How Well Can You Persist A Tempo?

A Study of Tempo Persistence and Musical Background

Jingjie Zhang
CCRMA, Stanford University
jingjiez@ccrma.stanford.edu

Abstract

The ability of stable and accurate tempo persistence is proven by prior studies to be significant during music ensembles where musicians need to synchronize with each other. This study tries to relate musicians' capability of persisting different tempos to their daily music exposure level, their experience in bands or musical groups and how often they use metronomes. The results so far suggest positive correlations between the tempo persistence ability and the daily music exposure level as well as the experience in musical groups, however, participants who usually use metronomes demonstrate a relatively worse performance.

Introduction

Previous studies have demonstrated the significance of tempo persistence during music ensembles where musicians need to synchronize with each other (Jungers, *et al.*, 2002). More specifically, in bands or other musical groups, there is usually at least one member, such as a conductor or a drummer, who controls the tempo of the whole ensemble, while all the other members should keep synchronized with them. For both types of roles in musical groups, the failure to persist the mutual tempo during an ensemble will result in a certain level of rhythmic dissonance, which might affect the performance of other members and is sometimes noticeable to the audience, especially in certain genres of music that are sensitive to tempo consistency.

This study focuses on the relationship between certain aspects of musicians' backgrounds and their capability of persisting different tempos. The hypothesis is that musicians with higher daily music exposure level and longer experience in bands or musical groups might be able to achieve better performances when they are asked to persist certain tempos. In addition, it is also widely believed that practicing with a metronome could improve such an ability. An experiment will be designed to test a subject's performance of persisting a few tempos range from slow to fast, and the results of subjects from different musical backgrounds will be analyzed and compared.

Experiment design

The subjects will be asked to listen to a few pulse trains with different tempos that range from slow to fast and try to track the tempos through tapping. Each pulse train will last several seconds. After the pulse trains stop, the subjects need to keep tapping and persist the tempo for a certain amount of time.

The recording will start while the pulse trains are still playing to make sure that the first few taps are on the beat. During the experiment, the subjects are allowed to create any context out of each tempo, such as imagining melodies and rhythms that fit in that tempo, because this study is more focused on subjects' ability of tempo persistence under the scenario of music ensembles. After the tests, the subjects need to answer a few survey questions regarding their musical background and their experience with musical groups and metronomes.

In this experiment, the order of the tempos are randomly chosen to avoid the situation where the tests are arranged simply from slow to fast or the other way around and thus reduce the influence between tests. To get the subjects ready, the very first test will serve as a training test without informing them in advance. The result of this training test will still be recorded for further study but will not be counted in the subjects' individual performances. Instead, the same pulse train will be played again as the last test.

Methods

The hardware equipment used for tapping is an AKAI® MPK mini MIDI controller, which has eight backlit pads that are mapped to MIDI note 44-51. Specifically, the pad corresponding to MIDI note 47 is used in this experiment. The MIDI tapping data acquisition and the user interface of the experiment is implemented in ChucK. Despite the fact that recording keyboard tapping is also possible through the HID support in ChucK, the MIDI pads are preferred due to their shorter key travel time so that the time measurement in ChucK can be more accurate.

The major task of each measurement is to detect the MIDI message "144-47-XXX", where 144 is the status byte of this MIDI message, while 47 and XXX are the data bytes of this message. Converting 144 into a hexadecimal number results in 90, which represents a MIDI note on message on channel 0. The data bytes "47-XXX" denote MIDI note 47 and its velocity information respectively. The ChucK program only monitors the MIDI note on messages with any velocity, while the MIDI note off messages are ignored, since the velocity and aftertouch information is relatively insignificant in this study. After recognizing a correct MIDI message, the corresponding timestamp is recorded and written into a .csv file by the ChucK program.

A simple state machine implemented in ChucK controls the whole experiment process, including five tapping tests under four different BPMs with random order and four survey questions in the end. For a BPM of 60, 90, 120 and 150, the corresponding test duration will be 30s, 40s, 50s and 70s respectively so that each test can have around 75 data points. Although the order of the four tempos are randomly chosen by ChucK, the first pulse train is generated again in the fifth test, since it serves as a training test as stated in the previous section. Before starting the recording of each test, the ChucK program clears the current MIDI message buffer to make sure that no previous data can burst out in the beginning of each recording. A higher-pitched pulse is used to indicated the start and end of each test.

The user interface of this experiment including instructions and survey questions is implemented using the console of miniAudicle. The four survey questions are:

A. Which of the following best describes your musical training background?
   (1) Never (2) <1 year (3) 1-3 years (4) >3 years

B. How much time do you spend practicing (or listening to music) everyday?
   (1) <0.5 hour (2) 0.5-1 hour (3) 1-3 hours (4) >3 hours

C. Which of the following best describes your experience in bands (or musical groups)?
   (1) Never (2) <6 months
   (3) >6 months (You are the one who follows the tempo)
   (4) >6 months (You are the one who controls the tempo)

D. When you practice alone, how often do you use a metronome?
   (1) Never (2) Seldom (3) Often (4) Always

The first two questions cover the basic musical background of the subjects. The purpose of question A is to divide the subjects into groups with different musical training experience, while

question B investigates the subjects' daily exposure to music. On the other hand, question C and D focus on more specific experience that is related to tempo persistence and music ensembles. Subjects with little experience in musical groups are distinguished from experienced band members by question C. Moreover, these musicians' roles in musical groups are also of great significance, so question C further differentiates between the musicians who control the tempo of the whole band and the others. In addition, since it is widely believed that practicing with a metronome may improve musicians' ability of accurate and stable tempo persistence, the subjects' experience with a metronome is covered in question D.
Finally, the .csv files containing subjects' tapping data and answers to the survey questions will be loaded and analyzed in MATLAB. In this experiment, the data analysis will mainly focus on the time differences between adjacent data points and their deviation from the period of the reference tempo. The subjects will be divided into different groups according to different survey questions so that the relationship between tempo persistence capability and different musical backgrounds can be studied.

Stimulus

The stimulus of this experiment are pulse trains under four different BPMs, which are chosen to be 60, 90, 120 and 150, since they are typical tempos used in music pieces. Each pulse is generated in ChucK by applying an ADSR envelope to the output of a sine wave oscillator, and its waveform is shown in Figure 1. The attack, decay and release time are 0.01ms, 8ms and 30ms respectively, and the sustain level is around -3dB. Each pulse train consists of 32 pulses, and the subjects will be given 16 pulses to get ready and another 16 pulses to keep themselves on the beat at the beginning of each recording. The 17th pulse of each pulse train has a higher frequency than the normal pulses, which will be played again at the end of each test. The frequency of the normal pulses are set to 1.1kHz, while the higher-pitched pulses are one octave up.

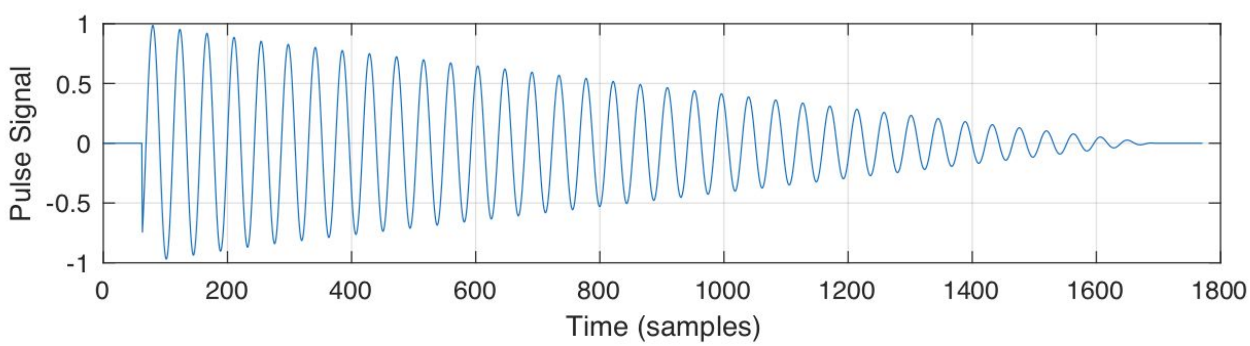
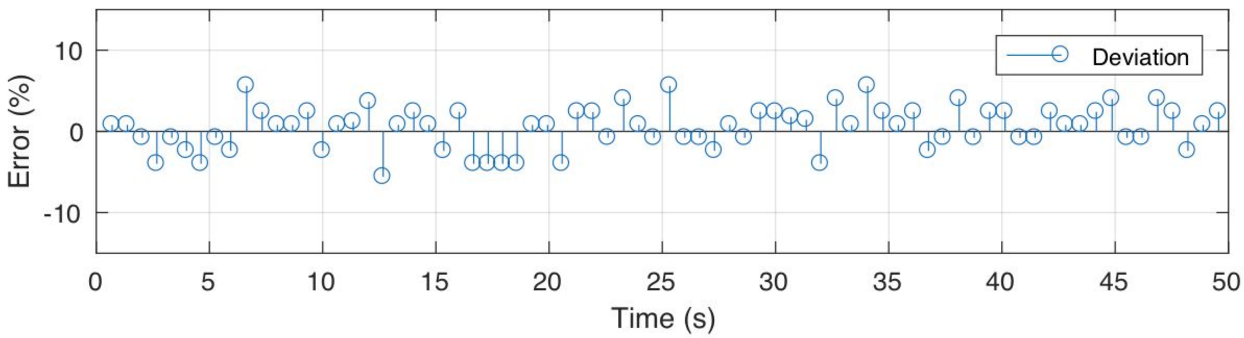


Figure 1: Waveform of the generated pulse signal.

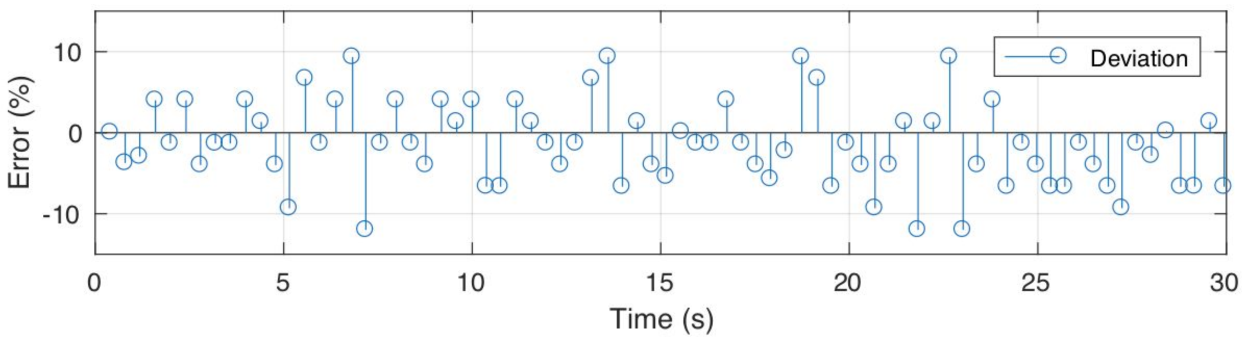
Data Analysis

The tapping data is basically a set of data points on the time axis, while the time differences $\Delta t[n]$ between adjacent data points are the subjects' instantaneous periods at that moment. It is more intuitive to quantify the deviation of instantaneous periods from the period T of the reference tempo by constructing an error signal $e[n]$, which is given by:

$$e[n] = \frac{\Delta t[n] - T}{T} \cdot 100\%. \tag{1}$$



(a) An error signal from a test under BPM 90.



(b) An error signal from a test under BPM 150.

Figure 2: Two error signals taken from different tests of two subjects.

However, to measure and compare different subjects' performances, the error signal itself is not enough. For example, Figure 2 shows two error signals taken from different tests of two subjects. It is clear that we cannot directly compare these two vectors. In this situation, a commonly used method is to calculate the mean squared error (MSE) of the vector (Wackerly *et al.*, 2007):

$$\text{MSE} = \frac{1}{N}\sum_{n=1}^{N} e^2[n]. \tag{2}$$

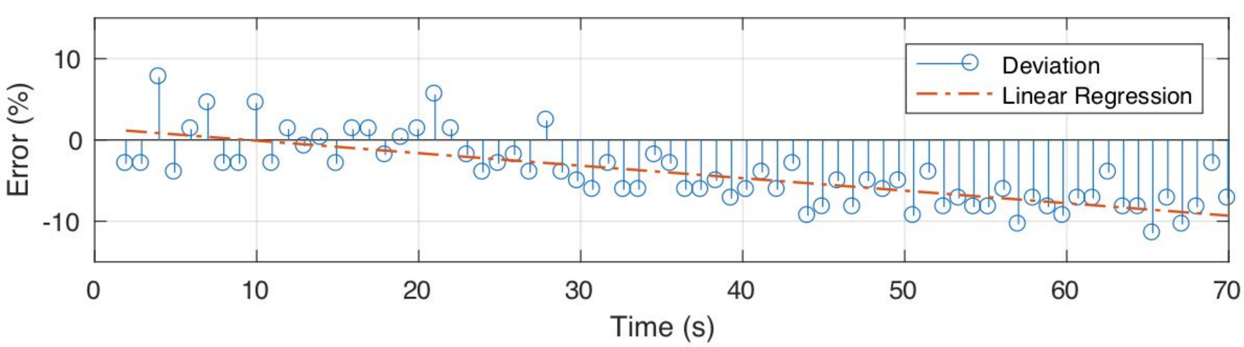


Figure 3: An error signal caused by drifting tempo and the corresponding linear regression line.

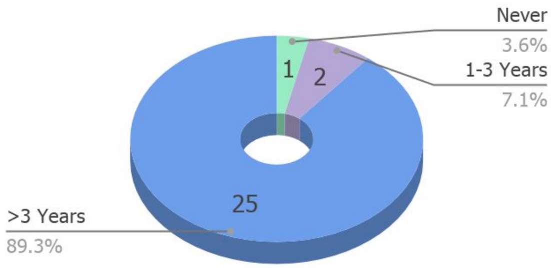
The variation tendency of each subject's instantaneous tapping tempo is also worth studying, and the most efficient approach is to apply a linear regression to the error signal, as demonstrated in Figure 3. By using linear regression, the overall error variation throughout one tapping test is approximated to be linear (Freedman, 2009; Yan and Su, 2009). The linear model of the error signal is given by:

$$e[n] = k \cdot n + b + \varepsilon_n, \tag{3}$$

where k and b denote the slope and y-intercept of the linear regression line respectively, while ε_n denotes the equivalent random noise at sample n.

Participants

Currently, there are 28 participants in this study, and 25 of them have more than 3 years of musical training, as presented in Figure 4(a). On the other hand, Figure 4(b) shows the subjects' daily music exposure levels, which range from less than 0.5 hour to more than 3 hours, and it is considered evenly distributed according to the categories defined by the second survey question.



(a)

(b)

Figure 4: (a) Musical training experience; (b) Daily music exposure level.

In addition, 22 out of the 28 subjects have more than 6 months' experience in bands or musical groups, and 6 of them are drummers, as illustrated in Figure 5. Moreover, Figure 6 shows that 19 subjects have little experience with metronomes, while 9 other subjects use them frequently.

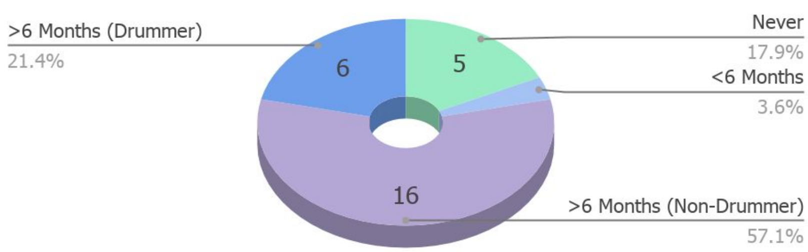


Figure 5: Experience in bands or musical groups.

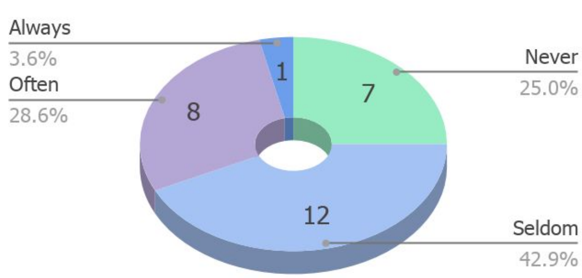


Figure 6: Frequency of using metronomes.

Results & Discussions

As mentioned above, to measure and compare the performances of different groups of subjects, the two major performance indicators used in this study are the mean squared error (MSE) and the linear regression slope k.

Figure 7 and 8 present the average MSE and linear regression slopes of subjects with different daily music exposure levels. As expected, the overall tendency of the data is that subjects with higher daily music exposure level tend to produce smaller error when persisting a tempo, and their instantaneous tempo is more stable so that the average linear regression slope is closer to zero. It is also worth mentioning that although the average MSE of the "0.5-1 Hour" group and the "1-3 Hour" group are close to each other, the "1-3 Hour" group demonstrates a better performance in the comparison of average linear regression slopes. Thus, according to the current results, there seems to be a positive correlation between the tempo persistence ability and the daily music exposure level, and it appears to have a more notable influence on the issue of drifting instantaneous tempo.

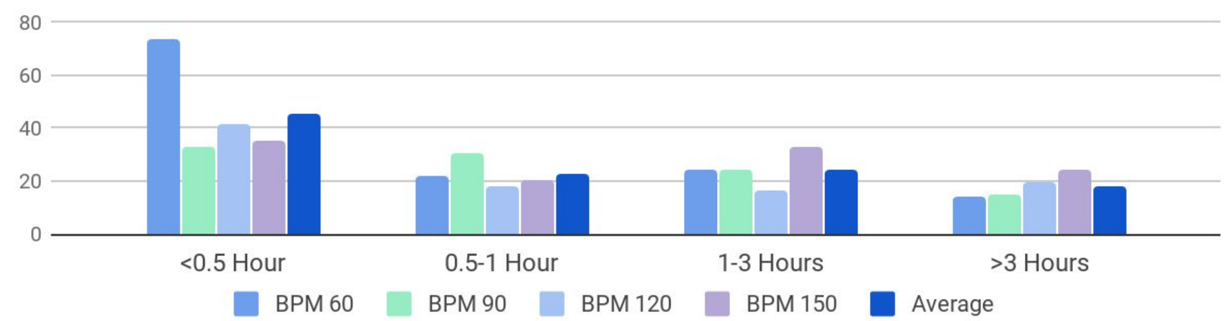


Figure 7: Average MSE under different daily music exposure level.

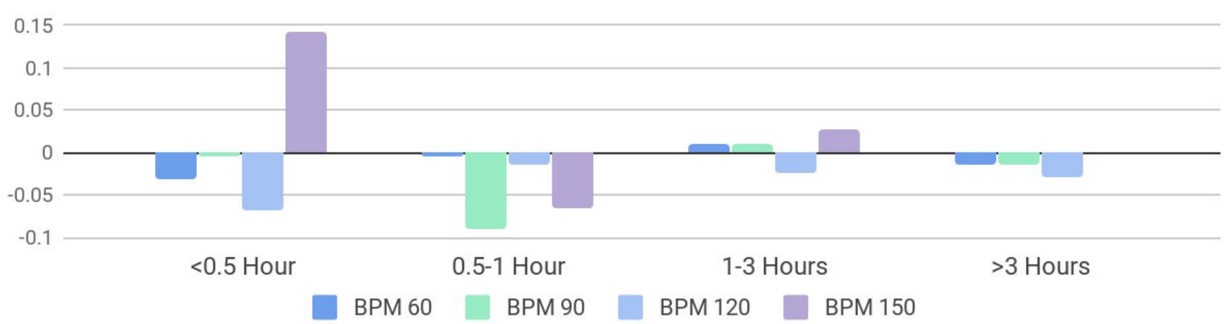
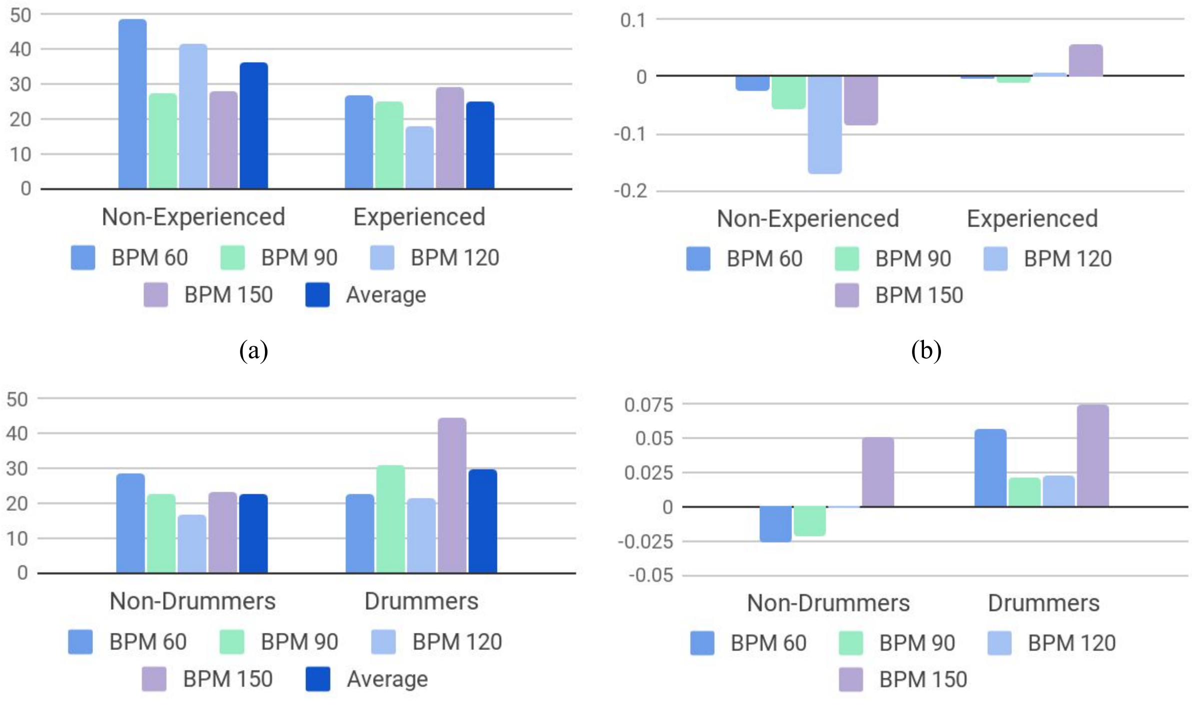


Figure 8: Average linear regression slope under different daily music exposure level.

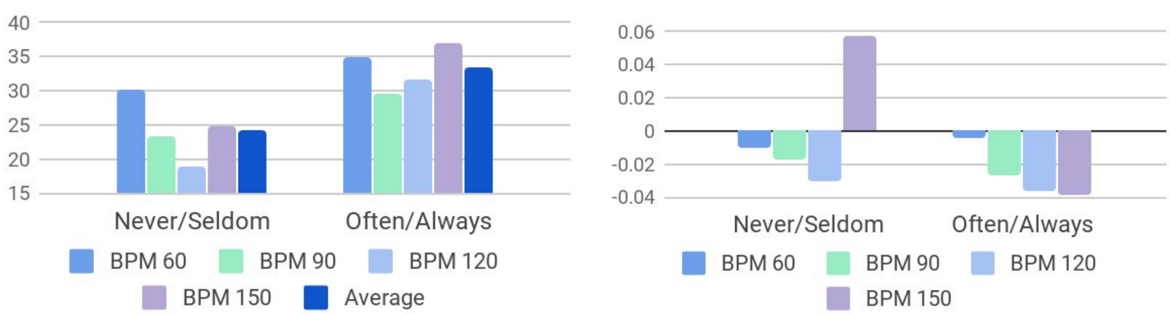
In regard to the influence of subjects' experience in musical groups on their performance, experienced band members demonstrate better performance than non-experienced subjects, which also fits the initial expectations. As shown in Figure 9(a) and 9(b), the average linear regression slopes of the experienced band members are especially remarkable, compared to that of non-experienced subjects. Therefore, with the group of experienced band members, the performances of drummers and non-drummers is further analyzed and compared. However, as presented in Figure 9(c) and 9(d), the average performance of drummers is actually worse than non-drummers. Such results might come from the small subject size, since there is only 6 drummers in this study. In addition, according to the individual performances of the 6 drummers, one of them actually produced a large number of errors, which strongly affected the average result of the drummers.



(c) (d)

Figure 9: (a) Average MSE and (b) average linear regression slope under different experience in musical groups; (c) average MSE and (b) average linear regression slope of different roles in musical groups.

Another unexpected result is that subjects who use metronomes more frequently tend to produce bigger errors in this study, and they always tend to be faster than the reference tempo, as shown in Figure 10. It is still unclear whether this result will be different under a larger subject size, but one of the possible explanations mentioned in the classroom discussions is that people who always use metronomes during practicing might develop a dependence to metronomes and therefore cannot reproduce the tempo accurately after the pulse trains stop.



(a) (b)

Figure 10: (a) Average MSE and (b) average linear regression slope under different metronome experience.

It is also interesting to study the average linear regression slope of all subjects so far, as presented in Figure 11. It seems that subjects tend to become slower under the BPM of 150 and faster under the other three tempos. However, the average linear regression slope under the BPM of 60 appears to be smaller than the average slopes under the BPM of 90 and 120, which is different from the expectation.

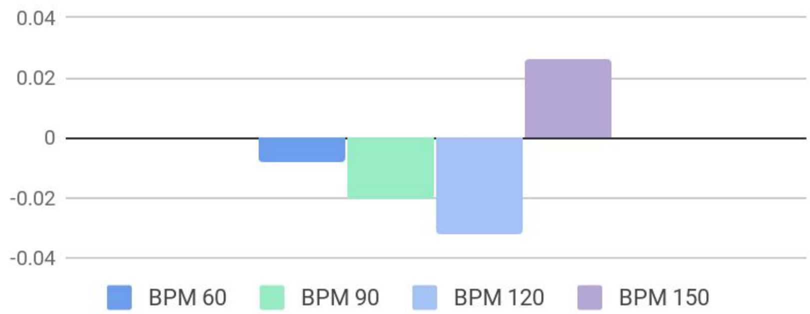


Figure 11: Average linear regression slope of all subjects.

Conclusion

In this work, the relationship between musicians' tempo persistence capabilities and certain aspects of their musical backgrounds was studied. The current results show positive correlations between this ability and the subjects' daily music exposure levels as well as their experiences in musical groups. On the other hand, due to small subject size and abnormal data points, the average performance of drummers appears to be worse than that of non-drummers. Moreover, participants who usually use metronomes actually demonstrate a relatively worse performance, which is another issue that remains to be studied further.

In the future, this study can also take the genres of music into consideration when there is more participants, since different genres of music might have different levels of sensitivity to tempo consistency. Moreover, there should also be some proper strategies of dealing with abnormal data points that will affect the average behavior of a certain group of subjects, so that some issues that are similar to the bad drummer problem in this study can be avoided in the future. Another direction of improvement is to find a better approach to calculate the average linear regression slope, since a positive slope from one subject and a negative slope from another subject might cancel each other out. One possible solution is divide the positive slopes and the negative slopes into different groups and analyze them separately.

References

Freedman, D. A. (2009). *Statistical models: theory and practice (2_{nd} ed.)*. New York, NY: Cambridge University Press.

Jungers, M. K., Palmer, C., and Speer, S. R. (2002). Time after time: The coordinating influence of tempo in music and speech. *Cognitive Processing, 1*(2), 21-35.

Wackerly, D., Mendenhall, W., and Scheaffer, R. L. (2008). *Mathematical statistics with applications (7_{th} ed.)*. Belmont, CA: Thomson Higher Education.

Yan, X., and Su, X. (2009). *Linear regression analysis: theory and computing*. Hackensack, NJ: World Scientific.