Tapping to Uneven Beats

Stephen Guerra, Julia Hosch, Peter Selinsky

Yale University, Cognition of Musical Rhythm, Virtual Lab

1. BACKGROUND AND AIMS [Hosch]

1.1 Introduction

One of the brain's most complex jobs is to interpret a myriad of stimuli and determine the most worthy of attention at any given time. Often when listening to music, one method of making sense of the multitude of musical stimuli present is to focus on a salient pulse and subdivide so that a listener gains expectancy for certain beats to come and becomes entrained to the rhythmic pulse. There have been multiple strategies suggested for processing this subdivision, such as entraining to the shortest pulse possible, focusing on a particular melodic line, or creating a beat from the shortest comfortable pulse (Martens, 2006). Also helpful when considering the "narrowing down" of over-abundant stimuli are strategies for discussing what types of stimuli might be the most prevalent to a listener, such as moving tones or complexity of polyrhythms (Handel, 1984).

This study intends to explore some of the questions surrounding the ability of musically trained auditors to entrain to non-isochronous rhythms. Specifically, do musically trained listeners demonstrate statistical tendencies in their metric entrainment of two different (number of attacks and size of inter-onset intervals) maximally even, non-isochronous rhythms posed against each other and a common unit (referent) pulse? If so, is there a tendency to hear one beat or the other as a NI meter, and what cues lead to a certain preference?

1.2 Previous Research

This study was highly related to Stephen Handel's 1984 study on polyrhythms, which discussed the point of using either non-isochronous rhythms or polyrhythms (or both) in research; namely that "dissonant rhythmic lines allow for the study of the emergent nature of rhythmic organization." Handel's discussion of a hierarchy to create musical rhythm as an overall impression included an interwoven "fast train" of pulses, or subdivision, and "slow train" of longer beats. Without the pulses, the longer beats did not make sense and were randomly occurring to participants, but without the slow beats, the fast beats (without spontaneous grouping) held no

meaning. Polyrhythms were merely the combination of more than one "fast train" of grouping, and usually led to cross-rhythm perception by a listener. However, factors that created this cross-rhythm perception included timing constraints such as tempo; the changing pitch of one or more of the polyrhythms; and configuration of the polyrhythms, which were useful to consider when determining factors relevant to a listener attempting to discern a melody line.

Regarding the subdivisions necessary for entraining to non-isochronous rhythms, Richard Parncutt's study on pulse salience and metrical accents demonstrated that without obvious subdivision, polyrhythms became significantly more difficult in terms of entrainment. In this study, participants were asked to tell whether a pitch given after a few cycles of a metrically complicated rhythm fell on or off the beat, and the participants were largely unsuccessful unless the pulse saliences that included the new event were very strong. In other words, if the stimuli provided the subdivision, even non-trained musicians could complete the task, but if not, everyone had difficulty (Parncutt, 1994).

1.3 Present Research

This study attempts to explore some of the questions surrounding musicians' ability to hear and subdivide non-isochronous rhythms. By creating a forced-choice, open response stimuli set for participants, we are able to take a broad look at whether factors such as pitch and timbre influence musically trained participants' choice of stimuli; whether participants entrain to one non-isochronous pulse over another regularly, showing the ease of a shorter or a longer subdivision or vice versa; and how the impact of years of musicianship affect the performance of the participants and their abilities to subdivide. We hope that with further research, studies like this one could impact further study of real-world ethnic rhythms, especially derivations of common Latin rhythms such as our stimuli.

2. METHOD [Guerra]

2.1 – Participants

59 participants took our online study; however, 8 did not complete all of the modules, and have therefore been excluded, leaving a sample size of 51. Our participants are aged between 18 and 59; their average age is 24, where 34 are between 18-20 yrs., 13 between 21 and 30 yrs., and 4 over 31 yrs. The gender distribution of our participants is close to balanced, with 28 male and 23 female. Our participants are largely musically trained: 33 self-report more than 10 yrs. of training

per their primary instrument. Only 2 (/51) reported 0 yrs. of training. Thus, we may generalize our analytical results to a population of male and female adult instrumental musicians.

2.2 Stimuli

Our experiment used five stimuli. All were created in Sibelius 7, exported as .aiff files (16-bit, 44.1 kHz), and converted to .mp3 files for online presentation. All stimuli are 0:58 m in duration. We set the notated quarter note across all stimuli equal to 90 bpm, with the notated 16th note thus equal to 166.$\overline{6}$ ms in absolute time. S1 through S4 use three Sibelius instruments: the flute [Fl.], the Bb clarinet [Cl.], and the wood block [WB], where the flute and clarinet sound the pitches F5 and C4, respectively. Because S5 tests the resultant pattern of NI5 and NI7, it is composed with only the flute and the wood block; in S5, the flute sounds a C5. The stimuli are identified as follows:

S1: Identity - NI5/NI5 (Fl./Cl. & WB)

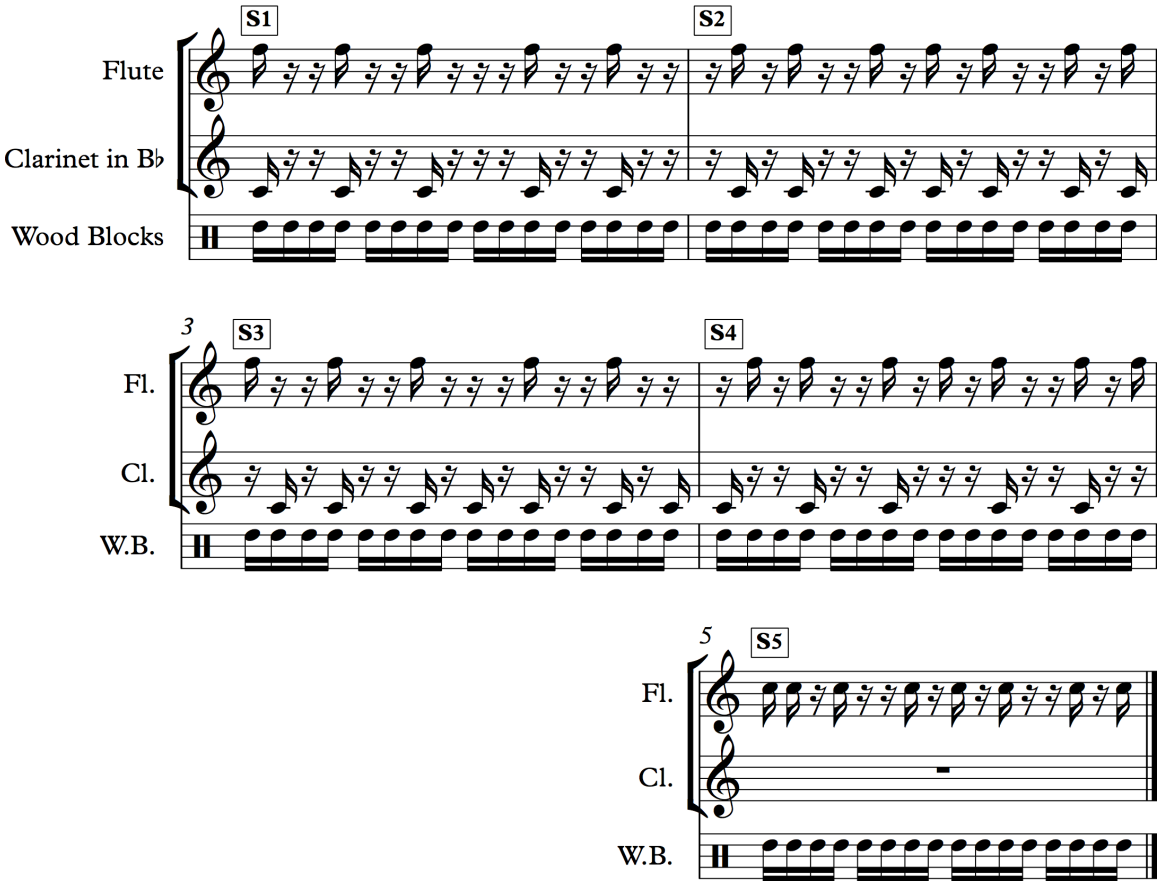
S2: Identity - NI7/NI7 (Fl./Cl. & WB)

S3: NI5/NI7 (Fl./Cl. & WB)

S4: NI7/NI5 (Fl./Cl. & WB)

S5: Resultant (Fl. & WB)

By way of summary, the figure below shows a single measure from each stimulus as composed in Sibelius where all instruments are present and sounding.



All stimuli used the same dynamic velocity template: each followed the same process of entrances, crescendos, plateaus, decrescendos, and exits. Crescendos and decrescendos were precisely configured using the 'velocity' settings in Sibelius via 'Transform Live Playback'. The figure below graphs the dynamic velocity template. N.b. S5 omits the clarinet; per S5, the template reads the same minus the clarinet row.

	Sibelius measure #:																				
Sibelius instrument:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Flute (F5)	-		0 to 85								85							85 to 0			-	
B♭ Clarinet (C4)	-		0 to 70								70							70 to 0			-	
Wood Blocks	0 to 60											60									60 to 0	

Sibelius values for dynamic playback velocities

2.3 Task & Procedure

Our study used both *tapping* and *forced-choice* paradigms. In general, the participant was asked to tap along to the 'prevalent beat' that he or she 'hears' and (in the case of trials using S3 and S4, where there are two different patterns) to choose which pitch he or she focused on: the 'higher'

or the 'lower'. The experiment was broken into six parts. The first two parts each include two trials where S1 and S2 are presented in a random order. For these trials, the participant was instructed:

> *Some beats can be uneven. Play the sound file below, where the pitches simulate the playing of an uneven beat. Try to tap along with the uneven beat.*

The third part presented an internal survey meant to assess familiarity with non-isochronous beats and to interrupt the study flow. The participant was forced to answer with either 'yes' or 'no'. The questions were as follows:

> *Do you play/study or have played/studied music with un-even beats?*
>
> *Do you play/study or have played/studied percussion or drumming?*
>
> *Do you play/study or have played/studied non-western music? (e.g., Indian, West African, East Asian, South American, American Indian, Jazz, Balkan, Middle-Eastern)*
>
> *Do you play/study or have played/studied any Afro-Latin music?*
>
> *Do you play/study or have played/studied any Brazilian music?*

The fourth and sixth parts each included two trials where S3 and S4 were presented in random order. The participant was instructed:

> *Remember that beats can be both **even** and **uneven**. When you press play in the box below, you will hear a woodblock and two different pitch sequences with two different rhythms. There will be a high pitched rhythm and a low pitched rhythm. After you hear these simultaneous rhythms, we will ask you a follow-up question. Please start tap-clicking to the **most prevalent beat that you hear** after the pitched sequences have begun. The sound file is about one minute long.*

The follow-up question forced the participant to choose whether he or she was focusing on the higher or the lower pitch. The fifth part interrupted the fourth and the sixth. The fifth part included one trial that presented S5. The participant was asked to tap-click to the most prevalent beat that he or she heard.

2.4 Data Collection & Analysis [Selinsky]

The data was collected through the NYU Music and Audio Research Laboratory [MARL] online platform and was organized in a Microsoft Excel spreadsheet. Participant taps were recorded as real time discrete series of timepoints in cumulative microseconds. Timepoints were divided by unit pulse (notated sixteenth note) duration, $166.\overline{6}$ ms. Differences between transformed, adjacent timepoints constituted IOIs (inter-onset-intervals) in unit pulses. IOIs were rounded to the nearest unit pulse so that all transformed IOIs were discrete quantities of pulses.

Participants were deemed to have 'successfully' performed a pattern, if their transformed IOI patterns matched one of three patterns (NI5, NI7, or composite) for a period of at least 32 consecutive unit pulses. When a participant's IOI pattern matched one of the same three patterns for at least two 16-pulse periods (but never for 32 consecutive pulses), that participant was deemed to have 'unsuccessfully' attempted that pattern. When a participant tapped a consistent IOI of 1, 2, 3, or 4 unit pulses for a period of at least 16 unit pulses, that participant was deemed to have successfully performed a isochronous IOI pattern. These patterns are denoted 'sixteenth' (1-1-1-etc.), 'eighth' (2-2-2-etc.), 'dotted-eighth' (3-3-3-etc.), and 'quarter' (4-4-4-etc.). If the participant tapped none of the above, a participant was deemed to have tapped no pattern. If a participant successfully performed multiple patterns to a single stimulus, the participant was deemed to have performed the final pattern. E.g. several participants tapped along with the woodblock only to change their patterns once other voices entered.

Analyses: (1) A distribution of performed rhythms (in percentages) was calculated for each stimulus. Distributions were aggregated and compared. Performances of NI5 and NI7 were compared. (2) For S3 and S4, participants were segregated by the pitch on which they reported focusing and distributions of performed rhythms were created. (3) For S1, S2, S3, and S4, participants were segregated by reported musical ability and distributions of performed rhythms were created.

3. RESULTS

3.1 Analysis & Figures 1 and 2 [Selinsky]

For each stimulus, participants were most likely to either perform or attempt to perform the pattern of the higher pitch (i.e. the flute): for S1 to S5 respectively, 78%, 75%, 40%, 72%, and

88% of participants performed the pattern of the higher pitch. Of the five stimuli, S3 had the widest spread distribution of performed rhythms; in S3, NI5 (successful) was the most commonly tapped pattern, but by only 31.3% of participants.

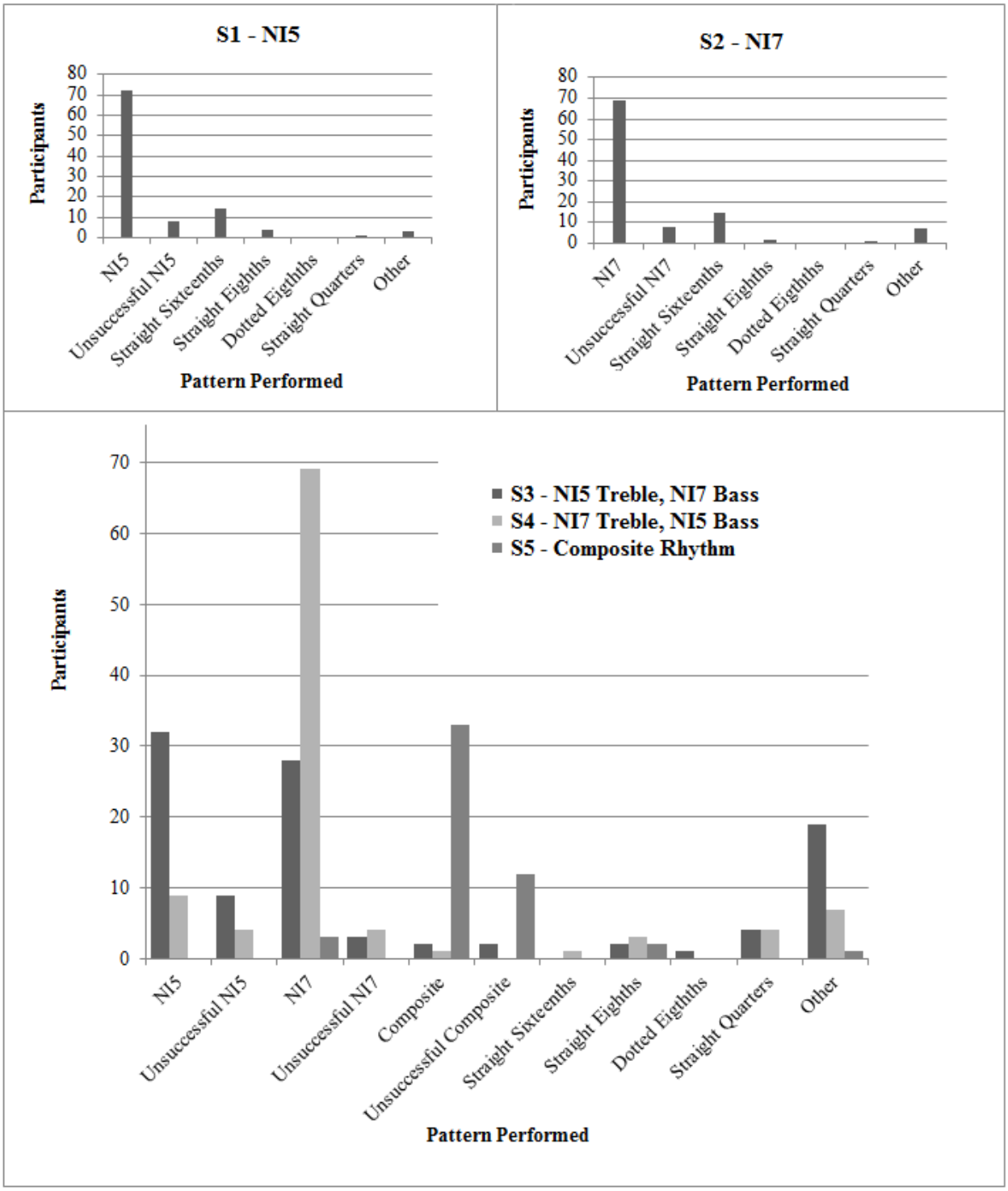


Figure 1: Distribution of Patterns performed for each stimulus

Between S3 and S4, in which NI5 and NI7 were presented simultaneously in distinct registers, participants were 92.6% more likely to tap NI7 than NI5. The proportion of unsuccessful performances for NI5 was .317 versus only .118 for NI7.

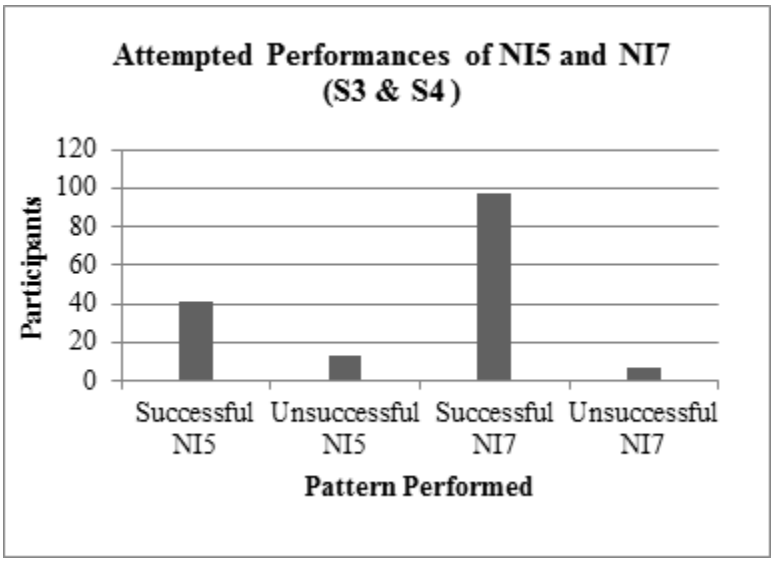


Figure 2: Participants' abilities to perform NI5 and NI7 when NI5 and NI7 are presented simultaneously

3.2 Analysis & Figures 3 and 4 [figures Guerra, prose Selinsky]

While listening to S3, 59.8% of participants reported focusing on NI5 (the higher pitch). Of those that reported listening to NI5, 52.4% successfully performed NI5.

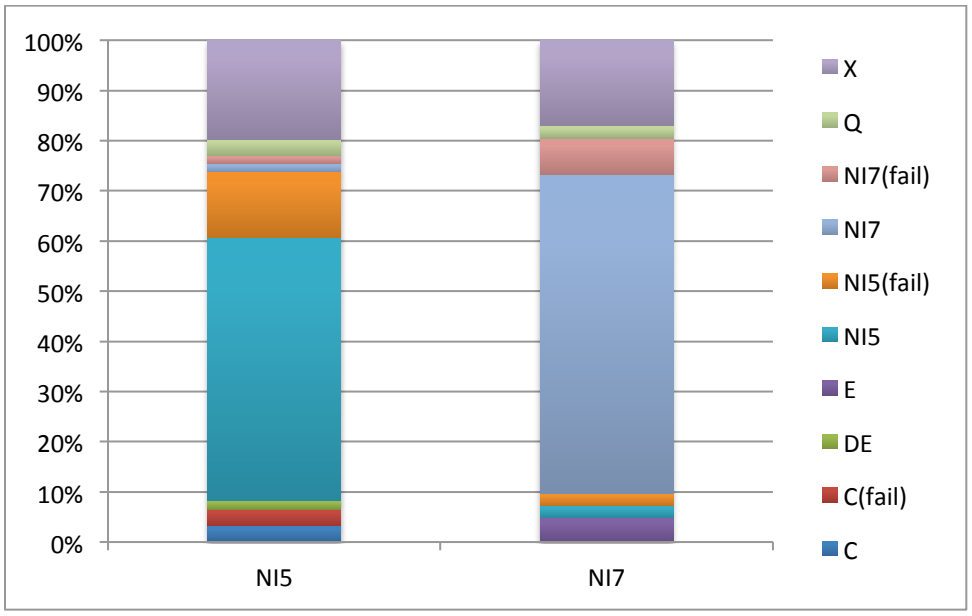


Figure 3: S3 and Reported Focus

When listening to S4, 82.3% of participants reported listening to NI7. Of those that reported listening to NI7, 79.7% successfully performed NI7. Between S3 and S4, 38.7% of participants reported listening to NI5, of whom 52% successfully performed NI5. Of the 61.3% of participants who reported listening to NI7 74.4%, performed NI7.

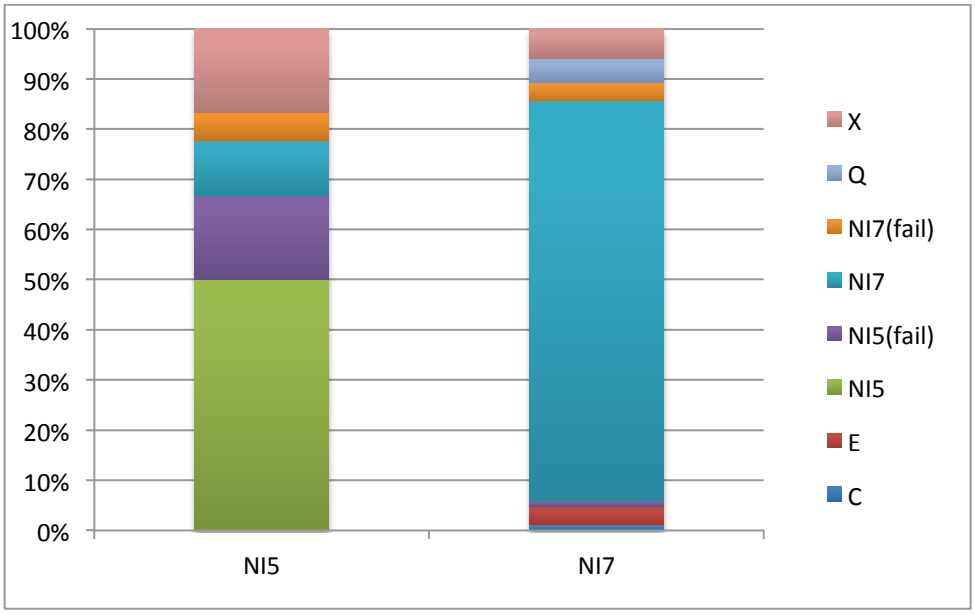


Figure 4: S4 and Reported Focus

3.3 Analysis & Figure 5 [Hosch]

In the initial survey of basic personal data to find demographics for the study participants, participants were asked to rate their level of musicianship from 0-5. Level 0 represented no experience, while level 5 represented professional musicianship. The intent of this analysis was to show that proficiency of subdivision demonstrated by professional musicians and overall time in spent in musical training would impact the overall success rates of musicians in the various tasks. However, overall results showed this trend less than would be expected. In all stimuli, the overwhelming majority of participants were able to correctly predict beats, regardless of level.

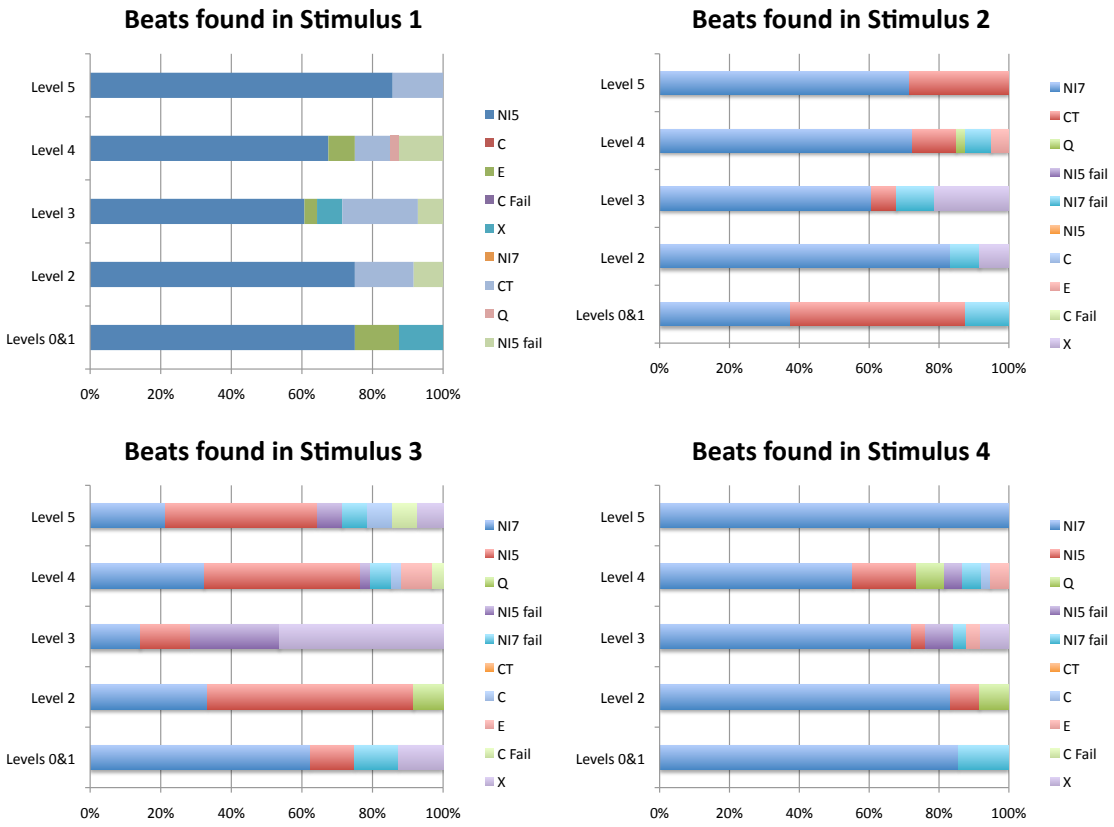


Figure 5: Beats found in Stimuli 1, 2, 3, and 4 by groups of participants based on self-reported expertise level

For example, in stimulus 1, Level 5 participants were most accurate at the task, with 84% of musicians correctly identifying the NI5 stimulus. However, the other levels also met with varying success, between 60% (Level 3) and 76-77% (Levels 0, 1, and 2).

A similar pattern appeared in Stimulus 2, with 70% of Level 5 musicians correctly finding the NI7 beat (and the rest clicking the 16th note subdivision), while Level 2, 3, and 4 musicians were close, ranging between 60% (Level 3) and 82% (Level 2). Levels 0 and 1 performed under 40% successfully, both erring in clicking through the 16th note, but also making NI7 clicking errors.

Stimuli 3 showed significantly fewer trends than would otherwise be expected. There was no significant difference between groups based on expertise. Stimulus 4, however, returned to the pattern, with 100% of Level 5 listeners correctly perceiving NI7, and the other levels ranged from 58% (Level 4), and 84% (Levels 0 & 1).

From this data, it can be determined that the ability to successfully subdivide may not have been a major factor in determining successful beat-keeping in the experiment, or that after a certain level of musicianship (perhaps a few years) people become generally successful and extra practice is unnecessary. A follow-up experiment should use more non-musicians as controls.

4. CONCLUSION

Our results suggested that listeners, when asked 'to tap to the most prevalent beat that [they] hear,' often choose to tap uneven beats if they are readily apparent on the musical surface. One surprising note was that although Handel's 1984 study showed that pitches of the "pulse train elements" were not a primary determinant of rhythm unless they were moving throughout the stimulus, our research showed different results (Handel, 1984, page 477). We found that a majority of listeners would tap to the highest pitch, because of the registrally-stratified patterns of our stimuli. However, even though participants generally focused on this one factor of pitch, they did not tap an even distribution of patterns between stimuli, which suggests that other confounding factors were involved in their choices of (uneven) beats.

Although listeners performed similarly when tapping to isolated NI5 and NI7 patterns, when NI5 and NI7 were presented simultaneously, participants seemed to perform NI7 more accurately than NI5. Further supporting this claim, trials of S3, in which NI5 was registrally highest, elicited the most frequent occurrence of 'other' patterns, some of which may have been attempted NI5 performances that were not accurate enough to fulfill our criteria. Because participants' performances of NI5 differed drastically between stimuli, we would like to advance the possibility that it is harder to conceptually reproduce in composite settings than NI7.

Alternately, because of participants' high proportion of 'unsuccessful' composite performances to 'other' performances for S5, it seems likely that the composite rhythm was

physically harder to reproduce than either NI5 or NI7. Given the strong effects of register and pattern on the determination and performance of uneven beats suggested by these results, further study attempting to isolate these features should be attempted.

There are three additional limitations to our method that should be addressed here and would benefit from further study. First, while our intent was to avoid having participants tap isochronous patterns, by informing our participants of the presence of uneven beats, we may have unfairly suggested that our participants tap in an unnatural manner. Second, our experiment was performed at a single tempo; if there are biases towards or against certain durations in beat selection, re-running the experiment at a variety of tempi should help to mitigate these factors. Finally, we suspect that the presence of an underlying isochronous pulse, sounded by the woodblock, may have confounded our results by providing our participants a means of organizing the NI5, NI7, and composite patterns.

REFERENCES

Handel, S. (1984). Using polyrhythms to study rhythm. *Music Perception: An Interdisciplinary Journal*,*1*(4), 465-484. Retrieved from http://www.jstor.org/stable/40285273?seq=19

Kivte, T.E. (2007). Categories and Timing: on the Perception of Meter. *University of Illinois Press.*

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

Parncutt, I. (1994). A Perceptual Model of Pulse Salience and Metrical Accent in Musical Rhythms. *Music Perception, 11(4)*, 409-464.