Title

Spontaneous Sensorimotor Coupling With Multipart Music

Journal

Journal of Experimental Psychology Human Perception & Performance, 40(4)

ISSN

0096-1523

Authors

Hurley, Brian K
Martens, Peter A
Janata, Petr

Publication Date

2014-08-01

DOI

10.1037/a0037154

Peer reviewed

Spontaneous Sensorimotor Coupling with Multipart Music

Brian K. Hurley

University of California, Davis

Peter Martens

Texas Tech University

Petr Janata

University of California, Davis

Author Note

Brian K. Hurley, Department of Psychology and Center for Mind and Brain, University of California, Davis; Peter Martens, School of Music, Texas Tech University; Petr Janata, Department of Psychology and Center for Mind and Brain, University of California, Davis.

We thank Corey Chomas, Hannah Whiteside, and Daman Preet Pannu for assistance with data collection. This work was supported, in part, by a Templeton Advanced Research Program grant from the Metanexus Institute to P.J.

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

Abstract

Music often evokes spontaneous movements in listeners that are synchronized with the music, a phenomenon that has been characterized as being in "the groove." However, the musical factors that contribute to listeners' initiation of stimulus-coupled action remain unclear. Evidence suggests that newly appearing objects in auditory scenes orient listeners' attention, and that in multipart music, newly appearing instrument or voice parts can engage listeners' attention and elicit arousal. We posit that attentional engagement with music can influence listeners' spontaneous stimulus-coupled movement. Here, two experiments – involving participants with and without musical training – tested the effect of staggering instrument entrances across time and varying the number of concurrent instrument parts within novel multipart music on listeners' engagement with the music, as assessed by spontaneous sensorimotor behavior and self-reports. Experiment 1 assessed listeners' moment-to-moment ratings of perceived groove, while Experiment 2 examined their spontaneous tapping and head movements. We found that, for both musically trained and untrained participants, music with more instruments led to higher ratings of perceived groove, and that music with staggered instrument entrances elicited both increased sensorimotor coupling and increased reports of perceived groove. While untrained participants were more likely to rate music as higher in groove, trained participants showed greater propensity for tapping along and they did so more accurately. The quality of synchronization of head movements with the music, however, did not differ as a function of training. Our results shed new light on the relationship between complex musical scenes, attention and spontaneous sensorimotor behavior.

Keywords: synchronization, groove, auditory scene, attention, emotion

Spontaneous Sensorimotor Coupling with Multipart Music

Listeners are often compelled to move in synchrony with music, and the ensuing experience which commonly involves overt movement, e.g. head bobbing, foot tapping, dancing, is generally pleasurable. This music-evoked urge to move and accompanying positive affect is referred to as *being in the groove* (Janata, Tomic, & Haberman, 2012). Recruitment of the action system is perhaps one of the most common manifestations of a listener's engagement with music, but this drive to action is poorly understood.

We postulate that the interaction of music, action, and affect is mediated by the brain's attentional system. Musical factors that are likely to increase a listener's attentional engagement are thus expected to increase the degree of overt sensorimotor coupling. Common compositional devices in music, such as the incremental entry of voices or instrument parts or increases in rhythmic complexity, are known to grab listeners' attention and elicit increased arousal (Grewe, Nagel, Kopiez, & Altenmüller, 2007; Huron, 1992; Keller & Schubert, 2011). Whether such predictors of attentional and affective engagement influence the degree and quality of listeners' motoric engagement with music is unknown. In the present study, we manipulated the number and entry timing of instruments in multipart music to test the hypothesis that distributing, across time, novel musical moments that are capable of attentional capture increases the amount and quality of listeners' spontaneous sensorimotor coupling and subjective measures of music-related positive affect, including ratings of perceived and experienced groove.

Music-evoked sensorimotor synchronization

At its core, moving in synchrony with a musical piece involves synchronizing to a beat, an isochronous pulse that typically forms the foundation of the piece's metric and rhythmic structure. The beat is often embedded in multiple periodicities whose time spans are integer subdivisions or multiples of the beat period. This temporal hierarchy is termed *mete*r, and the various levels of embedded periodicities correspond to levels of the metric hierarchy. The psychological mechanisms of synchronizing movements (particularly tapping) to a metronomic beat have been investigated extensively (for review, see Repp, 2005; Repp & Su, 2013), but recent research on sensorimotor synchronization that uses realistic, metrically structured music suggests that listeners move to music in complex patterns that may not be evident when synchronizing with a metronome (e.g., Martens, 2011; Toiviainen, Luck, & Thompson, 2010).

The tendency of music to induce movement is an integral component of musical engagement. Whereas a considerable amount of research has investigated cued auditory-motor synchronization, research on the aspects of music that influence the initiation and extent of spontaneous movement is largely missing. An emerging area of research on music-evoked movement focuses on the construct of groove (Janata et al., 2012; Madison, 2006; Madison, Gouyon, Ullén, & Hörnström, 2011). Janata et al. (2012) found that, across multiple samples of participants, individuals' self-generated definitions of groove and endorsements of questionnaire items converged on a notion of groove as an aspect of music that evokes a pleasurable urge to move. This, combined with the consistency of perceived groove ratings for the musical excerpts used across multiple experiments in their study, and of groove ratings obtained in a different study using a sample of the same musical stimuli (Stupacher, Hove, Novembre, Schütz-Bosbach, & Keller, 2013), suggests that the assessment of perceived groove is fairly stable among individuals. Echoing participant-generated definitions of groove, Janata et al. (2012) found that ratings of both perceived and experienced groove were positively correlated with ratings of enjoyment. Furthermore, the experience of groove decreased as synchronization became more difficult during bimanual tapping. Importantly, music rated highly in groove elicited a greater amount of spontaneous head and foot movement and greater stimulus-tapping synchronization than music rated low in groove (Janata et al., 2012), as well as greater corticospinal excitability in musicians (Stupacher et al., 2013).

Several rhythmic and spectral features have emerged as predictors of groove and movement characteristics. Madison et al. (2011) found that, across multiple musical styles, the salience of a beat and the number of acoustic events occurring between beats (event density) best predicted the extent to which listeners rated that a musical excerpt grooved. The importance of beat salience is supported by a study that parametrically varied the loudness of the bass drum in dance music and found that participants moved their bodies more and were better entrained to the tempo at louder bass drum levels (Van Dyck et al., 2013). A separate motion capture study found that "pulse clarity," which the authors define as an index of the beat's perceptibility and as inversely related to the entropy of a stimulus' fluctuation spectrum, was associated with the amount and speed of movement among many body parts (Burger, Thompson, Luck, Saarikallio, & Toiviainen, 2013). Burger et al. additionally found that low frequency spectral flux, which is most associated with instruments such as the kick drum or bass guitar that are often found at slower periodicities closer to the beat level, correlated positively with the speed of head movement, and high frequency spectral flux, which is most associated with instruments such as the hi-hat or cymbals and often found at faster periodicities, correlated positively with speed of head, speed of hands, changes in hand distance, and amount of movement.

While associations have been made between acoustical features of natural music and listeners' movement characteristics, it is unknown whether large-scale structural properties of music that ostensibly affect listeners' attentional engagement also influence listeners' motoric behavior with music. We fill this void by directly examining the ability of sequential instrument entrances to elicit sensorimotor coupling with multipart music.

Attentional capture in auditory scenes

The effectiveness of gradually introducing instrument entrances to engage attention and build arousal likely has its roots in perceptual asymmetries: additions or increases in the magnitude of stimulus properties are more salient than subtractions or decreases. For example, ramping (increasing intensity) sounds are perceived as having greater change in loudness than equivalent damping (decreasing intensity) sounds (Neuhoff, 1998), and listeners presented with ramped and damped sinusoids have been shown to associate ramped sinusoids with greater sinusoidal quality (Patterson, 1994). In the context of complex auditory scenes, it has been shown that listeners are considerably better at detecting appearances than deletions of auditory objects, suggesting that the appearance of new objects in the scene “pop out” (Cervantes Constantino, Pinggera, Paranamana, Kashino, & Chait, 2012). Psychophysical studies showing greater sensitivity to stimulus onsets than offsets are paralleled by neurophysiological evidence suggesting that the auditory system is more highly developed to represent sound onsets than offsets (for a review, see Phillips, Hall, & Boehnke, 2002). Similar effects have been observed in visual perception, in which object appearance in a scene is marked by increased saliency and attentional prioritization relative to object disappearance (e.g., Brockmole & Henderson, 2005; Cole & Kuhn, 2009).

Not surprisingly, asymmetries exist in music perception as well and may be utilized by composers to influence listeners’ engagement with the music. For example, increases in stimulus intensity and voice entrances are more perceptible to listeners than decreases in intensity and voice exits (Huron, 1992). Such asymmetries also appear to extend to the perception of music’s rhythmic qualities. For instance, Keller and Schubert (2011) showed that participants rated syncopated musical rhythms as more complex and arousing when a syncopated rhythm followed a non-syncopated variant than when it preceded a non-syncopated variant. Additionally, listeners are more successful in discriminating between an auditory rhythm and a variant with disrupted isochrony if tone onset perturbations are added rather than omitted (Bharucha & Pryor, 1986).

Huron (1992) proposed that composers engage and maintain listeners' attention in musical compositions by employing a "ramp archetype": incremental voice entries or gradual stimulus level increases followed by occasional large decrements in stimulus level or a large simultaneous reduction of multiple voices. In a survey of a diverse array of western art music, Huron observed that the dynamics of solo piano music contained more frequent crescendos than diminuendos and polyphonic (multi-voice) compositions contained more frequent voice entries than voice exits. Grewe et al. (2007) found, in a variety of recorded multipart music, that entrances of new parts were the best predictors of dynamic changes in participants' subjective feelings and physiological arousal. Although theory suggests that the entry of parts in multipart music engages listeners' attention, and part entries and increases in rhythmic complexity have been shown to influence listeners' arousal, such manipulations have not been tested in the context of music-induced movement.

In the current study, we examined the effect of ramping stimulus information on spontaneous sensorimotor coupling by temporally staggering the onset of new instrument streams in multipart music. We compared subjective and objective responses to musical stimuli in which a new instrument part entered periodically throughout the duration of a stimulus, with musical stimuli in which all instrument streams began simultaneously. In the first experiment, we examined the effect of staggered instrument entrances, as well as the number of instruments present in a musical scene, on dynamic, stimulus-coupled ratings of groove. In the second experiment, we investigated the degree to which participants spontaneously began tapping and/or moving their heads during the playback of musical stimuli that contained either temporally staggered or simultaneous instrument entrances. We predicted that as new parts entered, participants' overt engagement with the music would increase. Specifically, we hypothesized that listeners' dynamic reports of groove and their spontaneous stimulus-coupled movements would increase following staggered instrument entrances relative to the same time periods in musical stimuli with simultaneous instrument entrances.

Experiment 1

Method

Participants. Fifteen participants (8 females, mean age ± *SD*: 21.6 ± 1.9 years) were recruited from an undergraduate research participant pool at the University of California, Davis in exchange for course credit. All participants reported having normal hearing. All participants provided informed consent in accordance with a protocol approved by the UC Davis Institutional Review Board.

Stimuli. Ten stimuli were generated in Logic Pro 8 (Apple, Inc., Cupertino, CA) by combining individual audio tracks of looped instrument patterns (Apple Loops) into combinations of 3 – 4 instruments and repeating the combined loops to form 40 s excerpts of multi-part music. Stimuli represented the genres of rock, jazz, funk, bluegrass, hip-hop, reggae, new age, and electronic dance music, and the rhythmic structure of all stimuli was consistent with a 4/4 meter. These 10 excerpts served as the parent stimuli that were subsequently manipulated to create the conditions described below. Two versions of each excerpt were created such that 1) all instruments began playing simultaneously at the onset of the stimulus (*simultaneous*) or 2) instrument parts began playing sequentially at temporally offset locations in time (*staggered*) until the musical scene was fully populated with all instrument parts (Figure 1). Additional versions of each excerpt were created such that varying numbers of instruments were present: *full*, in which all instrument parts were present; *reduced*, in which 2 – 3 instrument parts were present; and *solo*, in which a single instrument was present.

Apparatus. Participants were tested individually in a sound-attenuating booth, where they sat at a desk equipped with a computer monitor, a keyboard, a mouse, and a MIDI slider apparatus identical to that used by Janata et al. (2012, Study 1). The analog slider signal was converted to MIDI parameter values (0 – 127) using a MIDITools computer[1], routed through a MOTU 828mkII mixer, and recorded using Digital Performer (MOTU, Inc., Cambridge, MA) at a sampling rate of 100 Hz. We concurrently recorded an audio track of the stimulus audio. Stimuli were presented at a comfortable listening level through Tannoy Reveal 6 speakers situated roughly 40 degrees to participants' right and left sides and roughly 45 inches away from participants. All instructions, stimuli, and questionnaires were delivered under the control of Ensemble (Tomic & Janata, 2007), a web interface for administering behavioral experiments.

Procedure. Near the beginning of the experiment, participants were asked if they had ever heard the term groove applied to music. If not, they were given the following definition of groove: "'the groove' is the aspect of music that compels the body to move." Participants were instructed to use the provided slider to continuously report their perceived level of groove on a scale of 0 (music does not groove at all) to 10 (music imparts a very strong feeling of groove) as each excerpt played. After each excerpt ended, participants were presented with three questions, to which they responded on 7-point scales: "How much did you enjoy what was just played (1 = not at all; 7 = very much)?", "To what extent did you feel compelled to move while listening to the musical excerpt (1 = not at all; 7 = very much)?", and "How much would you have liked to continue listening to this excerpt (1 = not at all; 7 = very much)?"

Prior to beginning the experiment, participants were given three practice trials for which stimuli were randomly selected from 20 looped multipart drum patterns that were used in previous groove studies (Janata et al., 2012). After completing the practice trials and confirming their understanding of the task, participants entered the experiment phase consisting of 40 trials. These were a random selection of 10 simultaneous stimuli, 10 staggered stimuli, 10 reduced stimuli, and 10 solo stimuli (note that the simultaneous stimuli also served as the *full* level of the musical scene density variable, since all instrument parts were present in this condition).

Preprocessing and data analysis. Data were preprocessed using custom scripts and third party toolboxes in MATLAB (MathWorks, Inc., Natick, MA). We used the MIDI Toolbox (Eerola & Toiviainen, 2004) to extract MIDI parameter values corresponding to changes in slider position across each participant's experiment session. We then segmented each participant's slider time series into trials using indices of silence from the ongoing stimulus waveform (-25 dB threshold). We matched each segment with the trial's stimulus and experimental condition based on Ensemble's record of stimulus presentation order.

We assessed changes in groove ratings following each new instrument entrance by calculating the slope of the inter-entrance linear trends (Figure 2A). To compare differences in changes of slope between simultaneous and staggered conditions, we fit linear trends between the time points in the simultaneous entrance condition that corresponded to instrument entrances in the staggered versions of the stimulus.

Given that all parts in the musical scene density conditions were present throughout a given trial and that participants' groove slider ratings for these conditions tended to increase during the first half of trials and plateau during the last half of trials, we averaged slider ratings across the final 10 s to obtain static groove ratings.

Statistical analyses. Linear mixed effects models were fit to the data using maximum likelihood (ML) estimation and implemented using the function *lmer* from the *lme4* library (Bates, Maechler, & Bolker, 2013) in R (R Core Team, 2012). Each model included random intercepts for individual participants and stimulus exemplars to control for those sources of random variation. We evaluated the fixed-effect factors of entrance type, entrance interval, and musical scene density using a likelihood ratio test to compare the fit between pairs of nested models: a general model with all factors and a restricted model with the factor of interest omitted. If the general model fit significantly better than the restricted model, we interpreted the omitted factor to have a significant main effect. We implemented contrasts via the *lsmeans* package (Lenth, 2013) in R, with *p*-values adjusted to control for false discovery rate (Benjamini & Hochberg, 1995).

Correlations between slider groove ratings and responses to post-stimulus questions within participants were calculated by regressing out between-subject variability from each correlation coefficient in order to account for repeated measures (Bland & Altman, 1995).

Results

Groove ratings. The slope of groove slider ratings generally decreased over the course of the trial (Figure 2B), with slopes under the simultaneous condition decreasing more rapidly than those under the staggered condition. There was a significant interaction between entrance type and entrance interval [$\chi^2(3) = 15.25$, $p < .01$]. Groove ratings for the simultaneous condition rose sharply near the beginning of the stimulus, but the slope of ratings quickly plateaued by the second entrance interval and settled towards zero in subsequent intervals. This is reflected by a trend across entrance intervals with both negative linear [$t(1021) = -11.58$, $p < .0001$] and quadratic [$t(1016) = 3.56$, $p < .001$] components. Conversely, groove ratings for the staggered condition continued to increase across entrance intervals with rating slopes decaying steadily, as reflected by a negative linear trend [$t(1021) = -8.05$, $p < .0001$] and an absence of a quadratic component.

We observed a main effect of musical scene density [$\chi^2(2) = 44.22$, $p < .001$], indicating that groove ratings were higher for music in which more instruments were present (Figure 3). Post-hoc comparisons revealed that this effect was driven by whether one instrument or multiple instruments were present, as both full [$t(574) = 6.36$, $p < .0001$] and reduced [$t(574) =5.57$, $p < .0001$] musical scenes elicited significantly higher groove ratings than solos, but were not significantly different from each other.

Post-stimulus ratings. Figure 4 presents mean ratings from the post-stimulus questions, indexing one's enjoyment, urge to move, and wanting to continue listening as a function of entrance type (Figure 4A) and musical scene density manipulations (Figure 4B). Here we report comparisons among modeled means of entrance type and scene density conditions, which we obtained from mixed-effects models with either entrance type or scene density as a fixed effect. Participants rated all three post-stimulus items higher for the staggered condition than for the simultaneous condition [urge to move: $t(275) = 2.42$, $p < .05$; enjoyment: $t(275) = 2.11$, $p < .05$; wanting the stimulus to continue: $t(275) = 3.03$, $p < .01$]. Participants rated all three post-stimulus items higher for full musical scenes than for solo instrumentations [urge to move: $t(574) = 6.34$, $p < .0001$; enjoyment: $t(574) = 6.45$, $p < .0001$; want to continue listening: $t(574) = 5.07$, $p < .0001$]. Participants also rated all three post-stimulus items higher for reduced musical scenes than for solos: urge to move[$t(574) = 5.12$, $p < .0001$; enjoyment: $t(574) = 5.29$, $p < .0001$; want to continue listening: $t(574) = 4.07$, $p < .0001$]. Mean urge to move [$t(574) =$

2.20, $p < .05$] and enjoyment [$t(574) = 2.16$, $p < .05$] ratings were higher for full than for reduced, although mean desire to continue listening did not differ significantly between full and reduced conditions.

As Table 1 (“Experiment 1” column) reports, all post-stimulus ratings were positively (.73-.82) and significantly (all *p*-values < .001) correlated with mean groove ratings. Groove ratings correlated most strongly with the urge to move.

Discussion

In line with our hypothesis, the periodic entrance of new instrument parts in the staggered condition appeared to sustain participants’ engagement, as evident by a continuing increase in listeners’ perceived groove and increased desire to continue listening relative to stimuli in which instruments entered simultaneously at the beginning. Moreover, stimuli with staggered entrances garnered stronger urge to move and enjoyment ratings than did stimuli with simultaneous entrances. The same pattern of results was obtained for musical scene density. One possible explanation for increased groove with multipart music is that, when the temporal patterns of the various instrument parts do not overlap fully, the the number of perceptible events that subdivide the beat period increases, thus increasing event density, a feature that is predictive of groove ratings (Madison et al., 2011). Another explanation, along the lines of Keller and Schubert (2011), stems from the observation that the initial instrument part in the staggered condition of each excerpt tended to be less syncopated than the parts that entered second, third, or fourth; this situation is typical in popular and dance-oriented musical genres. Syncopation in these subsequent parts is thus more salient, with a corresponding increase in arousal.

Groove ratings correlated positively and strongly with the extent to which participants experienced enjoyment, wanted a stimulus to continue playing, and experienced the urge to move, echoing the findings of Janata et al. (2012).

Experiment 2

Given that the explicit judgments of groove in Experiment 1 confirmed our predictions, in Experiment 2 we sought to determine whether staggered entrances under similar stimulus conditions would have a potentiating effect on overt movements, both in terms of the amount of movement and the quality of coupling between the movements and the music. Theories of dynamic attending and neural resonance (Jones & Boltz, 1989; Large & Jones, 1999; Large & Snyder, 2009) and evidence supporting them (e.g., Kung, Tzeng, Hung, & Wu, 2011; Nozaradan, Peretz, Missal, & Mouraux, 2011; Snyder & Large, 2005) predict that internal, oscillating "attending rhythms" entrain to hierarchical levels of periodic stimuli and facilitate the dynamic allocation of neural resources to metrically salient points in time. Thus, in addition to our main manipulation of staggered and simultaneous instrument entrances, we manipulated the metrical position of entrances to test whether rhythmic attending interacted with the temporal location of instrument stream onsets to influence tapping or head movement behavior.

Method

Participants. Fifty-nine individuals (37 females, mean age ± *SD*: 21.0 ± 2.4 years) participated in the current experiment. Fifty-eight of the participants were recruited from an undergraduate research participant pool at the University of California, Davis in exchange for course credit, and one participant was a graduate student from the UC Davis Department of

Psychology. All participants reported normal hearing. All participants provided informed consent in accordance with a protocol approved by the UC Davis Institutional Review Board.

Stimuli. Using the same stimulus composition method as in Experiment 1, we generated 20 excerpts that consisted of five exemplars in each of the following genres: electronic dance music, folk, rock, and soul/funk. All excerpts comprised four instrument parts and were in 4/4 time. We manipulated the instrument part onsets of each exemplar to create simultaneous and staggered conditions similar to those described in Experiment 1 (Figure 1). We created two staggered conditions for the present experiment: one condition with instrument entrances at metrically strong positions (Beat 1 of a measure), and a separate condition with instrument entrances at relatively metrically weak positions (Beat 2 or 4 of a measure or off-beat). To maintain the same degree of rhythmic and tonal coherence between the two staggered conditions, instruments in the metrically weak staggered condition began playing at the position where the part would have been had it entered on the first beat of the measure.

Apparatus. Participants were assessed individually in the same setting as in Experiment 1, with the addition of a MIDI drum pad and a motion capture system. The drum pad (Roland HPD-15) stood 26 inches from the ground and was partially occluded by a desk so that only the bottom left and bottom right pads of the drum's surface were exposed. Participants sat at the desk with the drum pads directly in front of them and a computer mouse within easy reach (Figure 5). The drum pad's audio was disabled so that participants did not hear any auditory feedback from their taps, aside from sounds generated by their hands striking the pad. We did not differentiate between taps to the left and right pads in the tapping analyses.

We measured spontaneous head movements during stimulus playback using a Zebris CMS-HS 3D motion capture system (Zebris Medical GmbH). The system transmits ultrasonic pulses at a rate of 100 Hz from body markers and records the pulses with a measuring sensor to calculate the position of the markers in three-dimensional space. The measuring sensor was situated on the opposite side of the desk from the participant with the head of the unit inclined toward the participant at 60 degrees (Figure 5). Body markers were attached to each participant's forehead and chest using adhesive patches. Recording from the chest marker enabled us to assess head movements relative to chest position, thereby controlling for changes in the position of a participant's torso. Data were recorded using WinData (Zebris Medical GmbH), and were exported as ASCII files for further preprocessing.

MIDI events from the drum pad were routed through a MOTU 828mkII mixer and recorded in Digital Performer alongside an audio track of the stimulus audio, which we used to parse the tapping and motion capture data into trials based on time indices of silence in the stimulus audio feed. Additionally, we used synchronization signals from the Zebris CPU unit to demarcate when motion capture recording started and stopped in order to co-register the motion capture recording and the stimulus audio. Custom MATLAB scripts controlled stimulus selection, and Ensemble (Tomic & Janata, 2007) controlled stimulus presentation and the acquisition of question responses.

Procedure. Participants were asked if they had ever heard the term "groove" applied to music, and if so, to give their definition of groove. All participants (regardless of their familiarity with the term) were subsequently given the same definition of groove as was given in Experiment 1[2]. Next, participants were told that they would hear various clips of music and to keep their left hand resting on the left drum pad section and their right hand resting on the right drum pad section as the music played. Participants were instructed to listen to the music and to tap or drum along on the drum pad in any way they wished if the music compelled them to do so.

We encouraged participants' natural engagement and attempted to minimize performance anxiety by emphasizing that we were not assessing their musical ability, and that they should interact comfortably with the music as they would at home. They were given the option of not tapping at all.

Once participants confirmed that they understood the instructions, they were given three practice trials to allow them to practice tapping on the drum pad and to become familiar with the experimental setting. Practice stimuli were selected randomly from the drum loops used as practice stimuli in Experiment 1. Each experiment trial consisted of stimulus playback, during which time spontaneous tapping and head movements were recorded. After stimulus playback, participants rated the following three questions on seven-point scales: "To what extent did you feel that the musical excerpt grooved (1 = least groove; 7 = most groove)?", "How much did you enjoy the musical excerpt (1 = not at all; 7 = very much)?", and "How much would you have liked the music to continue (1 = not at all; 7 = very much)?" Submission of the rating form initiated the next trial. Aside from the stimuli, practice trials were identical to experimental trials. Following the practice phase, the experimenter addressed any of the participant's remaining questions. The participant then started the experiment phase, which was divided into two blocks with a short break between blocks. During each block, participants were given 10 trials with simultaneous stimuli and 10 trials with staggered stimuli in randomized order. Stimuli were selected such that no two stimuli within each of the blocks were derived from the same parent excerpt. There were a total of 40 trials. Following the final trial, participants rated the degree to which they felt self-conscious while tapping towards the beginning and towards the end of the experiment.

Preprocessing and Data Analysis. Data were preprocessed and analyzed using a combination of custom scripts and third party toolboxes in MATLAB and R.

Tapping. Tap onsets were extracted from each participant's MIDI recording using the MIDI Toolbox (Eerola & Toiviainen, 2004). Each MIDI recording was parsed into trials and matched with trial conditions using the same trial parsing method as was used in Experiment 1.

Selectivity groups. During data collection, we observed considerable variability in how participants approached the tapping portion of the experiment. There were those who tapped enthusiastically for the majority of stimuli (exhibited *low selectivity*), some who tapped on some trials but not others (*high selectivity*), and those who refrained from tapping for the majority of the experiment (*non-tappers*). We designated participants who produced fewer than 10 taps on 85% or more of the trials as non-tappers[3]. We then calculated each remaining participant's tapping selectivity as the ratio of between-trial tapping variability to the number of taps averaged across trials:

$$selectivity_i = \frac{SD_{taps_i}}{M_{taps_i}}$$

where $_{\text{taps}}$ refers to the number of taps within each trial and $_i$ refers to individual participants. Thus, participants who varied more in the average amount they tapped from trial to trial had higher selectivity indices. We formed high and low selectivity participant groups using a median split of the selectivity index scores (Figure 6).

Tapping rate. We calculated each participant's average tapping rate for each trial by dividing the total number of taps across the trial by the duration of the trial. We also calculated inter-entrance tapping rates as the number of taps within 1 s prior to the second instrument entrance (*baseline* period) and the number of taps/s within 3.9 s intervals following each subsequent entrance. We chose a post-entrance interval of 3.9 s because the intervals between instrument entrances varied among stimuli, and 3.9 s was the shortest among all inter-entrance intervals.

Motion capture. Motion capture analyses were performed for 55 of the 59 participants (data were not recorded properly for four participants). We imported the 3-dimensional marker coordinate data from the Zebris-generated ASCII files into MATLAB using the Motion Capture Toolbox (Toiviainen & Burger, 2013). Using the Zebris synchronization signal, we determined the onset asynchrony between the stimulus audio and motion capture recordings. We then parsed participants' motion analysis recordings into trials and matched the trials with their corresponding stimuli using the trial parsing technique described in Experiment 1. Transient drops in the motion capture signal were replaced with a spline interpolation of the values surrounding the dropped signal. Trials containing more than 25% signal loss were omitted from further analyses, which resulted in the omission of 23 trials. The remaining trials contained an average of 2% signal loss. The motion analysis data were subsequently transformed from absolute to relative coordinate space by calculating the difference in position (in x-, y-, and z-planes) between the head and chest markers. The resulting signals were passed through a fourth-order Butterworth high-pass filter, with a frequency cutoff of .5 Hz, to attenuate low frequency artifacts related to changes in head and body position during recording.

Resonator model. We assessed the temporal structure of participants' spontaneous movements in relation to the temporal structure of the stimuli by processing stimulus, tapping, and head movement data through a resonator model developed by Tomic and Janata (2008)[4]. [See also the *Modeling Music and Behavior* section of Janata et al. (2012).] Of primary importance to the present paper are two representational stages of the model: the average periodicity surface (APS) and mean periodicity profile (MPP). The APS is a time-frequency plot depicting the amount of energy through time in a bank of 99 reson filters (tuned between 0.25 and 10 Hz), and the MPP is an amplitude spectrum that averages the APS across time (Figure 7). For the motion capture data, each of the three (*x*-, *y*-, and *z*-plane) preprocessed signals was passed into a separate resonator bank, leading to three periodicity surfaces. The periodicity surfaces were averaged to create an APS, from which the MPP was calculated.

Model metrics. We calculated an APS and an MPP for each stimulus, and we calculated an APS for participants' spontaneous head movements and tapping responses from each trial. Given that modeling the periodic structure of a participant's tapping required at least several taps within a trial, we omitted trials from model-dependent tapping analyses in which participants generated fewer than two taps.

We calculated each stimulus' prevalent periodicities by identifying MPP peak heights that were greater than 5% of the MPP's amplitude range. We then assessed the width of MPP peaks by measuring the full-width at half-maximum of each peak. We matched peak widths that fell within a single resonator frequency to the corresponding frequency, whereas we matched peak widths that spanned more than one resonator frequency to the corresponding range of neighboring frequencies. We classified peak frequencies as metrically related frequencies, and we classified non-peak frequencies as non-metric frequencies. We then calculated a *stimulus-matching ratio* (SMR) for each time sample of each motion capture and tapping APS by dividing the mean resonator energy (rms) occurring within metrically related frequencies by the mean resonator energy occurring in non-metric frequencies (Figure 7). This calculation resulted in two time series for each trial, indicating the time-varying degree to which spontaneous (1) tapping and (2) head movement periodicities matched the stimulus' rhythmic structure. We assessed changes in SMR following instrument entrances in staggered conditions by calculating the mean

SMR across the baseline interval and across 3.9 s intervals following instrument entrances. In order to compare pre-and post-entrance stimulus-movement synchronization between staggered and simultaneous conditions, we calculated the mean SMR for each simultaneous stimulus across the same pre- and post-entrance intervals as the staggered metrically strong condition derived from the same parent excerpt. Although we recognize that, in the strictest since, sensorimotor synchronization involves matching movements to both the period and phase of a stimulus (London, 2004), we consider the matching between stimulus and movement periodicities here to be evidence of synchronization, and we thus refer to SMR and synchronization interchangeably.

In addition to assessing stimulus–movement synchronization, we calculated the magnitude of spontaneous head movements by calculating the mean resonator energy output across all 99 motion capture APS resonators for each time sample. This resulted in a time series of mean *head movement energy*. We assessed changes in head movement energy following instrument entrances in staggered conditions by calculating mean head movement energy across the baseline interval and across 3.9 s intervals following instrument entrances. In order to compare time-varying sensorimotor coupling among entrance conditions, we calculated the mean head movement energy and SMR for simultaneous stimuli across the same pre- and post-entrance intervals as the staggered metrically strong stimuli derived from the same parent excerpt.

Results

We found no significant difference between the staggered condition with metrically strong instrument entrances and that with metrically weak instrument entrances in any of the movement responses. We thus collapsed across the two staggered conditions, such that all statistical analyses reported here compare the simultaneous condition with the pooled staggered condition. Using the same statistical procedures as in Experiment 1, we fit separate linear mixed-effects models for each selectivity group (high-selectivity, low-selectivity, and non-tappers) and for each dependent measure (tapping rate, tapping SMR, head movement energy, and head movement SMR). Each model included entrance type (simultaneous or staggered) and entrance period (baseline, second, third, and fourth entrances) as fixed-effect factors, and individual participants and individual stimuli as random intercepts.

A paired t-test comparing participants' reported self-consciousness towards the beginning and towards the end of the experimental session revealed that, on a seven-point rating scale, participants on average rated their task-related self-consciousness 1.14 points higher for the beginning than for the end of the experiment [$t(58) = 8.81$, $p < .0001$]. We thus included experiment block number (1 or 2) as an additional covariate in all statistical models to account for differences in responses between the first and second halves of the experiment.

Tapping rate. For both low and high selectivity groups, stimuli with staggered instrument entrances elicited a greater increase in mean tapping rate from each measured time window to the next than did stimuli in which the instruments entered simultaneously (Figure 8A), as indicated by significant Entrance Type X Entrance Period interactions [low selectivity: $\chi^2(3) = 48.13$, $p < .001$; high selectivity: $\chi^2(3) = 37.76$, $p < .001$]. The low selectivity group's mean tapping rate increased linearly across entrance periods both for simultaneous [$t(3948) = 11.68$, $p < .0001$] and staggered [$t(3948) = 21.23$, $p < .0001$] conditions, with a significant cubic trend in the staggered condition [$t(3948) = -3.42$, $p < .01$]. The mean tapping rate in the staggered condition increased to a greater degree over the course of a given trial for the low selectivity group, as reflected by the significantly lower mean tapping rate for the staggered than for the simultaneous condition within the first two measured time windows [baseline: $t(3948) = -5.29$, $p < .0001$; Entrance 2: $t(3948) = -3.89$, $p < .001$] followed by a significantly greater mean tapping rate for the staggered than for the simultaneous condition within the final time window [Entrance 4: $t(3948) = 3.03$, $p < .01$]. We observed similar effects in the high selectivity group, such that tapping rates increased linearly across entrance periods for both the simultaneous [$t(3789) = 3.45$, $p < .01$] and staggered [$t(3789) = 12.00$, $p < .0001$] conditions, with tapping rates increasing more rapidly throughout the trial for the staggered condition than for the simultaneous condition. The latter observation is indicated by significantly lower mean tapping rates for staggered than for simultaneous conditions during the baseline period [$t(3789) = 2.37$, $p < .05$] followed by a significantly higher mean tapping rate for staggered than for simultaneous stimuli by the final Entrance 4 period [$t(3789) = 5.68$, $p < .0001$]. We found no significant effects for the non-tapper group, as the mean tapping rate for this group was at or near zero taps/s for all measured time windows.

Tapping synchronization. The low selectivity group (Figure 8B, top panel) exhibited an increase in mean SMR across the duration of the stimulus, as evident from a main effect of entrance period [$\chi^2(3) = 172.77$, $p < .0001$] and a significant linear trend across entrance periods [$t(3020) = 13.17$, $p < .0001$]. The low selectivity group generally exhibited a greater mean SMR for the simultaneous than for the staggered condition [$\chi^2(1) = 16.35$, $p < .0001$]. There was no significant interaction between entrance type and entrance period for the low selectivity group. The high selectivity group's tapping also became increasingly attuned to stimulus periodicities throughout the course of a trial, regardless of entrance type (Figure 8B, center panel), as indicated by a main effect of entrance period on mean SMR [$\chi^2(3) = 27.34$, $p < .0001$] and a significant linear trend across entrance periods [$t(1260) = 4.87$, $p < .0001$]. No other factors or interactions were significant for the high selectivity group's mean SMR. Intriguingly, even though participants in the non-tapper group tapped very little throughout the experiment, the rhythmic quality of tapping for the trials that they did tap for differed distinctly between simultaneous and staggered conditions (Figure 8B, bottom panel), as reflected by a significant Entrance Type x Entrance Period interaction [$\chi^2(3) = 15.27$, $p < .01$]. Mean tapping SMR for the simultaneous condition showed a trend toward a linear decrease [$t(70) = -2.42$, $p = .09$], whereas mean SMR for the staggered condition showed a trend toward a significant increase [$t(70) = 2.21$, $p = .09$], such that mean SMR was significantly greater for staggered than for simultaneous stimuli by the final entrance period [Entrance 4: $t(59) = 3.82$, $p < .01$]. The difference between simultaneous and staggered conditions during the other entrance periods were not significant, likely due to the small number of tapping trials available from the non-tapper group for SMR calculation.

Head movement energy. Figure 9A-C depicts mean head movement energy as a function of entrance type and entrance period. Differences in the magnitude and selectivity of participants' spontaneous tapping also extended to the magnitude of participants' spontaneous head movements, as indicated by a significant effect of selectivity on mean head movement energy [$\chi^2(2) = 9.06$, $p < .05$]. In accordance with the tapping results, head movement energy for the low selectivity group was greater than that for the high selectivity group [$t(52) = 2.31$, $p < .05$] and the non-tapper group [$t(52) = 2.75$, $p < .05$], although the difference in head movement energy between the high selectivity and non-tapper groups did not differ significantly.

The magnitude of head movement in the low selectivity group differed between simultaneous and staggered conditions over the course of a trial (Figure 9A), as revealed by a significant Entrance Type X Entrance Period interaction [$\chi^2(3) = 10.64$, $p < .05$]. In the simultaneous condition, mean energy increased slightly during the earlier portion but plateaued for the latter portion, such that the linear increase in head movement energy across entrance periods was only marginally significant [$t(3411) = 2.29$, $p = .07$]. In the staggered condition, mean energy increased to a much greater degree across the course of a stimulus, with a strong linear trend across entrances [$t(3411) = 6.44$, $p < .0001$]. Although mean head movement energy was significantly lower in the staggered condition than in the simultaneous condition towards the beginning of a stimulus [baseline: $t(3422) = -2.67$, $p < .05$; Entrance 2: $t(3411) = -3.25$, $p < .01$], the greater increase in energy across entrances for the staggered condition led to a non-significant difference in energy between the simultaneous and staggered conditions for the latter entrance periods.

Across entrance periods, head movement energy for the high selectivity group was generally greater for the simultaneous condition than for the staggered condition (Figure 9B), as indicated by a significant effect of entrance type [$\chi^2(1) = 8.34$, $p < .01$]. There was also an effect of entrance period [$\chi^2(3) = 14.22$, $p < .01$], such that mean head movement energy increased linearly [$t(3609) = 3.77$, $p < .001$] throughout a trial. No significant interaction existed for the high selectivity group. Head movement energy for the non-tapper group (Figure 9C) was highly variable and did not show significant variation as a function of entrance type, entrance period, or an interaction between entrance type and entrance period.

Stimulus-head movement synchronization. The matching of head movements to the rhythmic structure of stimuli as a function of entrance type and entrance period is presented in Figure 9D-F. For the low selectivity group, the trajectory of the SMR among successive entrance periods differed between entrance types (Figure 9D), as indicated by a Entrance Type X Entrance Period interaction [$\chi^2(3) = 15.46$, $p < .01$]. As with head movement energy, the low selectivity group's mean SMR increased during the early portion of stimuli with simultaneous entrances but plateaued in the latter portion of stimuli, as indicated by both linear [$t(3411) = 5.52$, $p < .0001$] and negative quadratic [$t(3411) = -3.29$, $p < .01$] components in the change of mean PMR across entrance periods. Conversely, the mean SMR for staggered stimuli increased steadily with a strong linear trend [$t(3411) = 7.85$, $p < .0001$] across entrance periods. Furthermore, although the SMR for staggered and simultaneous conditions were statistically indistinguishable during the baseline period and that SMR means were smaller for the staggered condition than for the simultaneous condition during Entrance Periods 2 [$t(3411) = -2.63$, $p < .05$)] and 3 [$t(3411) = -2.95$, $p < .05$), the SMR for the staggered condition became higher than that for the simultaneous condition during the final measured period (Entrance 4), although the difference only approached significance [$t(3411) = 1.97$, $p = .06$]. For the high selectivity group, the mean SMR increased with successive entrance periods, as evident from a main effect of entrance period [$\chi^2(3) = 88.51$, $p < .0001$] and a significant positive linear trend of SMR across entrance periods [$t(3609) = 9.25$, $p < .0001$] (Figure 9E). There was no significant effect of entrance type, nor was there an interaction, for the high selectivity group. The non-tapper group's stimulus-head movement period matching varied as a function of entrance period [$\chi^2(3) = 10.08$, $p < .05$], with changes in mean SMR across entrance periods containing marginally significant linear [$t(1562) = 2.38$, $p = .05$)] and quadratic [$t(1562) = -1.89$, $p = .09$)] components in both simultaneous and staggered conditions (Figure 9F). The trend of the SMR across entrance periods in non-tappers was similar to that seen in the simultaneous conditions for the low selectivity and high selectivity groups. There was no significant effect of entrance type, nor was there an interaction, for the non-tapper group.

Post-stimulus ratings. Participant's mean ratings of the music's groove (Figure 10A), enjoyment of the music (Figure 10B), and the extent to which participants wanted to continue with the trial (Figure 10C) were all higher for the staggered condition than for the simultaneous condition [music grooved: $t(70) = 12.51$, $p < .0001$; enjoyed music: $t(70) = 13.84$, $p < .0001$; wanted to continue: $t(70) = 15.59$, $p < .0001$].

As displayed in Table 1 (Experiment 2 columns), all three post-stimulus ratings correlated positively and significantly with participants' trial-wise tapping rates (correlation coefficients = .38 – .48, all *p*-values < .001), mean head energy (correlation coefficients = .12 – .18, all *p*-values < .001), and stimulus-head movement synchronization (correlation coefficients = .08 – .17, all *p*-values < .001). Perceived groove ratings were correlated most strongly with each of the above movement measures, and enjoyment ratings were correlated second most strongly with each of those movement measures. There were no significant correlations between any post-stimulus questions and trial-wise mean tapping SMR.

Discussion

This experiment assessed the amount and quality of spontaneous tapping and head movement elicited by multipart music as a function of when instrument parts entered relative to each other. As predicted, individuals tapped more when instrument entrances were staggered, complementing the observed increases in groove ratings following staggered instrument entrances in Experiment 1. As tapping rates increased, the synchronization between participants' tapping and the stimulus' rhythmic structure also generally increased. The amount of tapping differed among individuals such that some listeners were not very selective in their tapping: they tapped to most stimuli and at high rates, whereas others were more selective in the music that they chose to tap to, and still others tapped very little or not at all. Interestingly, the relatively few trials that participants in the non-tapper group tapped to became more synchronized to the stimulus following staggered entrances, whereas their tapping became less synchronized throughout the course of stimuli with simultaneous entrances. This suggests that individuals in the non-tapper group were unable or unmotivated to maintain a consistent level of stimulus-tapping synchrony throughout the duration of the trial, whereas synchrony was somehow facilitated as staggered instrument entrances occured. Possible reasons for the difference between individuals in the non-tapper group and the other groups is given in the General Discussion.

Groupwise differences in movement selectivity were consistent across movement modalities, suggesting that differences in the magnitudes of movement among participants were not peculiar to specific modes of measurement. Tapping and head movements were recorded with different devices (drum pad versus motion capture) which represented movements much differently (discretely versus continuously).

We varied the metrical salience of entrance times between metrically weak and metrically strong staggered conditions, but found no reliable difference in movement responses. This temporally focal metricality manipulation was very subtle. Any attendant expectancy violations or processing decrements that are associated with less salient time points had an inconsequential effect on the macroscopic behaviors examined here.

Participants' perceived groove, enjoyment, and desire to continue listening were all higher for music with staggered instrument entrances than for music with simultaneous instrument entrances, corroborating the results of Experiment 1, and further suggesting a general increase in aesthetic and affective experience when instrument part onsets were temporally separated. Ratings of perceived groove, enjoyment, and wanting to continue were all positively correlated with the magnitude and synchronization of movements with music (except for the tapping SMR). Perceived groove was most strongly associated with movement magnitude and synchronization, providing further evidence that spontaneous movement is an important aspect of groove. Extending the results of Janata et al. (2012), who reported a strong association between enjoyment and groove, the current results show a similar relationship between enjoyment and the magnitude and synchronization of movement.

Effects of Musical Training on Perceived Groove and Spontaneous Movement

In Experiments 1 and 2 we found that staggered instrument entrances in multipart music led to increases in perceived groove and spontaneous movements. Given that both experiments involved participants with and without musical training, we were able to assess whether the effects we observed varied as a function of participants' musical training. We did not expect to find notable differences between musicians' and non-musicians' responses to multipart music with staggered and simultaneous instrument entrances, and for this reason we did not recruit separate groups *a priori*.

We examined whether musical training (*musically untrained* = fewer than two years of training; *musically trained* = more than two years of training) influenced perceived groove (Experiment 1) or spontaneous tapping and head movements (Experiment 2) either directly or by interacting with manipulations of entrance type (simultaneous or staggered) and musical scene density (full, reduced, or solo; Experiment 1 only). Seven of the 15 participants in Experiment 1 had less than two years of musical training (*M* ± *SD:* 0.14 ± 0.38). The remaining 8 participants had an average of 5.62 ± 3.89 y of training. In Experiment 2, 22 of 59 participants had less than two years of training (0.23 ± 0.43 y), and the remaining 37 participants had 6.19 ± 3.94 y of training on average.

Results

Groove ratings. Influences of entrance type and scene density on dynamic groove ratings were similar to those observed in Experiment 1, regardless of musical training (Figure 11). Musical training did not interact with entrance type [$\chi^2(1) = 0.48$, *n.s.*], nor did we find an Entrance Type X Entrance Interval X Musical Training interaction [$\chi^2(3) = 4.16$, *n.s.*], suggesting that the effect of staggered entrances on groove rating dynamics did not differ between musically trained and untrained listeners. However, we observed an Entrance Interval X Musical Training interaction [$\chi^2(3) = 10.06$, $p <.05$], such that the slope of the groove rating increase was steeper among untrained than trained participants during the initial entrance interval of a stimulus [$t(39) = 3.08$, $p < .05$], although both groups increased their ratings similarly across subsequent entrance intervals. The effect of musical scene density did not interact with musical training [$\chi^2(2) = 0.06$, *n.s.*]. However, when disregarding entrance type and scene density manipulations, untrained listeners produced higher mean groove ratings (as averaged over the last 10 s of trials) ($M = 69.52$, $SE = 1.74$) than did musically trained listeners ($M = 53.93$, $SE = 1.61$) [$\chi^2(1) = 5.64$, $p <.05$].

Movement. Spontaneous movement behaviors in both groups of participants were largely consistent with the aggregate movement-related results of Experiment 2 reported above. Participants from both musical training backgrounds were distributed across the movement selectivity groups (Table 2), though untrained individuals were more concentrated in the high selectivity than the low selectivity group, whereas the reverse was true for participants with musical training. The non-tapper group included participants from both backgrounds.

Movement magnitude. Figure 12A illustrates participants' time-varying tapping rates as a function of entrance type (staggered or simultaneous) and musicianship (untrained or trained). The effect of entrance type on participants' overall tapping rates differed between musically trained and untrained individuals [Entrance Type X Musical Training interaction: $\chi^2(1) = 9.11$, $p < .01$], such that musically trained participants' trial-wise tapping rates (averaged across entrance periods) were lower for the staggered than the simultaneous condition [$t(9350) = -2.46$, $p < .05$], whereas untrained participants' trial-wise tapping rates were higher for the staggered than the simultaneous condition [this difference approached significance: $t(9350) = 1.91$, $p = .056$]. However, time-varying tapping rates revealed that stimuli with staggered instrument entrances elicited a greater increase in participants' tapping rates as trials progressed than did stimuli with simultaneous entrances, regardless of whether participants were musically trained [Entrance Type X Entrance Period interaction: $\chi^2(3) = 80.38$, $p < .0001$]. Participants initially produced lower tapping rates for the staggered than for the simultaneous condition during baseline [$t(9350) = -5.19$, $p < .0001$] and Entrance 2 [$t(9350) = -3.28$, $p < .01$] periods, whereas tapping rates for the staggered condition increased beyond those for the simultaneous condition during the Entrance 3 [$t(9350) = 2.57$, $p < .05$] and Entrance 4 [$t(9350) = 5.91$, $p < .0001$] periods. This interaction effect did not differ between musically trained and untrained individuals [Entrance Type X Entrance Period X Musical Training interaction: $\chi^2(3) = 3.42$, *n.s.*]. Moreover, overall tapping rates did not differ significantly between musically trained and untrained participants [$\chi^2(1) = 2.8$, *n.s.*].

The magnitude of head movements in response to the entrance type manipulation did not differ significantly between musically trained and untrained individuals [Entrance Type X Musical Training interaction: $\chi^2(1) = 3.37$, *n.s.*], nor did the changes in movement energy between entrance periods [Entrance Period X Musical Training interaction: $\chi^2(3) = 0.71$, *n.s.*] (Figure 12B). Both musically trained and untrained individuals produced less head movement during the early portion of staggered compared to simultaneous trials, but the amount of head movement increased during later entrance periods under the staggered condition to reach similar levels of head movement as in the simultaneous condition. However, the interaction of entrance type and entrance period only approached significance [$\chi^2(3) = 7.70$, $p = .052$]. The magnitude of participants' head movements in response to entrance conditions across time did not differ between musically trained and untrained individuals [$\chi^2(3) = 0.49$, *n.s.*], and overall amount of head movement did not differ between musically trained and untrained participants [$\chi^2(3) = 0.57$, *n.s.*].

Stimulus-movement synchronization. Figure 12C illustrates the degree to which participants synchronized their tapping with stimuli as a function of entrance type and musical training. Musically trained participants better synchronized their tapping to the temporal structure of stimuli (as indexed via SMR) than untrained participants [$\chi^2(1) = 6.21$, $p < .05$]. Furthermore, the quality of participants' stimulus-tapping synchronization differed between simultaneous and staggered entrances as a function of their musical training [Entrance Type X Musical Training interaction: $\chi^2(1) = 4.15$, $p < .05$], such that musically trained individuals better synchronized their tapping on a trial-wise basis with simultaneous stimuli than with staggered stimuli [$t(4387) = 4.12$, $p < .0001$], whereas untrained participants' trial-wise stimulus-tapping synchronization did not differ between entrance types [$t(4397) = 0.46$, *n.s.*]. Moreover, although we observed a significant Entrance Period X Musical Training interaction [$\chi^2(3) = 10.27$, $p < .05$], suggesting that participants' stimulus-tapping synchronization increased differently between successive entrance periods as a function of musical training, polynomial contrasts revealed that synchronization increased linearly between entrance periods for both trained and untrained individuals, albeit with a stronger linear trend for musically trained [$t(4392) = 12.63$, p

$< .0001$] than for untrained participants [$t(4391) = 5.16$, $p < .0001$]. There was no significant Entrance Type X Entrance Period X Musical Training interaction [$\chi^2(3) = 1.27$, *n.s.*].

Figure 12D presents participants' stimulus-head movement synchronization as a function of entrance type and musical training. Overall, participants did not differ in stimulus-head movement synchronization as a function of training [$\chi^2(1) = 2.56$, *n.s.*], nor was there an interaction of entrance type and training on stimulus-head movement synchronization measure [Entrance Type X Musical Training interaction: $\chi^2(1) = 0.015$, *n.s.*]. Synchronization of head movements with stimuli increased steadily across the course of trials under the staggered condition but tapered toward the end in the simultaneous condition, regardless of musical training [Entrance Type X Entrance Period interaction: $\chi^2(3) = 16.58$, $p < .001$]. We did find a significant Entrance Period X Musical Training interaction [$\chi^2(3) = 9.83$, $p < .05$], likely due to a steeper slope in the rate of improved synchronization among participants with training.

Intriguingly, untrained participants were at a synchronization disadvantage compared to musically trained participants for tapping, but not for head movements. This was reflected by a Movement Type X Musical Training interaction [$\chi^2(1) = 4.14$, $p < .05$], and significantly lower synchronization quality among untrained compared to trained individuals during tapping [$t(90) = -2.90$, $p < .01$], but not during head movements [$t(58) = -1.72$, *n.s.*].

Discussion

Across both experiments, effects pertaining to staggered instrument entrances and the density of musical auditory scenes were consistent across musically trained and untrained individuals, suggesting that effects of incremental ramping of stimulus information on listeners' perception of groove and overt sensorimotor coupling do not depend on musical training. Nevertheless, non-musicians tended to rate stimuli as higher in groove than musicians, and do so more quickly, suggesting that musically trained participants were perhaps more critical of the stimuli and thus more conservative in their assessment of groove despite a greater propensity on their part to tap along with the music.

Musically trained participants synchronized their tapping to the temporal structure of stimuli better than did untrained participants, but both trained and untrained participants synchronized their head movements equally well to. Tapping one's hands on a drum pad could be conceptualized as a form of musical instrument performance, an activity in which musically trained participants were presumably more experienced, and one that allows for more complex rhythmic behavior than does head movement. Conversely, moving one's head with a musical rhythm is a subtler entrainment behavior that is likely performed by music listeners of diverse musical (and non-musical) backgrounds, allowing untrained participants to more comfortably synchronize with the music in this way.

General Discussion

We performed two experiments designed to examine the intersection of perception, action, attention, and affect in the context of music and the psychological construct of groove. We manipulated the number of concurrent instrument parts and, more importantly, their entrance timing while assessing individuals' spontaneous movements and subjective assessments of groove and experienced affect. As predicted, listeners generally exhibited increased stimulus-coupled movement and reported a greater degree of perceived groove when instrument entrances were staggered across time and when musical scenes comprised multiple instruments, regardless of whether participants were musically trained. We also provided preliminary evidence for individual differences in participants' motoric engagement with music. In the remainder of this paper, we discuss potential mechanisms that may induce spontaneous movement with music.

Neuroimaging evidence suggests that humans' proclivity to move with music arises from interactions between the brain's auditory and motor systems during auditory rhythm processing (Zatorre, Chen, & Penhune, 2007). Brain regions important for motor control and implementation, including the basal ganglia, supplementary motor area, and premotor cortex are especially sensitive to auditory rhythms containing or inducing the perception of a regular beat and become engaged when one listens to such stimuli, even in the absence of motor demands (Bengtsson et al., 2009; Chen, Penhune, & Zatorre, 2008; Grahn & Brett, 2007; Grahn & Rowe, 2009). Beat-based rhythms have also been shown to selectively entrain neuronal populations (Nozaradan et al., 2011; Nozaradan, Peretz, & Mouraux, 2012) and to synchronize modulations of neural beta-band oscillations to the stimulus' rhythm (Fujioka, Trainor, Large, & Ross, 2012). Given the propensity for functional coupling of auditory and motor regions in humans, how might the apparent strength of the functional coupling be modulated in complex musical scenes?

We propose that engagement of the brain's arousal and attention systems may mediate auditory-motor synchronization and facilitate action. The entrance of instruments into the musical auditory scene increases physiological arousal (Grewe et al., 2007) and periodic entrances and ramping of intensity are musical devices that can maintain a listener's attention (Huron, 1992). Studies in audition have demonstrated that newly appearing objects in a scene pop out and draw attention (Cervantes Constantino et al., 2012). When attention is directed to a stream within a complex auditory or musical scene, listeners are better able to discriminate changes in the stream than when attention is undirected (Eramudugolla, Irvine, McAnally, Martin, & Mattingley, 2005; Janata, Tillmann, & Bharucha, 2002). Given that attending to auditory objects enhances sensory processing in auditory cortex (e.g., Fritz, Elhilali, & Shamma, 2007; Petkov et al., 2004; Woldorff et al., 1993), orienting of attention to an entering instrument part would be expected to strengthen responses to features of that part. To the extent that those features enhance beat saliency/pulse clarity or increase perceived event density, features known to correlate with groove and movement characteristics (Burger et al., 2013; Madison et al., 2011), the periodic drive of auditory-motor loops would be expected to increase.

An alternative explanation for the observed effects of staggered entrances is that of expectancy. The violation and fulfillment of musical expectations are known to facilitate emotional responses such as surprise, awe, pleasure, and disappointment (Huron, 2006; Juslin & Västfjäll, 2008). As one listens to music with initial patterns clearly repeating, as participants did in the present experiments, one may develop the expectation that some aspect of the music will change. The entrance of a new instrument part into the musical scene may thus generate a reward as the expectation of change is fulfilled, leading to increased affective response and engagement with the music, while the continuation of simultaneous repetitive parts may leave one's expectation of change unfulfilled, leading to a decrease in positive affective response and musical engagement. A test of this hypothesis could entail musical stimuli in which musical changes other than part additions occur periodically, such as switching from one timbre to another.

We observed individual differences in the way that listeners engaged motorically with the music, depending on both musical background and the selectivity with which individuals produce stimulus-coupled movements. In terms of movement selectivity, some participants tapped at high rates to most musical excerpts, others tapped more selectively to some excerpts but not others, and some tapped very little or not at all throughout the experiment. Although the relevant data were not gathered, it could be that the particular instruments or genres in which participants were experienced directed their tapping behavior (Martens, 2011). We observed similar differences in the magnitude of spontaneous head movement. Differences in the magnitude of listeners' movements may have stemmed from differences in motivation or self-consciousness. Variation in the amount of music-evoked movement among individuals may have also been related to differences in participants' personality traits, as Luck, Saarikallio, Burger, Thompson, and Toiviainen (2010) found that different personality traits were associated with different patterns of music-evoked movement. In terms of non-tapping individuals, some of these listeners may have found synchronizing movements to stimulus periodicities difficult. Although non-tapping participants achieved similar levels of SMR as other groups during at least some periods of time, the lowest mean SMR for any entrance period was that of non-tapping participants, as observed in both tapping and head movement. Moreover, some individuals experience considerable difficulty synchronizing movements to a beat, a phenomenon referred to as "beat deafness" (Phillips-Silver et al., 2011). Given that Janata et al. (2012) found that participants tended to move less and rate music lower in groove when they found synchronization to be difficult, individuals in Experiment 2 may have been less inclined to move with the music if they found it difficult to synchronize movements with the music. On the other hand, the percentage of musically trained individuals (whose training presumably entailed some aspect of movement timing) classified as non-tappers was slightly higher than that of untrained participants, calling into question the possibility that synchronization difficulty underlies some participants' absence of tapping.

Aside from movement selectivity, we observed differences in the way that musically trained and untrained individuals moved with the music. Most notably, untrained participants synchronized their tapping to stimuli less accurately than did musically trained participants. However, both untrained and trained participants synchronized their head movements to stimuli equally well. This suggests that non-musicians may show task-specific sensorimotor synchronization deficits compared to musicians, and that such specificity should be considered when drawing conclusions about general sensorimotor abilities of non-musicians versus musicians.

Conclusion

In support of our hypothesis, we found multipart music in which instrument parts enter in a staggered sequence, as opposed to music in which instruments begin playing simultaneously, to evoke increased spontaneous sensorimotor coupling, as observed in spontaneous tapping and head movements, as well as reports of increased groove, enjoyment, and wanting the music to continue. These results support theoretical assertions that incrementally increasing stimulus information increases musical engagement and provide, for the first time, evidence that such high-level structural organization principles in music influence motoric engagement. Our results thus provide initial empirical insights into the relationship between complex musical scenes, attentional engagement as manifested in spontaneous overt sensorimotor coupling, and emotion.

References

Bates, D. M., Maechler, M., & Bolker, B. (2013). lme4: Linear mixed-effects models using S4 classes (Version 0.999999-2). Retrieved from http://CRAN.R-project.org/package=lme4

Bengtsson, S. L., Ullen, F., Ehrsson, H. H., Hashimoto, T., Kito, T., Naito, E., . . . Sadato, N. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex, 45*(1), 62-71.

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.

Bharucha, J. J., & Pryor, J. H. (1986). Disrupting the isochrony underlying rhythm: An asymmetry in discrimination. *Perception & Psychophysics, 40*(3), 137-141.

Bland, J. M., & Altman, D. G. (1995). Calculating correlation coefficients with repeated observations: Part 1--Correlation within subjects. *BMJ, 310*(6977), 446.

Brockmole, J., & Henderson, J. (2005). Object appearance, disappearance, and attention prioritization in real-world scenes. *Psychonomic Bulletin & Review, 12*(6), 1061-1067.

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S., & Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology, 4*, 183.

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.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to Musical Rhythms Recruits Motor Regions of the Brain. *Cerebral Cortex, 18*(12), 2844-2854.

Cole, G. G., & Kuhn, G. (2009). Attentional capture by object appearance and disappearance. *The Quarterly Journal of Experimental Psychology, 63*(1), 147-159.

Eerola, T., & Toiviainen, P. (2004). MIDI Toolbox: MATLAB Tools for Music Research. Kopijyvä, Jyväskylä, Finland: University of Jyväskylä. Retrieved from http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/miditoolbox/

Eramudugolla, R., Irvine, D. R. F., McAnally, K. I., Martin, R. L., & Mattingley, J. B. (2005). Directed attention eliminates 'Change deafness' in complex auditory scenes. *Current Biology, 15*(12), 1108-1113.

Fritz, J. B., Elhilali, M., & Shamma, S. A. (2007). Adaptive changes in cortical receptive fields induced by attention to complex sounds. *Journal of Neurophysiology, 98*(4), 2337-2346.

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *The Journal of Neuroscience, 32*(5), 1791-1802.

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

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *The Journal of Neuroscience, 29*(23), 7540-7548.

Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Emotions over time: Synchronicity and development of subjective, physiological, and facial affective reactions to music. *Emotion, 7*(4), 774-788.

Huron, D. (1992). The ramp archetype and the maintenance of passive auditory attention. *Music Perception, 10*(1), 83-92.

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*: MIT Press.

Janata, P., Tillmann, B., & Bharucha, J. J. (2002). Listening to polyphonic music recruits domain-general attention and working memory circuits. *Cognitive, Affective, and Behavioral Neuroscience, 2*(2), 121-140.

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.

Jones, M. R., & Boltz, M. (1989). Dynamic Attending and Responses to Time. *Psychological Review, 96*(3), 459-491.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *The Behavioral and Brain Sciences, 31*(5), 559-575; discussion 575-621.

Keller, P. E., & Schubert, E. (2011). Cognitive and affective judgements of syncopated musical themes. *Advances in Cognitive Psychology, 7*, 142-156.

Kung, S.-J., Tzeng, O. J. L., Hung, D. L., & Wu, D. H. (2011). Dynamic allocation of attention to metrical and grouping accents in rhythmic sequences. *Experimental Brain Research, 210*(2), 269-282.

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., & Snyder, J. S. (2009). Pulse and Meter as Neural Resonance. *Neurosciences and Music III: Disorders and Plasticity, 1169*, 46-57.

Lenth, R. V. (2013). lsmeans: Least-squares means (Version 1.10-4). Retrieved from http://CRAN.R-project.org/package=lsmeans

London, J. (2004). *Hearing in time: Psychological aspects of musical meter*. New York, NY: Oxford University Press.

Luck, G., Saarikallio, S., Burger, B., Thompson, M. R., & Toiviainen, P. (2010). Effects of the Big Five and musical genre on music-induced movement. *Journal of Research in Personality, 44*(6), 714-720.

Madison, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception, 24*(2), 201-208.

Madison, G., Gouyon, F., Ullén, F., & Hörnström, K. (2011). Modeling the tendency for music to induce movement in humans: First correlations with low-level audio descriptors across music genres. *Journal of Experimental Psychology: Human Perception and Performance*.

Martens, P. A. (2011). The Ambiguous Tactus: Tempo, Subdivision Benefit, And Three Listener Strategies. *Music Perception, 28*(5), 433-448.

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. *The Journal of Neuroscience, 31*(28), 10234-10240.

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *The Journal of Neuroscience, 32*(49), 17572-17581.

Patterson, R. D. (1994). The sound of a sinusoid: Spectral models. *Journal of the Acoustical Society of America, 96*(3), 1409-1418.

Petkov, C. I., Kang, X., Alho, K., Bertrand, O., Yund, E. W., & Woods, D. L. (2004). Attentional modulation of human auditory cortex. *Nature Neuroscience, 7*(6), 658-663.

Phillips, D. P., Hall, S. E., & Boehnke, S. E. (2002). Central auditory onset responses, and temporal asymmetries in auditory perception. *Hearing Research, 167*(1–2), 192-205.

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piché, O., Nozaradan, S., Palmer, C., & Peretz, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia, 49*(5), 961-969.

R Core Team. (2012). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969-992.

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review, 20*(3), 403-452.

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.

Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition, 82*(2), 127-136.

Toiviainen, P., & Burger, B. (2013). MoCap Toolbox Manual. Jyväskylä, Finland: University of Jyväskylä. Retrieved from http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: hierarchical eigenmodes in music-induced movement. *Music Perception, 28*(1), 59-70.

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.

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2013). The impact of the bass drum on human dance movement. *Music Perception, 30*(4), 349-359.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America, 90*(18), 8722-8726.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience, 8*(7), 547-558.

Footnotes

---

[1] http://www.miditool.com

[2] Forty-nine out of the 59 participants reported familiarity with the term "groove" as applied to music. A multinomial regression revealed no significant relationship between familiarity with the term and membership in the tapping-selectivity groups formed in the analysis.

[3] This criterion was arbitrary, but we felt that it sufficiently distinguished participants who hardly tapped at all on most trials, from those who may not have tapped much on any given trial, but nonetheless tapped on a greater proportion of trials, or from those participants who tapped more intensively, but on very few trials.

[4] MATLAB code for the resonator model can be downloaded as part of the Janata Lab Music Toolbox (jlmt), within which the resonator model is referred to as Beyond the Beat (BTB), from http://atonal.ucdavis.edu/resources/software/jlmt/

Table 1

Correlations Between Self-Reported Ratings and Movement Measures

	Experiment 1	Experiment 2			
Rated item	Groove	Tapping Rate	Tapping SMR	MoCap Head SMR	MoCap Head SMR
df	13	57	57	53	53
Urge to move	.82***	-	-	-	-
Enjoyment	.73***	.42***	.01	.15***	.11***
Wanting to continue	.73***	.38***	.00	.12***	.08***
Perceived groove	-	.48***	.04	.18***	.17***

Note. Correlations between variations in movement measures and subjective ratings within participants were calculated by regressing out between-subject variability, as described by Bland and Altman (1995), in order to account for repeated observations. *df* = degrees of freedom. MoCap = motion capture. SMR = stimulus matching ratio. Dashes (-) indicate that a rated item was not collected for the experiment in which the dependent movement-related measure was taken.

*$p < .05$. **$p < .01$. ***$p < .001$.

Table 2

Percent of musically trained and untrained participants within movement selectivity groups

	Movement Selectivity		
Musical Training	High Selectivity	Low Selectivity	Non-Tapper
Untrained	59.09%	27.27%	13.64%
Trained	29.73%	51.35%	18.92%

Note. Musically trained participants were defined as those with two or more years of musical training. Movement selectivity was determined using the methods described in the *Preprocessing and Data Analysis* section of Experiment 2, which are also summarized in Figure 6.

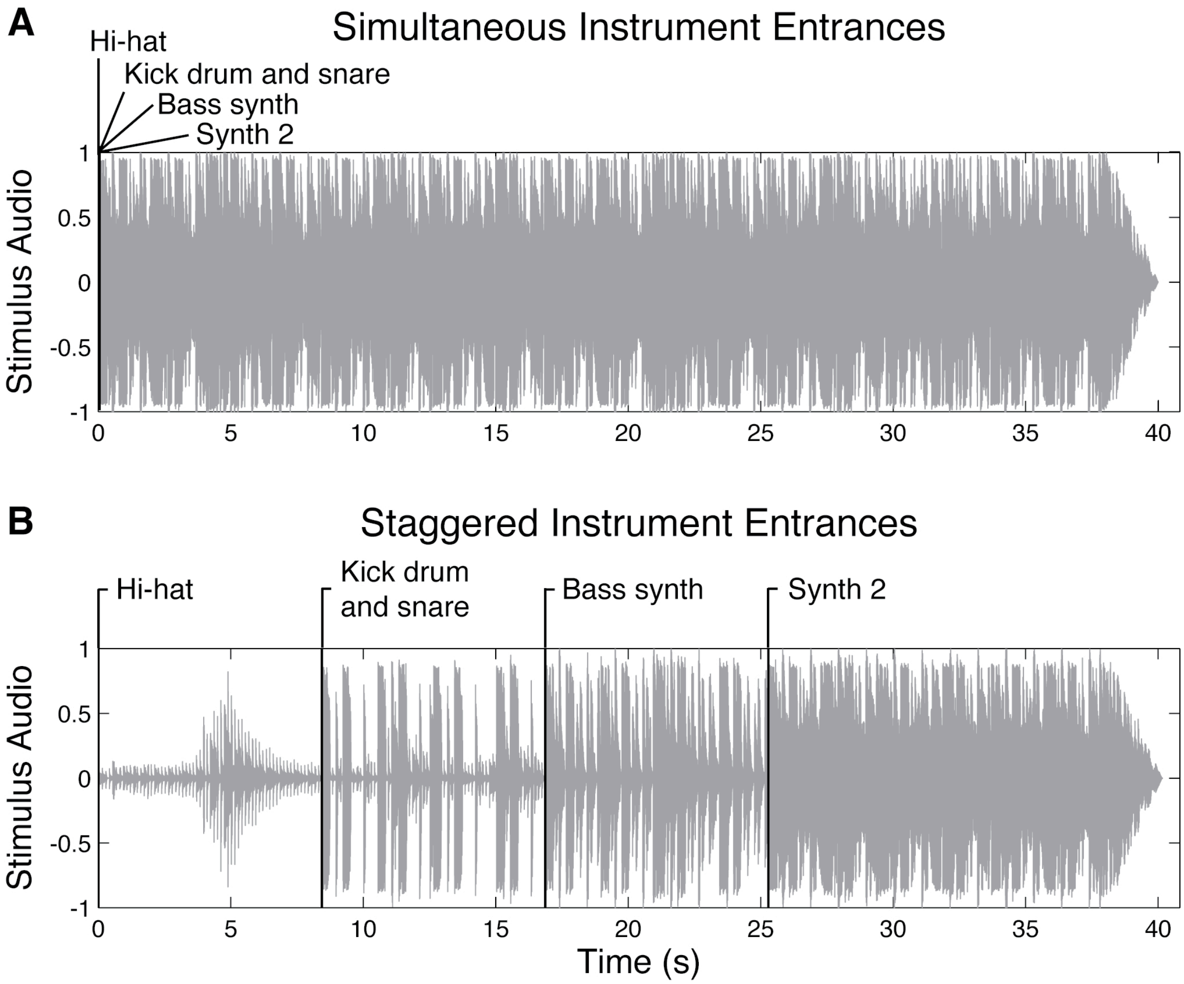

Figure 1. Instrument entrance conditions. Vertical and angled bars next to instrument names indicate the times at which instrument parts began playing. (A) Simultaneous instrument entrances. All instrument parts began playing simultaneously at the onset of the trial and continued to play concurrently throughout the trial. (B) Staggered instrument entrances. One instrument began playing at the onset of the trial, and subsequent instruments began playing at temporally distributed time points until the musical scene was populated with all instruments. Note that this figure corresponds to the entrance conditions of a single exemplar out of the 10

exemplars used in Experiment 1. The particular instruments and the order in which they entered differed among exemplars for both Experiments 1 and 2.

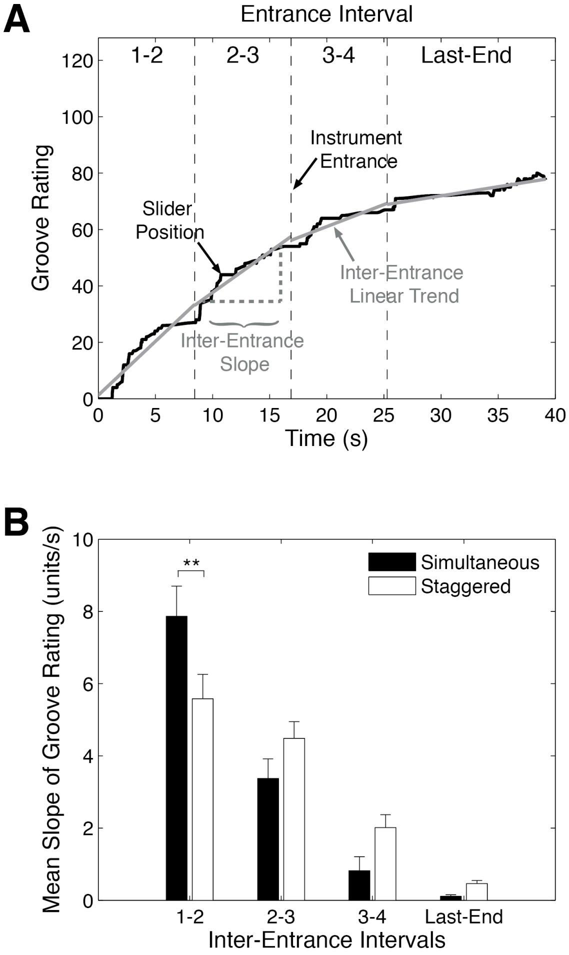

Figure 2. Analysis and results of groove ratings among entrance intervals. (A) Linear trends of slider movement were fit between each instrument entrance, and between the final entrance and the end of the stimulus. We measured the slope of each fitted line to assess changes in groove ratings following entrances. (B) Mean slope of continuous groove ratings across entrance intervals as a function of simultaneous and staggered instrument entrances. Note that some of the stimuli contained 3 instrument parts, whereas others contained 4 parts. Thus, the 3-4 entrance interval applies only to 4-part stimuli, whereas last-end refers to the interval between the final

entrance (third entrance for 3-part and fourth entrance for 4-part stimuli) and the end of the stimulus. Error bars indicate standard error of the mean (*SEM*). ** $p < .01$

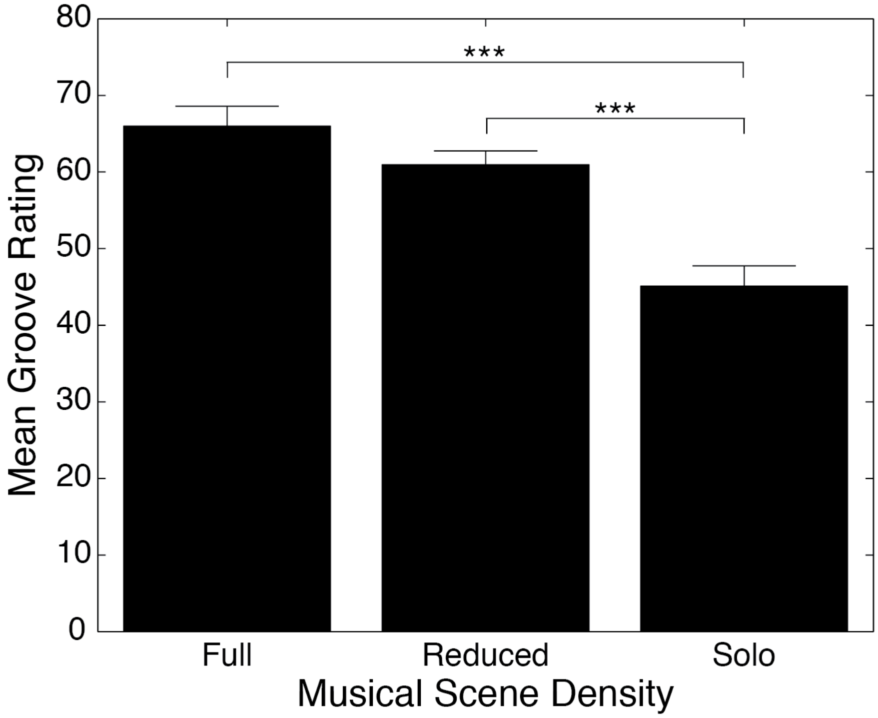

Figure 3. Mean groove rating during the final 10 s of trials as a function of the number of instrument parts present (*full* = 3 instruments for 3-part stimuli and 4 instruments for 4-part stimuli, *reduced* = 2-3 instruments, *solo* = 1 instrument). Error bars indicate *SEM*. *** $p < .001$

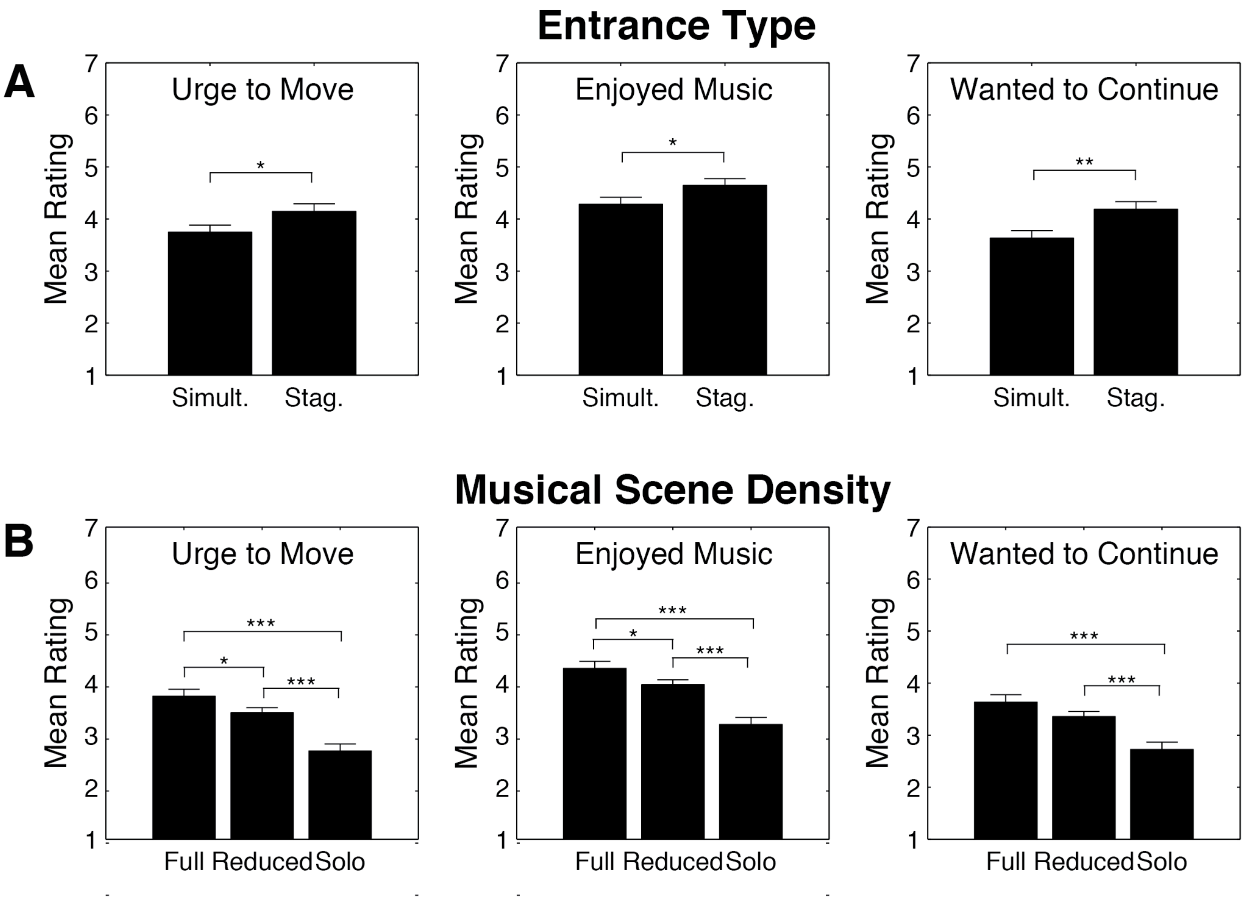

Figure 4. (A) Mean self-reported urge to move to the music (left), enjoyment of the music (center), and wanting to continue listening to the music (right) as a function of simultaneous and staggered instrument entrances (simult. = simultaneous, stag. = staggered). (B) Mean self-reported urge to move (left), enjoyment of the music (center), and wanting to continue listening to the music (right) as a function of musical scene density density (full, reduced, or solo instrumentation). Error bars represent *SEM*. * $p < .05$, ** $p < .01$, *** $p < .001$

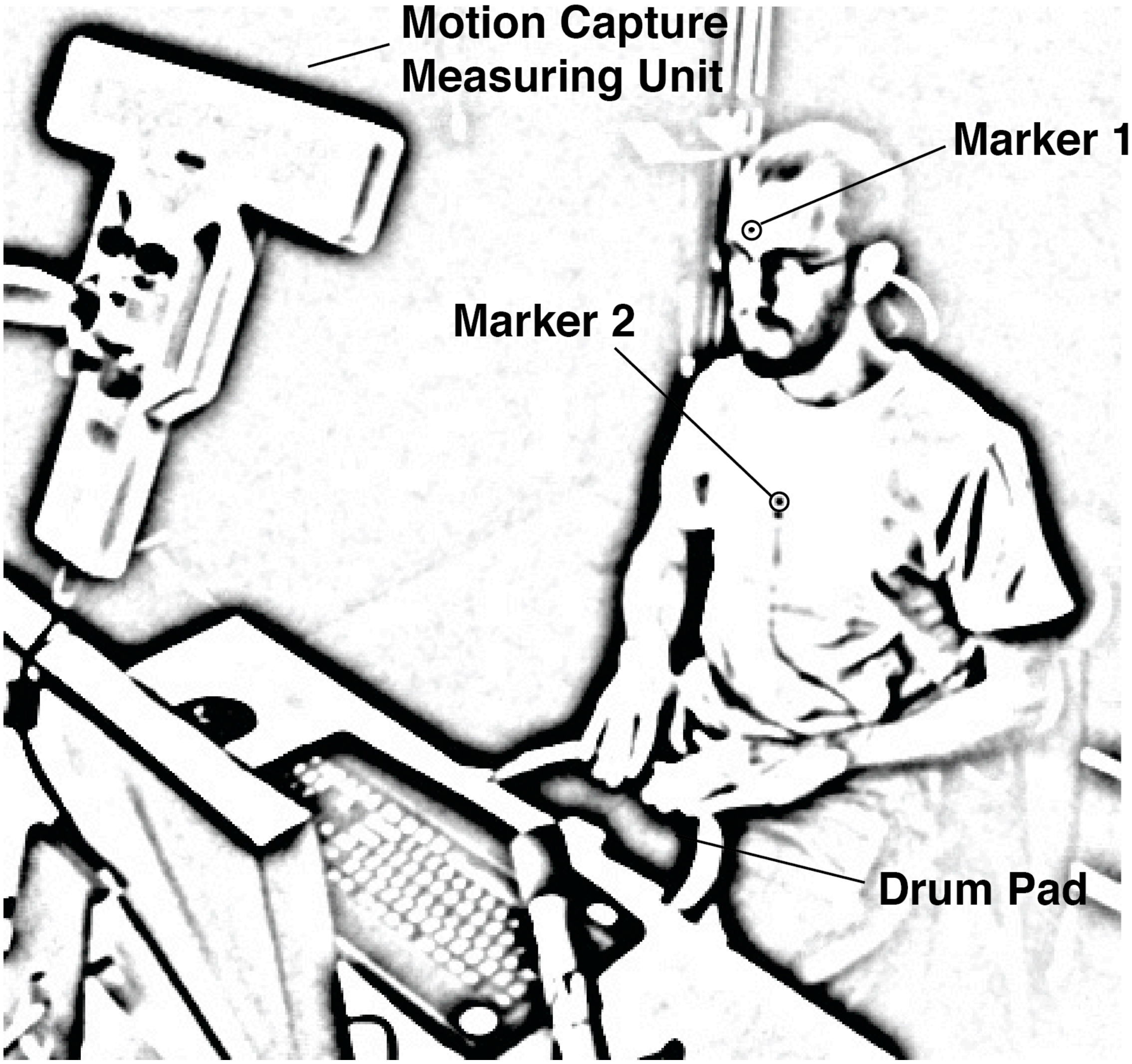

Figure 5. Arrangment of movement acquisition apparatus used in Experiment 2. Participants sat with their hands resting on a drum pad and were given the option to begin tapping if the music compelled them to do so. Spontaneous head movements were recorded using an ultrasonic motion capture system, which consisted of 2 ultrasound emitting markers (placed on forehead and chest) and a measuring unit. (Photograph used with permission of participant).

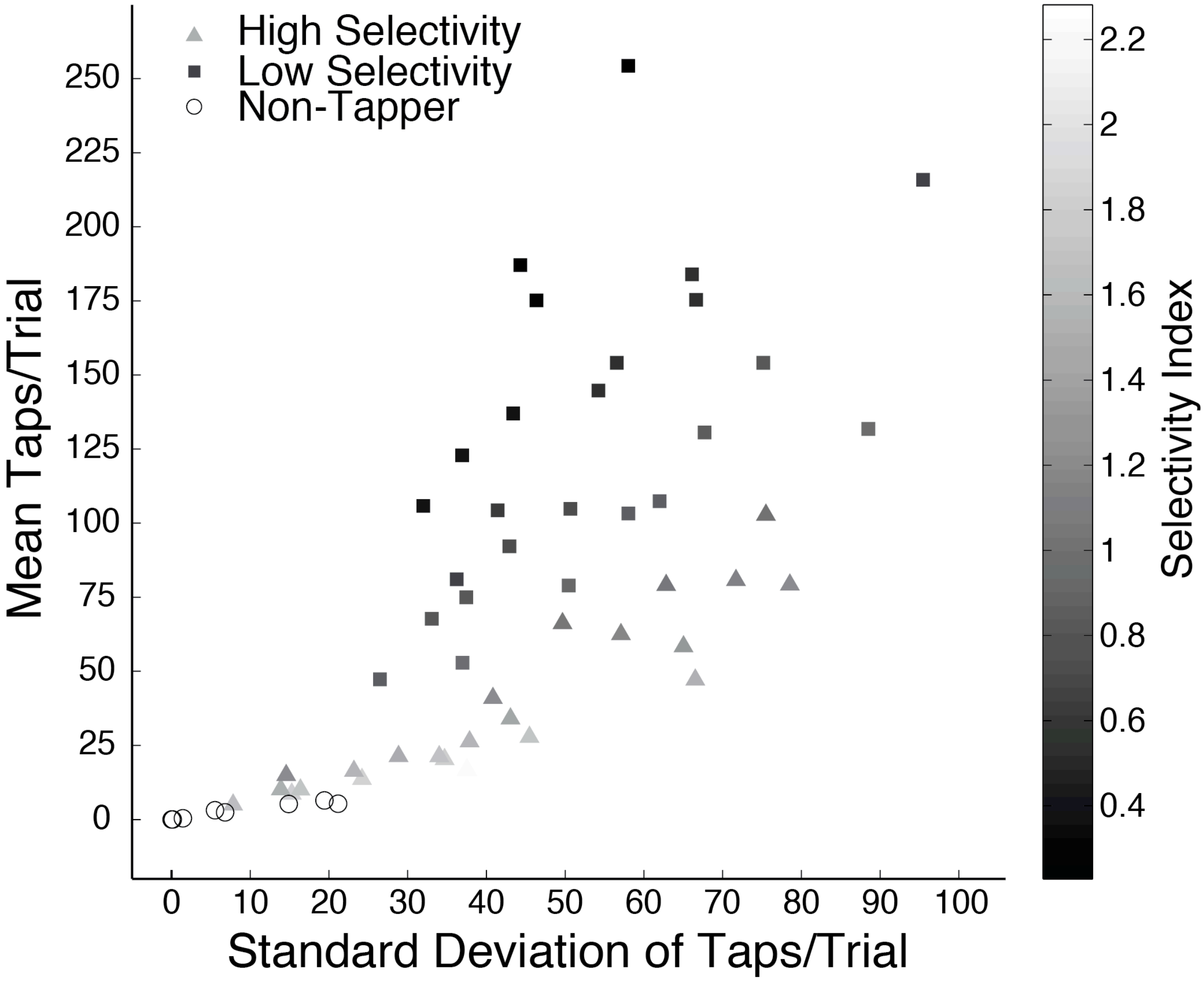

Figure 6. Individual differences in tapping selectivity. Each data point corresponds to a participant. Participants who produced fewer than 10 taps on ≥ 85% of trials were designated as non-tappers. Selectivity indices were calculated for the remaining participants by taking the ratio of $SD_{\text{taps per trial}}/M_{\text{taps per trial}}$. Participants with selectivity indices greater than the median were classified as high selectivity tappers, whereas participants with selectivity indices less than or equal to the median were classified as low selectivity tappers. Note that participants with similar means and standard deviations could be classified in either the non-tapper or high selectivity group depending on whether their tapping was restricted to < 15% of trials, thus giving rise to the close proximity between some non-tapper and low-selectivity data points.

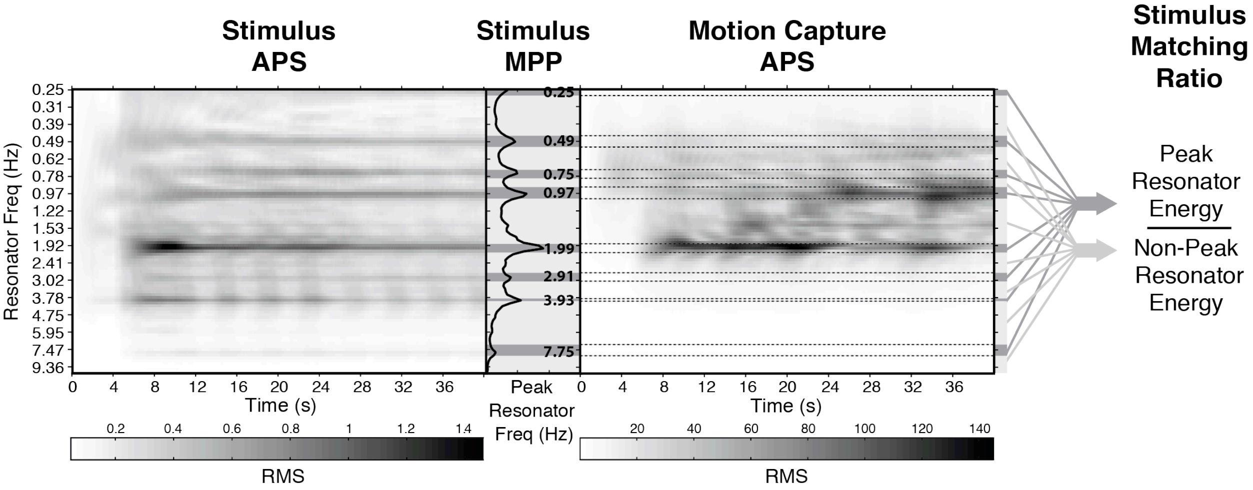

Figure 7. Calculation of a stimulus matching ratio (SMR) as a metric for stimulus-movement synchronization quality. The average periodicity surface (APS) is akin to an amplitude spectrogram, illustrating the root mean square (RMS) of each reson filter's output. The mean periodicity profile (MPP) illustrates an excerpt's average periodicity structure. Stimulus-related periodicities were identified as peaks in the MPP (labeled as "peak resonator energy"). The temporal structure of participants' spontaneous movements (tapping and head movements) was also modeled, and the degree of synchronization between a participant's movements and the corresponding stimulus on any given trial was calculated by taking the ratio of the participant's movement energy within peak (stimulus-related) resonator frequencies to movement energy in non-peak frequencies. A larger SMR indicates greater stimulus-movement synchronization.

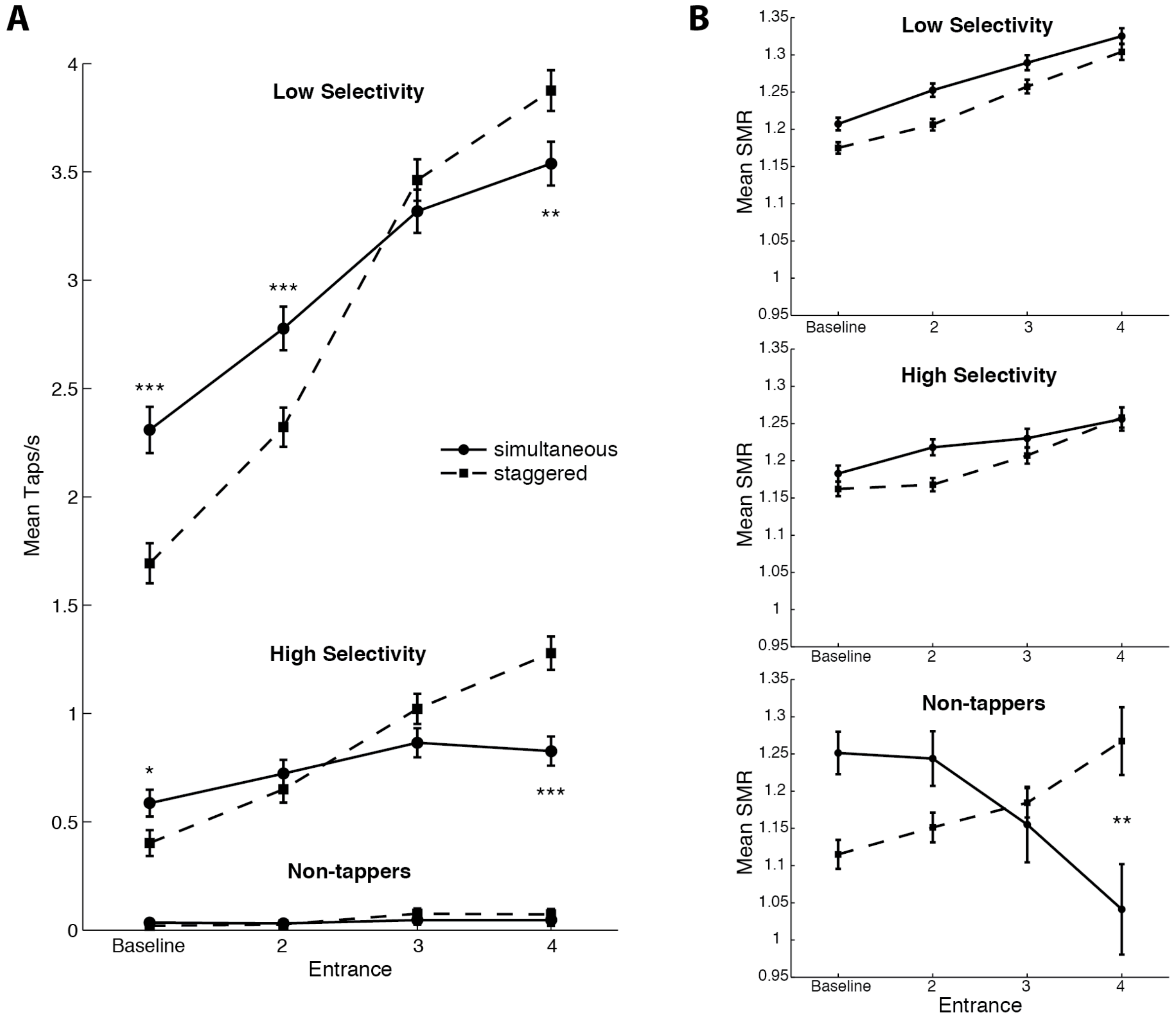

Figure 8. Spontaneous tapping. (A) Mean tapping rate as a function of instrument entrance type (simultaneous and staggered) and entrance period for low selectivity (top), high selectivity (center), and non-tapper (bottom) groups. (B) Mean stimulus matching ratio (SMR) as a function of entrance type and entrance period for low selectivity (top), high selectivity (center), and non-tapper (bottom) groups. In both A and B panels, baseline refers to the 1 s period preceeding Entrance 2, and Entrances 2, 3, and 4 corresponds to 3.89 s periods following each of those entrances. Error bars indicate *SEM*. * *p* < .05, * *p* < .01, *** *p* < .001

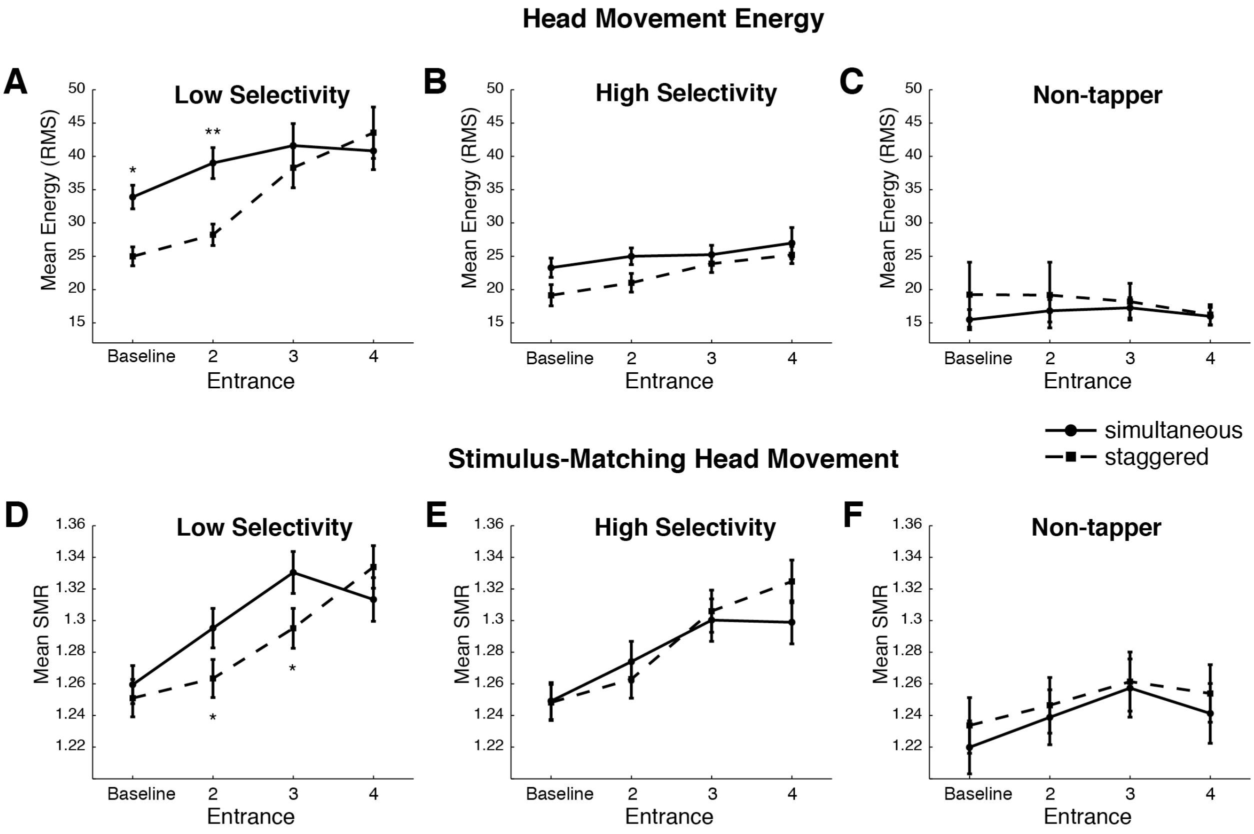

Figure 9. Spontaneous head movements as recorded by motion capture. (A-C) Head movement energy (root mean square of motion capture signal) as a function of instrument entrance type (simultaneous and staggered) and entrance period for (A) low selectivity, (B) high selectivity, and (C) non-tapper groups. (D-F) Mean stimulus-head movement stimulus matching ratio (SMR) as a function of entrance type and entrance period for (D) low selectivity, (E) high selectivity, and (F) non-tapper groups. Entrance periods are the same as described in Figure 8. Error bars indicate *SEM*. * $p < .05$, ** $p < .01$

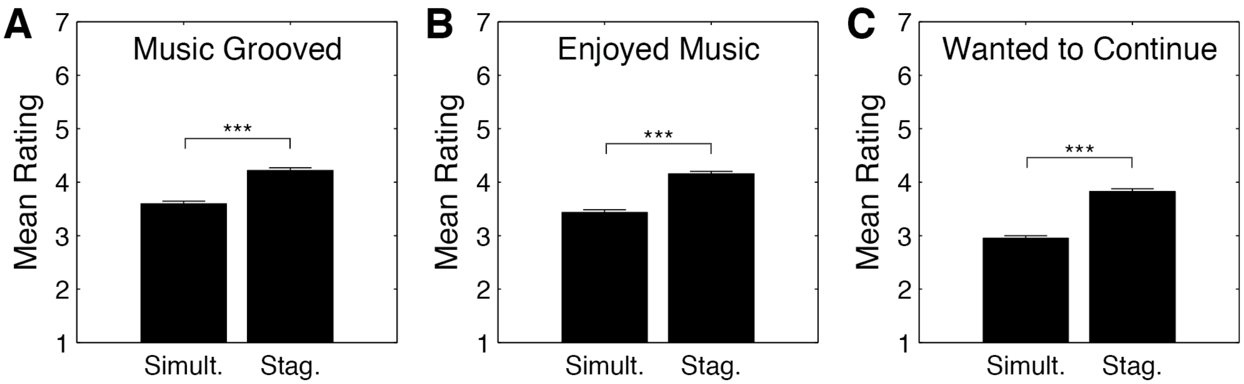

Figure 10. Mean self-reported degree to which participants (A) felt that the music grooved, (B) enjoyed the music, and (C) wanted to continue listening to the music as a function of simultaneous and staggered instrument entrances (simult. = simultaneous; stag. = staggered). Error bars indicate *SEM*. *** $p < .001$

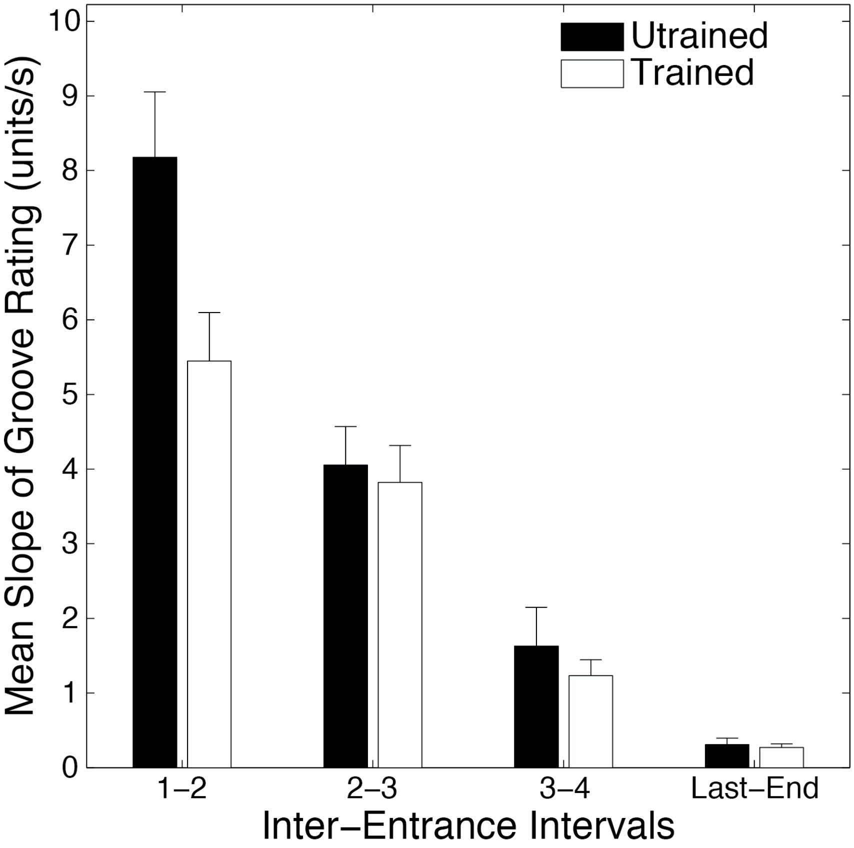

Figure 11. Effects of musical training on the rate of increase in groove rating between successive instrument entrances (inter-entrance intervals).

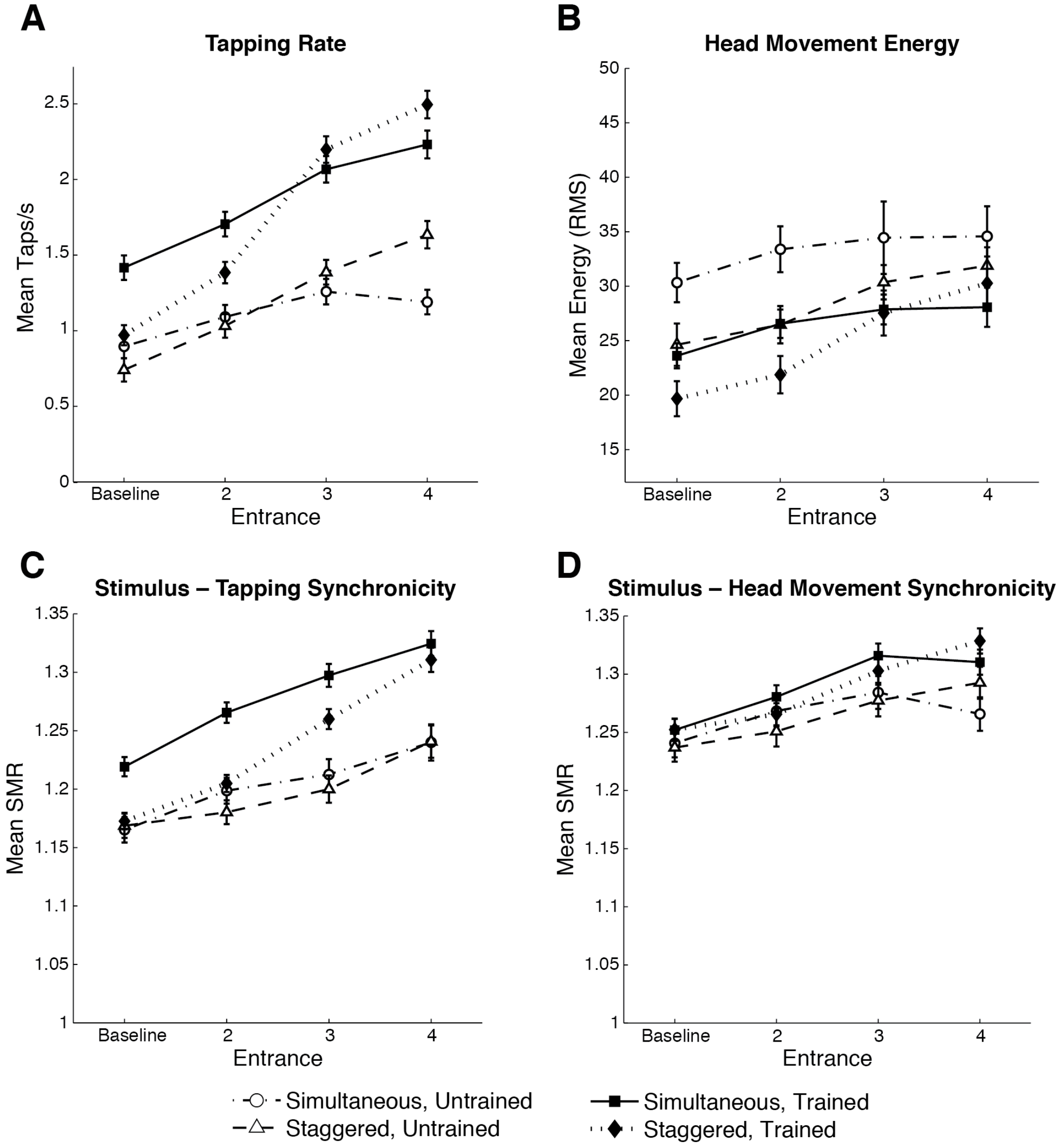

Figure 12. Effects of musical training on the magnitude of participants' spontaneous (A) tapping and (B) head movements (the latter assessed as RMS of motion capture recorded from head movements), as well as the degree to which participants' (C) tapping and (D) head movements were synchronized to stimuli (SMR stimulus matching ratio).

Supplemental Material accompanying, “Spontaneous Sensorimotor Coupling with Multipart Music,” by Hurley, Martens and Janata.

Audio examples. Stimuli with simultaneous entrances and staggered entrances at metrically strong locations are labeled *Sim* and *Strong*, respectively.

Club2-Sim.mp3 – the stimulus depicted in Figure 1A.

Club2-Stag.mp3 – the stimulus depicted in Figure 1B.

Dance1-Sim.mp3 – used in Experiment 2

Dance1-Strong.mp3 – used in Experiment 2

Rock3-Sim.mp3 – used in Experiment 2

Rock3-Strong.mp3 – used in Experiment 2

Funk1-Sim.mp3 – used in Experiment 2

Funk1-Strong.mp3 – used in Experiment 2

Folk1-Sim.mp3 – used in Experiment 2

Folk1-Strong.mp3 – used in Experiment 2