold ~ resonator intercept as a function of isochronous intensity deviance discrimination. Participants' slopes and intercepts in B and C were estimated using the same procedures as described in Figure 4. Decrement thresholds are

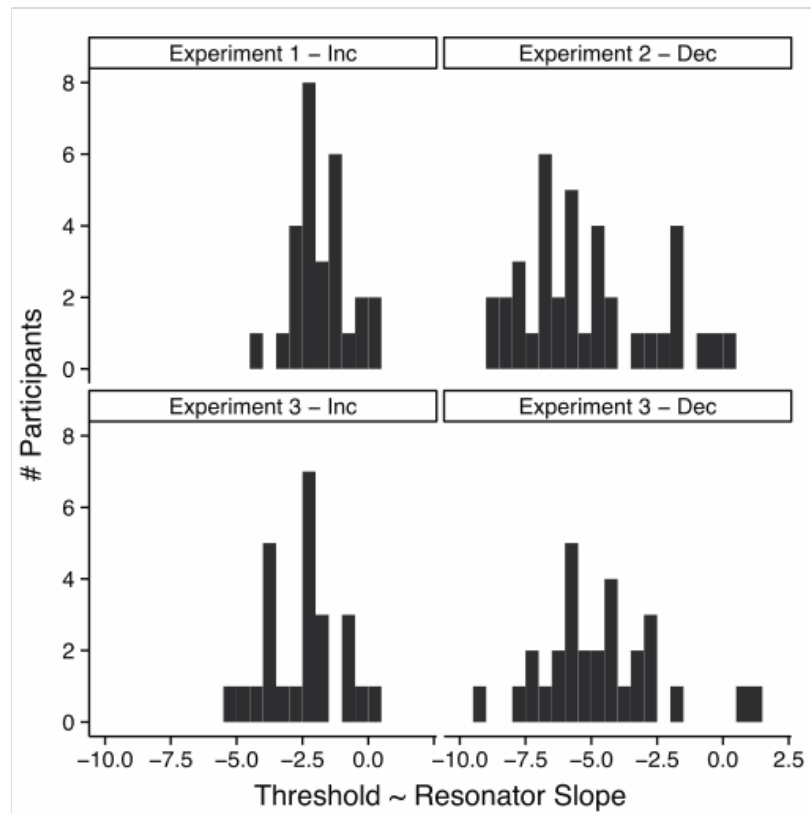MAPPING DYNAMIC TEMPORAL ATTENTION

expressed as absolute values.

Figure 7. Distribution of resonator prediction effects for each experiment. The horizontal axis of each histogram denotes the slope of the relationship between resonator amplitude and intensity threshold, as assessed by a mixed effects model with subject's threshold intercepts and resonator slopes modeled as random effects. Inc = increment detection. Dec = decrement detection.