Getting in “the groove” while tapping

Petr Janata, Stefan Tomic & Jason Haberman

Center for Mind and Brain, Univ. of California, Davis, Davis, CA.

Introduction

People often find themselves moving along rhythmically with the music they are listening to, and this form of engagement with music seems to be quite pleasurable. Music that spontaneously engages our action system is commonly described as having a high degree of “groove.” Although many studies have examined how people find and synchronize with the perceived beat of the music, the degree to which spontaneous movement patterns occur on multiple metric levels is not well understood. Similarly, the relationships among the perceived rhythmic properties of the music, perceptions of one’s own movements, actual performance measures, and affective responses are poorly understood. In this study we examined the properties of bimanual tapping in silence and in response to music under both restrictive and unrestrictive tapping instructions in order to explore the relationships among these variables.

Subjects

Thirty four students with normal hearing and full mobility of their limbs were recruited from Psychology courses at UC Davis. No requirement for prior musical training were imposed. A brief survey indicated that 12 of the subjects had at 2 or more years of training on a musical instrument, and of those, 8 had 5 or more years of training.

Stimuli

Properties of the psychological construct of “groove” were assesed using a survey administered across six different experiments (Figure 1). Musical excerpts were the 30s free samples provided by the Apple Music Store. Forty eight excerpts were divided equally into 3 categories of perceived groove (low, medium, and high) based on groove ratings by 19 subjects in a pilot experiment. The excerpts were drawn from R&B/Funk, Pop/Rock, Jazz, and Folk genres and varied in tempo. The fourth stimulus category was silence.

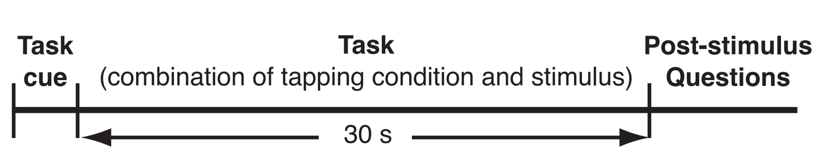
Tasks & Procedure

Subjects were seated in a sound-attenuating chamber in a comfortable straight-backed chair 2 m from a computer monitor and two speakers situated 1m to either side of the monitor. Bimanual tapping data were collected via the two large sensor pads closest to the subject on a Roland HPD-15 drum pad. The drum pad was mounted on a stand and placed between the subjects’ legs at a comfortable distance. All subject interactions with the computer were accomplished with a wireless mouse on an adjacent stand.

Subjects started with 5 practice trials to familiarize them with the different task demands, and then completed 66 trials from each combination of tapping condition and groove category (except no tapping in silence). Stimuli were sampled from each groove category at random without replacement. 6 isochronous and 15 freeform tapping trials were performed in silence.

Tapping Conditions: No Tap, Bimanual Isochronous, Bimanual Freeform

Stimulus Categories: Silence, Low Groove, Mid Groove, High Groove



Groove attribute ratings (N=166)

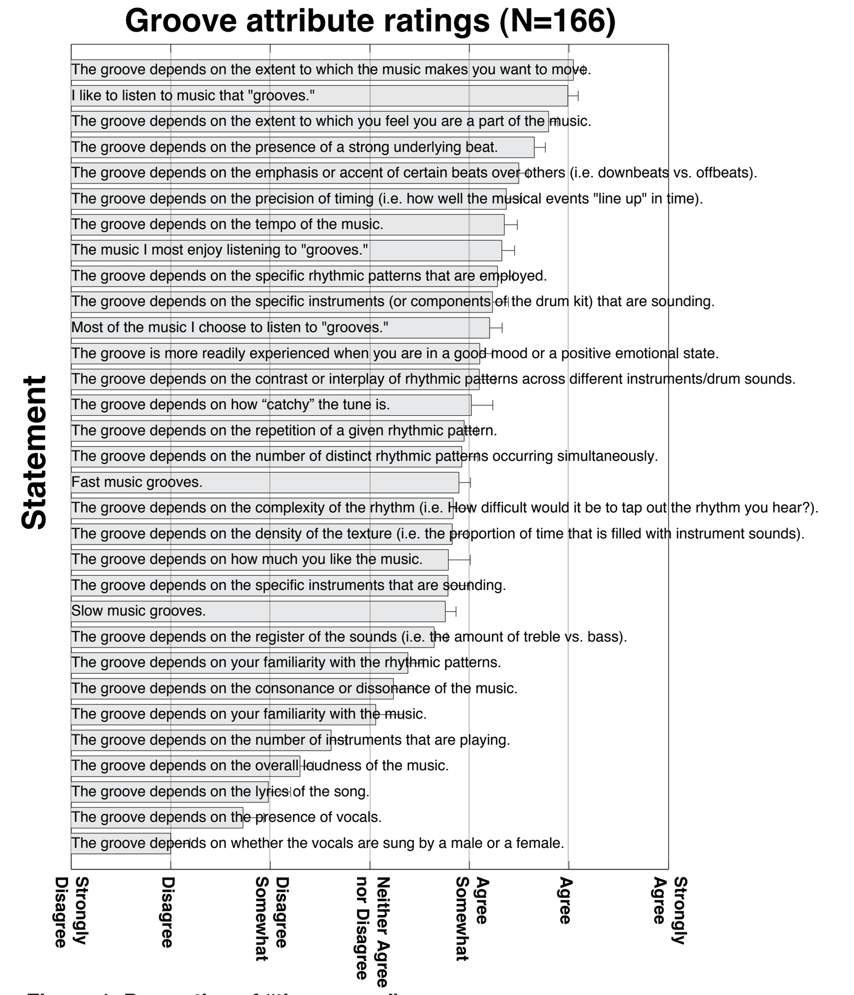


Figure 1. Properties of “the groove”
Using a 7-point scale, 166 subjects endorsed statements about properties of “the groove” and their disposition toward music that they perceive as having groove. The statements have been ordered by the average amount that subjects agreed with the statement.

Analyses

We analyzed the tapping data via two different methods. The first method used a traditional approach of calculating inter-tap-interval distributions, finding the peak in the distribution and calculating the variance of the distribution about the peak at the peak frequency. This approach worked well for the isochronous tapping conditions in which the hands effectively acted as one.

Development of the second method arose from the desire to have an efficient means of comparing rhythmic patterns in the tapping data with rhythmic patterns in music for which specific information about note onsets of all the instrument parts, e.g. MIDI information, is unavailable. To this end, we adapted a beat-finding model (Scheirer, 1998) so that we could visualize and quantify certain aspects of the time-varying rhythmic structure of both the musical excerpts and subjects’ performance.

Rhythm Profile Calculation

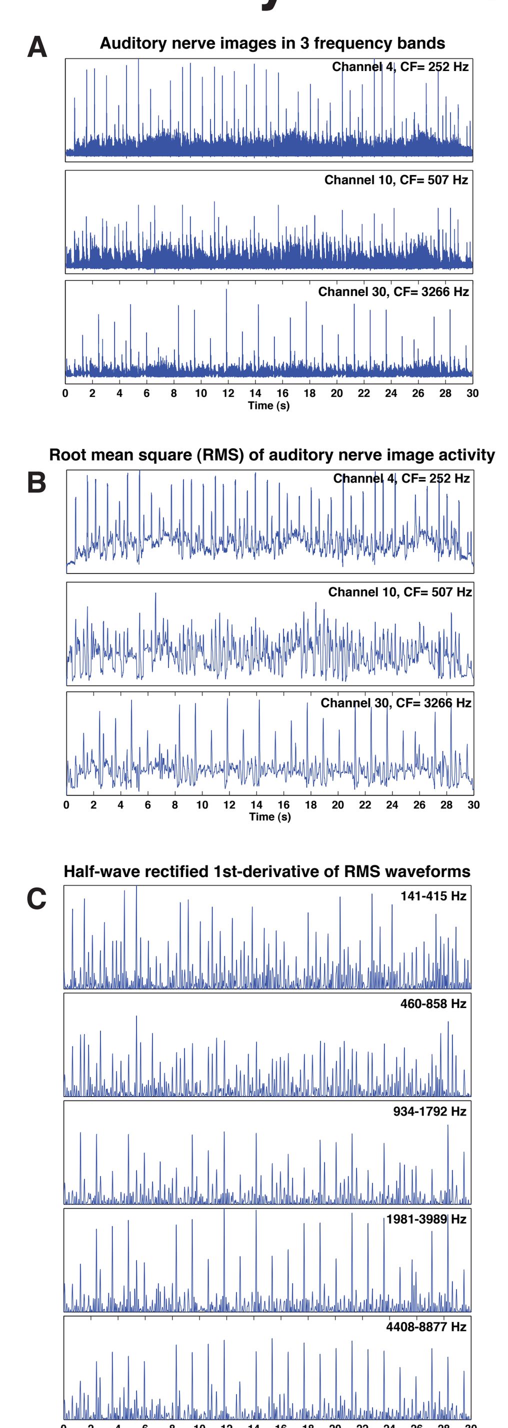


Figure 2. Stimulus rhythm profiles

A) An audio signal is processed through the auditory peripheral module of the IPEM Toolbox (Leman, Lesaffre, & Tanghe, 2001). The output of the module is an auditory nerve image (ANI), which is a physiologically motivated representation of the auditory information. The ANI consists of 40 frequency bands approximating critical bands spaced equally between center frequencies of 141 to 8877 Hz. Three of the channels are shown at the left.

B) The ANI is smoothed with an RMS (root-mean-square) calculation (50 ms sliding window), downsampled to a 100 Hz and filtered with a low pass filter. This calculation provides an estimate for the envelope of the ANI.

C) First order differencing and half-wave rectification are applied to the RMS output of step B. Half-wave rectification simply removes all negative values by setting them to zero. This output provides an estimation of the onset magnitudes and velocities of the original signal. Non-overlapping sets of 8 successive frequency bands are then summed to produce the 5 channels shown in this graph. Since the original channels are narrow, summation facilitates analysis of signal patterns that span several frequency bands. For example, vocal sound may have energy across a few channels, while percussive sounds will have energy across a much broader frequency range.

D) Each of the 5 summed channels is passed through a bank of 99 resonator filters. Each resonator is a band pass filter with an impulse response that behaves like a damped oscillator. The resonator center frequencies range between .25 Hz and 10.09 Hz, which are the rhythmic frequencies of interest. The onset patterns produced in the previous stage excite the resonators at varying degrees. Resonators with center frequencies closest to the most prominent frequencies of the onset pattern have stronger oscillatory patterns. Each of the filters has the same Q (quality) factor. The Q factor is the ratio of the filter's center frequency to its bandwidth. Q factor also indicates the number of periods at which the resonator decays by approximately 27 dB. Therefore, higher frequency filters decay faster in time. The resonators are spaced one bandwidth apart from each another. Note that the bandwidth decreases as center frequency increases, as all filters have the same Q factor. The filter bank output of the third summed channel (934-1792 Hz band in C) is shown in this graph.

E) The time-varying energy output of each resonator in each of the five filter banks is estimated using an RMS calculation in a 2s sliding window, and these estimates are summed together to give a composite rhythm profile which shows the relative prominence of the different periodicities in the input signal.

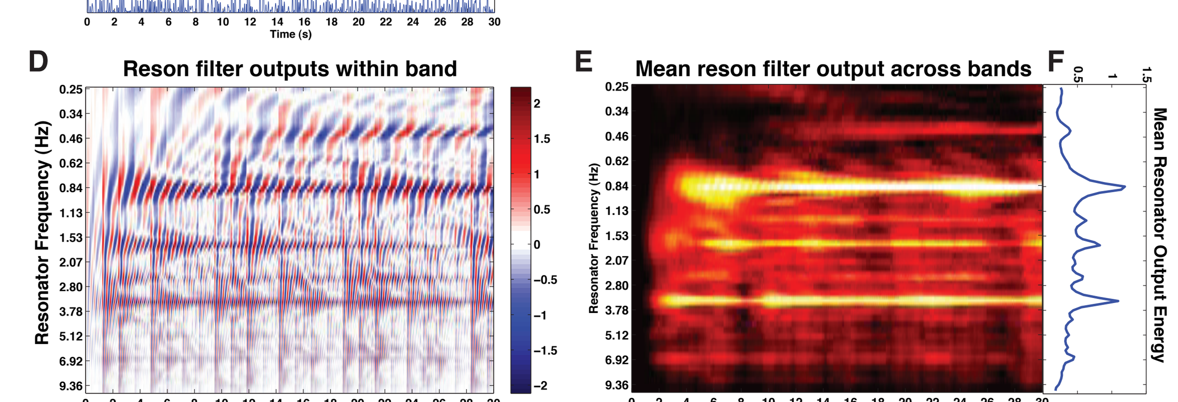
F) An average rhythm profile is obtained by averaging the filter energy outputs across the entire duration of the input signal.

Calculation of rhythm profile entropy

The complexity of the average rhythm profile is expressed in terms of entropy, **H**. The response profile is first normalized to unit area.

$$H = \left(-\sum p(r_i)\log(p(r_i))\right)/\log(N)$$

where r_i is the output of the ith reson filter and N is the total number of reson filters. If the energy is constrained to few filters, entropy will be lower than when there is a weaker periodic structure in the input signal.



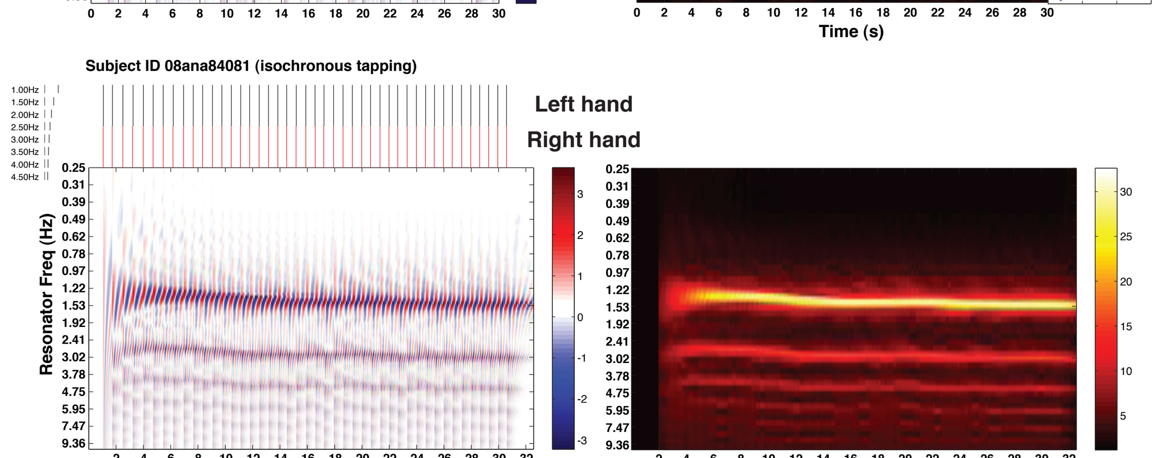
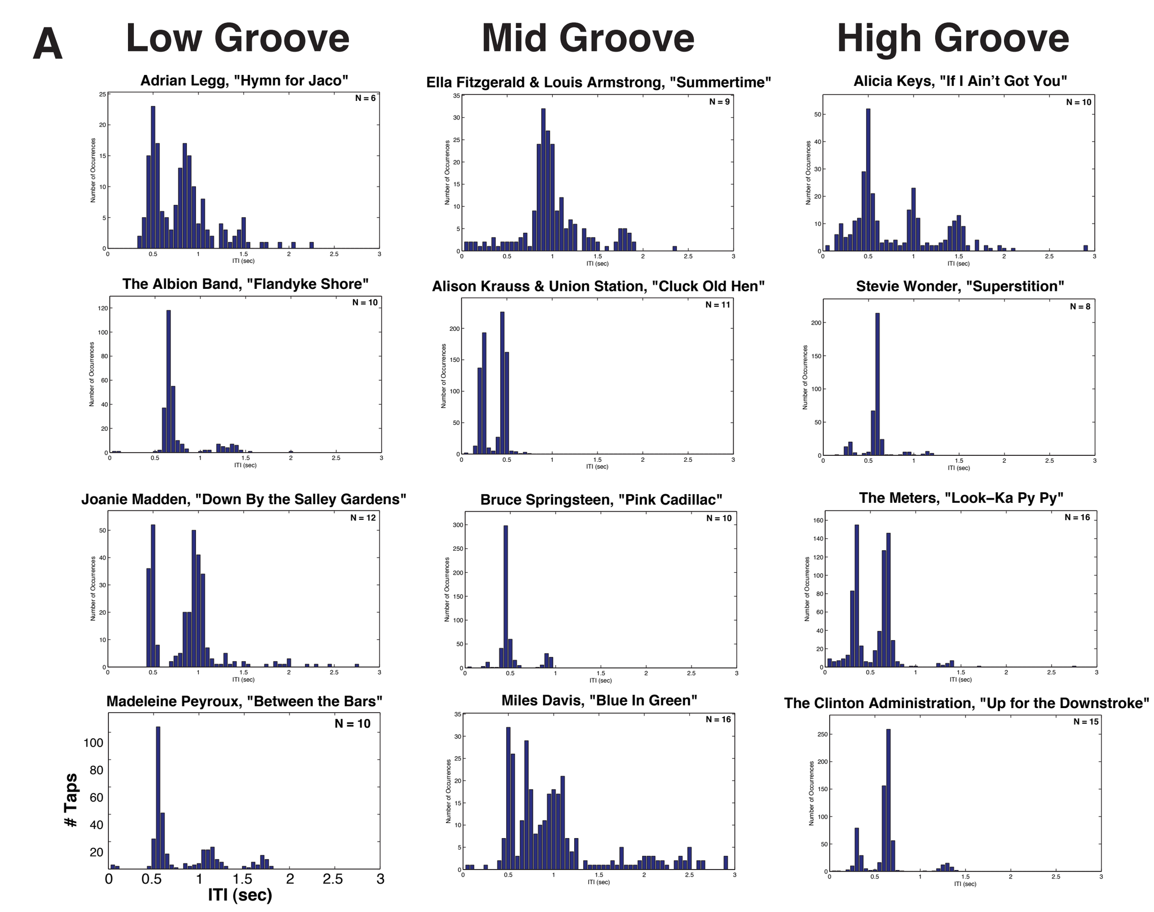


Figure 3. Calculation of response rhythm profiles
Rhythm profiles for subject’s tapping responses are calculated by modeling onsets of the MIDI events as Dirac impulses and feeding them directly into a single bank of reson filters (step shown in Figure 2D). For all of the tapping rhythm profiles shown in this poster, the left and right hand event streams are merged before they are fed through the filter bank. Thus, certain within-hand metric levels and relationships between metric levels may not be captured.

Isochronous tapping ITI distributions



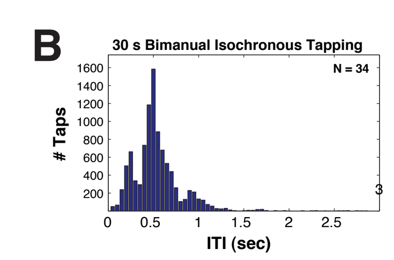


Figure 4. A) Cumulative ITI distributions obtained during isochronous tapping during the presentation of musical excerpts from different groove categories. Four examples were selected from each category. The examples were chosen to highlight the fact that some songs in a groove category prompted accurate tapping at a single metric level, whereas other songs prompted more variable tapping at individual metric levels. Some songs in each category prompted tapping at different metric levels. While it is not shown here, most subjects stayed at the same metric level throughout the 30s excerpt for most songs, so the peaks in these distributions reflect preferences of different subjects for different metric levels when responding to these stimuli. **B)** Distribution of inter-tap intervals across all trials of isochronous tapping in silence.

Rhythm profiles capture multiple metric levels

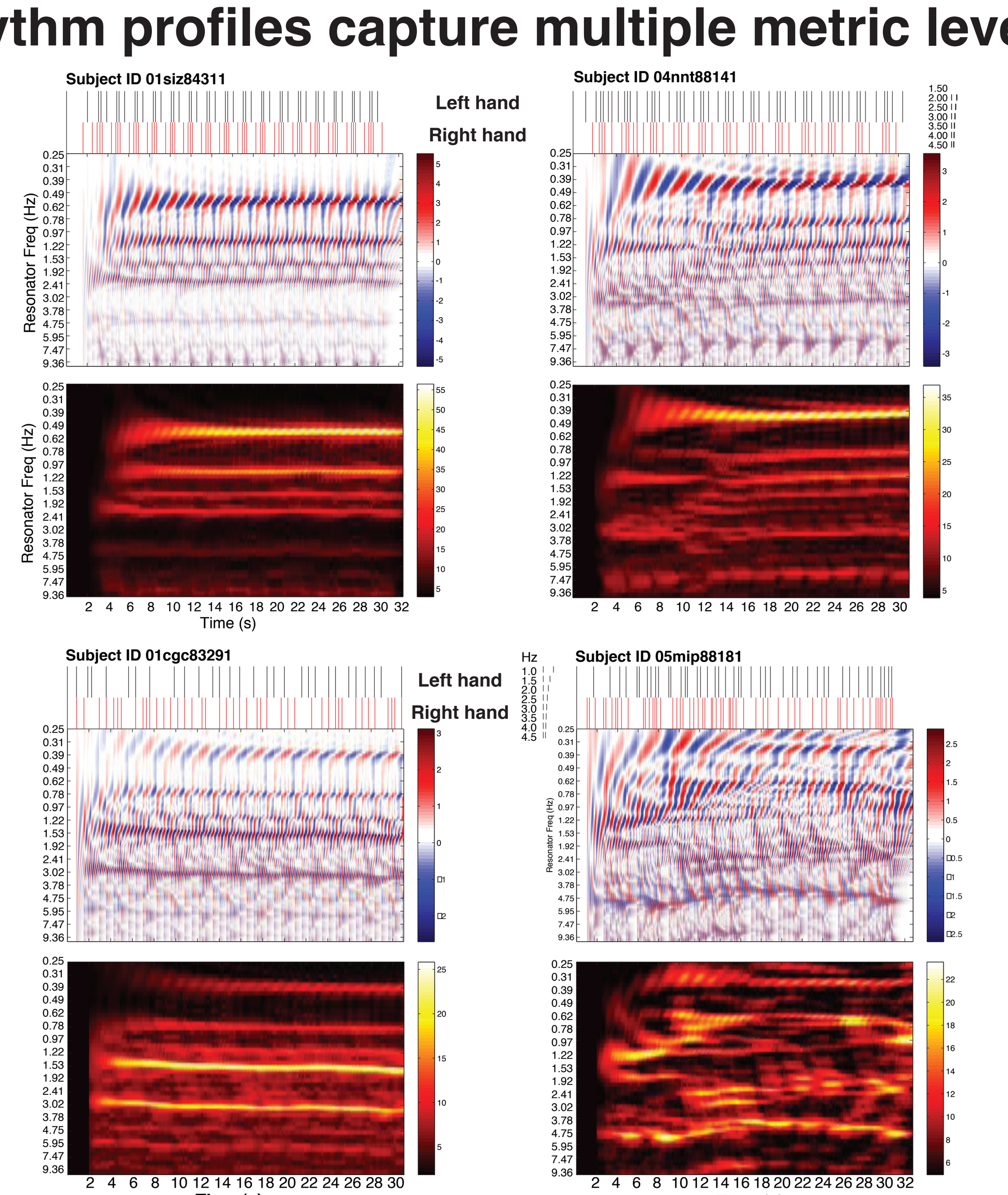


Figure 5. Rhythm profiles of freeform bimanual tapping in silence
Single trials for each of 4 subjects are shown in which they tapped as they saw fit for 30s in silence. Tap onsets for each hand are shown as raster plots that are aligned with the reson filter bank outputs in order to illustrate how the taps line up with the damped oscillations within each filter. These examples were chosen to illustrate the capacity of the reson filter model to capture the presence of rhythmic patterns in the tapping data, something that examination of 1st order intertap interval distributions cannot do. The bottom right panel shows the data from a trial in which the subject did not generate very good rhythmic patterns.

Correlations of Performance Measures and Subjective Ratings

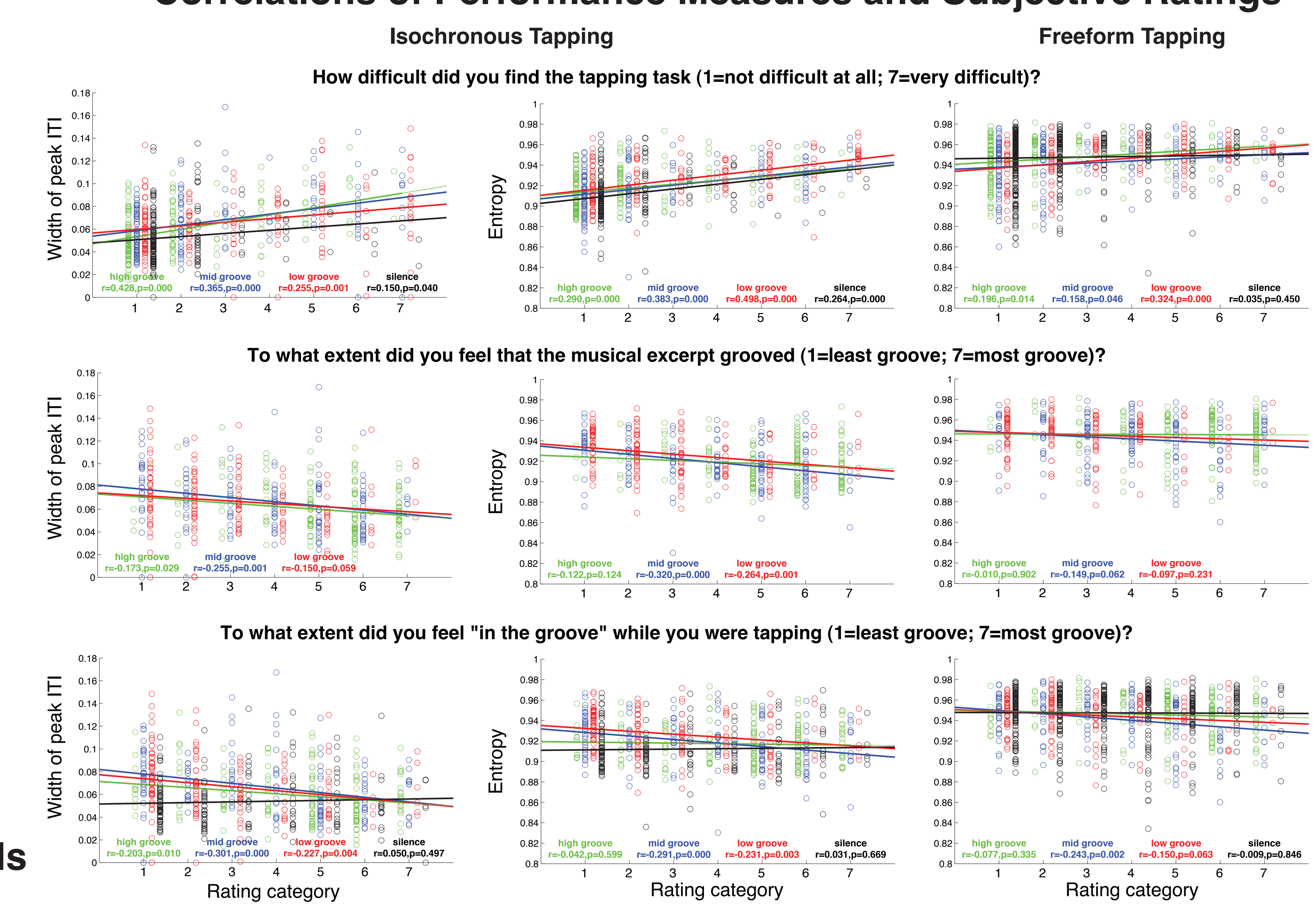


Figure 6. For two different performance metrics of tapping variability we calculated the correlation with subjective assessments made following each trial. For each trial, we calculated the ITI distribution, identified the peak, divided the period scale by the period of the peak ITI, and measured the width of the peak at half-height. The entropy measure was calculated as described in Figure 2F. Each circle in the plots above represents the data from a single trial. At each rating level, the circles are grouped by the stimulus category into which the stimuli were assigned prior to the experiment: red=low groove, blue=mid groove, green=high groove, black=silence. In the center panel, the distributions of circles from each groove category across the different rating levels indicate that groove ratings of the stimuli in this group of subjects largely matched those of the subjects in the pilot experiment. Performance measures were correlated with subjective measures most strongly in the isochronous (beat-finding) condition. Increases in perceived task difficulty were associated with increased variability. Tapping variability also decreased the more the person felt in the groove (except for the entropy measure during high-groove stims). The most notable exception was during tapping in silence. Isochronous tapping in silence did not tend to elicit feelings of groove. Free-form tapping elicited such feelings more often. As expected, freeform tapping was associated with greater entropy in the rhythm profiles than was isochronous tapping.

Mean stimulus and performance rhythm profiles

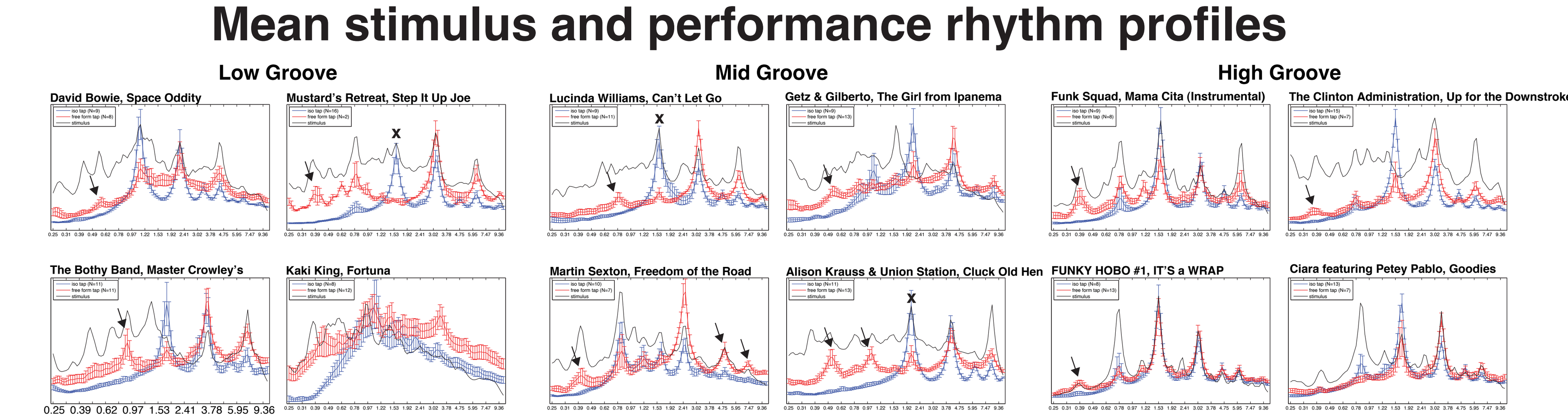


Figure 7. Mean rhythm profiles during isochronous tapping (blue) and free-form tapping (red) are superimposed on the rhythm profile of the stimulus (black) for 4 stimuli from each of the groove categories. The tapping rhythm profiles are the average of individual subject rhythm profiles and the error bars are 1 SEM. Arrows mark locations at which peaks emerged in the free-form tapping conditions. In most cases, the peaks in the tapping profiles reflected peaks that were also present in the stimulus profiles. The ‘Xs’ illustrate situations in which subjects commonly tapped at a particular rate when constrained to tap isochronously, but abandoned that rate during the free-form conditions. In a number of instances (Master Crowley’s -low groove; The Girl from Ipanema - mid groove; Up for the Downstroke - high groove), the major peak in the isochronous tapping profiles did not correspond to a peak in the stimulus profile.

Discussion

We examined the timing properties of bimanual tapping under both constrained and unconstrained tapping regimes in silence and while hearing music that varied in the degree of “groove.” The concept of groove, from the perspective of undergraduate students, refers primarily to the extent that a piece of music makes a person want to move. We sought to characterize the tapping performance with both a traditional measure of inter-tap interval distributions, as well as an adaptation of a beat-finding model that allowed us both to examine the metric structures in non-isochronous tapping data and to compare the tapping data directly with the stimulus in a common descriptive framework. Decreased variability in the isochronous tapping data corresponded with the perceived groove in the music and the degree to which a person felt in the groove while tapping. Isochronous tapping in silence was rarely associated with feelings of groove. Although free-form tapping in silence was more commonly asociated with feeling the groove, there was no correlation between this feeling and the tapping variability metrics we used. Thus, the feeling of being in the groove is largely potentiated by tapping along with music. When unconstrained to find the beat, subjects will often reproduce other rhythmic patterns that are present in the music. Analyzing bimanual tapping data within the same framework as analyses of the music appears to hold a lot of promise, though better metrics for comparing the music and tapping remain to be developed.

Acknowledgments: We thank Rawi Nanakul and David Horton for assistance with data collection. This work was supported in part by a Templeton Advanced Research Project grant from the Metanexus Institute.

References: Leman M., Lesaffre M., & Tanghe K. IPEM Toolbox (Version 1.01) [Computer Software]. Retrieved January 25, 2005, from http://www.ipem.ugent.be/Toolbox. Ghent, Belgium.

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