Dependency of Pitch Recall and Reproduction on Musical Ability

Feasibility Study

Anandwade R. , Brendal B. , Harr J. , Pandya S.
University of Illinois Urbana-Champaign, Grainger College of Engineering, Department of Physics, PHYS 398 DLP

Abstract

Music is something that is encountered by people nearly constantly in their everyday lives. There are also multiple perceptions of what it means for music to be good. We believe one of these factors depends on the accuracy of each note played by the musician. There have been multiple biological studies[1] conducted on how the brain perceives music and how the auditory processes in the brain perceives music. There is also a neurological study conducted examining pitch memory in musicians and nonmusicians[2]. We intend to combine both of these kinds of studies together by collecting various environmental, biological, and psychological factors about our participants. Throughout this paper we will discuss the accuracy of pitch reproduction and its dependency on the above mentioned factors.

Keywords: pitch recall, pitch reproduction, musical ability

---

1. Introduction and Background Information

1.1 Melodies and Pitch

Musical melodies have two main factors to them, pitch and rhythm.[3] Pitch corresponds to the frequency of the note played while rhythm corresponds to the timing of playing the notes. In order to have a satisfactory aural experience while listening to music, both the pitch and the rhythm have to be accurate and correct. In our work, we will focus specifically on pitch.

Biologically, humans' ability to hear pitches depends on different kinds of specialized cells in the inner ear.[1] That information is then processed within the brain, which leads us to believe there are various neurological[2] and psychological factors that affect the ability to process that sound.

1.2 Categorizing Musicians vs Non-Musicians

Previous work has been done examining what sections of the brain are stimulated[2] in musicians versus non-musicians while recalling and remembering various pitches.

For this reason, we decided to have our subjects self-identify their musical ability.

1.2.1 Absolute and Relative Pitch

Additionally, some people have the ability to hear and identify absolute and relative pitches, while others are referred to as being "tone deaf". Absolute pitch is the ability to hear a frequency and identify exactly what note it is without a reference note and relative pitch is the ability to do so with a reference note. We believe this ability might depend on the musical abilities of the participant. This extra ability arises because melodies depend on set differences in pitches.

1.3 Semitones

Western music divides the range of audible frequencies into a recurring sequence of 12 notes: C, C#, D, D#, E, F, F#, G, G#, A, A#, and B.[4] A semitone is the smallest interval noted in those 12 notes. Mathematically, the ratio of frequencies present in one semitone is $\sqrt[12]{2}$[5]. The radical present here puts musical notes on a logarithmic scale.

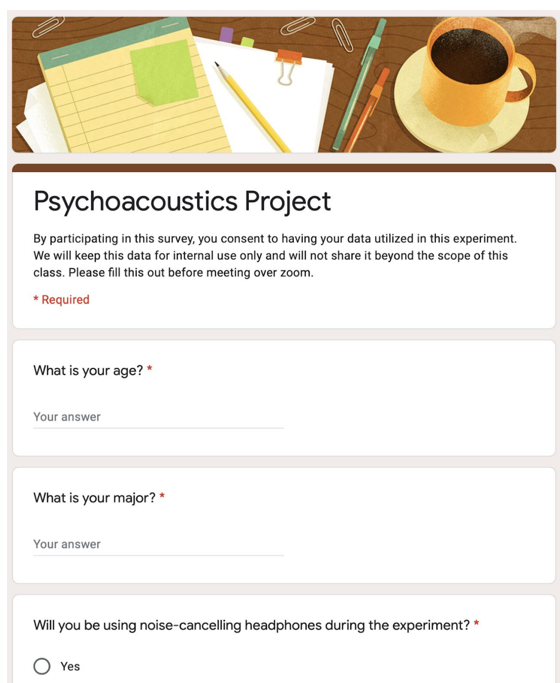
1.4 Behavioral Factors

In addition to discrepancies arising from physiological, neurological, and psychological factors, there are behavioral factors contributing to this study. Different languages and instruments have different tonal qualities, and so a subject's immersion in such a language or use of such an instrument may influence their ability for pitch recall and reproduction. For the sake of narrowing down our argument we have decided to only consider the effect that a subject's experience with a given instrument may cause.

For example, somebody who plays the flute and piccolo would be better accustomed to higher register notes. Whether or not the participants sing is another important factor to consider. In order to produce a note you undergo a physical change. We would also like to consider these discrepancies as factors.

2. Methods

2.1 Survey


Psychoacoustics Project

By participating in this survey, you consent to having your data utilized in this experiment. We will keep this data for internal use only and will not share it beyond the scope of this class. Please fill this out before meeting over zoom.

* Required

What is your age? *

Your answer

What is your major? *

Your answer

Will you be using noise-cancelling headphones during the experiment? *

Yes

Figure 1: Screenshot of Google form used to ask participants questions.

Each participant was pre-screened with a survey written using Google Forms. This consisted of 14 questions, ranging from basic personal information to assessing how musically-inclined one might be. We ended up having to adapt this survey to include additional questions regarding Zoom. With these data, we would be able to make conclusions that relate the accuracy of the subject's results with their self-defined musical ability, as well as the other factors listed in the introduction.

2.2 Procedure

Once the subjects complete the questionnaire, we meet with our subject over Zoom to conduct the experiment. The subject will be tested on 6 notes: a middle and high A, C, and F. Before manipulation

	FREQUENCIES (HZ)		
NOTES	Primary Overtone	Secondary Overtone	Tertiary Overtone
Middle F	176.3	353.8	N/A
High A	446.5	890.1	1,333.8
Middle C	266.7	532.6	798.6
Middle A	222.9	443.9	887.6
High C	528.3	2,115.0	1,056.4
High F	353.2	1,059.6	706.8

Table 1: Frequency data for all notes used in testing

begins we play the note for the subject five times consecutively. During manipulation, the subject is allowed one additional replay of the original note during coarse adjustments.

2.2.1 Test Samples

In order to collect the samples of 6 notes to play for the subjects, we record the 6 notes on the one guitar with the same microphone. These notes are recorded as .bin files. The .bin files are then converted to .wav files. A sample waveform is shown in Figure 3.

2.2.2 Audacity

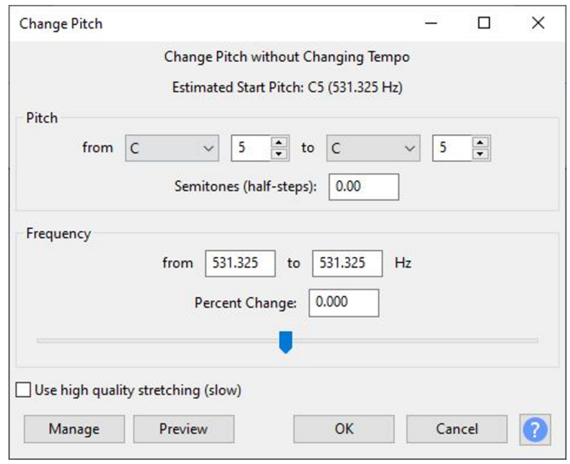


Figure 2: Screen capture of the "Change Pitch" function inside of Audacity used to pitch the notes up and down

Audacity allows a .wav file to be played back at a different pitch by a difference of a certain percentage than the file originally is. We use this function to change the pitch of a note. This process has been standardized by pitching the tone either up 74% (up 9 semitones) or down 74% (down 24 semitones). The Audacity interface is shown in Figure 2.

2.2.3 User Experience

The nature in which this experiment was conducted was less than ideal, and the implications and initial plans will be discussed later. Our user experience is as follows.

The subject meets with the experimenter via Zoom to conduct the experiment. Prior to the meeting, the subject has completed a questionnaire to assess musical ability. The subject signals audio manipulation to the experimenter via hand signals. The subject also verbally adds comments and requests whenever necessary. A "thumbs up" signal is used to indicate that the subject wants to pitch up, "thumbs down" for pitch down, and palm for stop.

The subject navigates frequency differences through coarse changes (intervals of 10% change, 1.82 semitones down/1.65 semitones up) and fine changes (intervals of 1% change, <1 semitone up/down). The subjects are allowed to jump multiple coarse/fine intervals at once should they choose.

The audio playback from Audacity is output directly to Zoom, so it is as if the subject is hearing the sound from their own machine with no audio deprecation or latency issues in the frequency transmitted. This was tested by playing a monotone frequency with an online monotone generator on the host computer and then using the "share computer sound" feature in the Zoom software to play the tone through the subject's computer. The subject was then instructed to try and match the host's frequency with a monotone generator on their own computer and they were able to match the tone exactly.

We recommend all users wear headphones to reduce distractions (noise cancelling preferred) but we understand when they are not available.

2.3 Data Acquisition Code

Our data acquisition is primarily done through data manipulation in python, utilizing python modules pyaudio, scipy, and matplotlib. With the .wav files for each note, a fast fourier transform (FFT) can be done to isolate peak frequencies. We graphically determine the top three most prominent frequencies and quantitatively identify our notes hereafter through these frequencies.

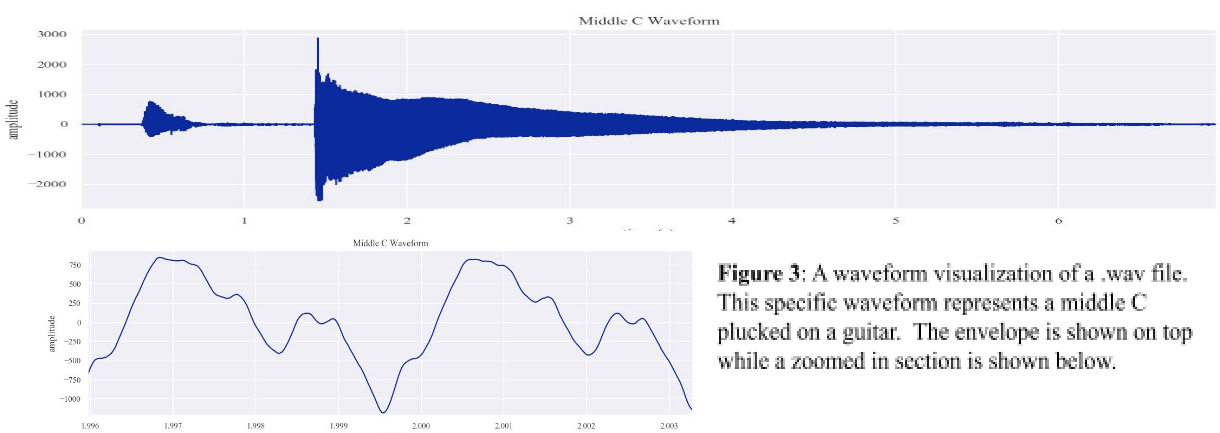


Figure 3: A waveform visualization of a .wav file. This specific waveform represents a middle C plucked on a guitar. The envelope is shown on top while a zoomed in section is shown below.

After testing subjects, we utilize the FFT function present within Audacity to record their output frequencies. It is worth noting that the two FFT processes might differ, due to a difference of ~2 Hz.

3. Hardware

3.1 Breadboard

The first stage of our hardware was to get different electrical components installed and communicating on the breadboards. The breadboard served as a tool to learn how the different hardware and software components worked and to become comfortable with writing code to accomplish different tasks through the Arduino and with the installed sensors. The components used are as follows: Adafruit MAX4466 Electret Microphone, Adafruit BME680 Temperature, Humidity, Pressure, and Gas Sensor, MicroSD card breakout board, LCD screen, 3x4 Matrix Keypad, MCP4725 Breakout Board 12-Bit Digital to Analog Converter, PAM8302A Audio Amplifier with Speaker, and INA219 DC Current Sensor. The Aruduino used was the Arduino MEGA2560.

3.2 MAX4466 Electret Microphone

The MAX4466 Electret Microphone was meant to be used to record the guitar notes that would then be played back for the subjects during the experiment. Arduino code[6] provided by George Gollin was used which saved audio recorded from the microphone to the SD card as .bin files. The MAX4466 refers to the amplifier on the breakout board, which has a 20-20kHz electret microphone soldered to the breakout board itself (refer to Appendix A).

3.3 Printed Circuit Board

The next stage was to build printed circuit boards (PCB) to be more compact. We only integrated sensors and other components that would be necessary to our specific project. These components were as follows: the Arduino

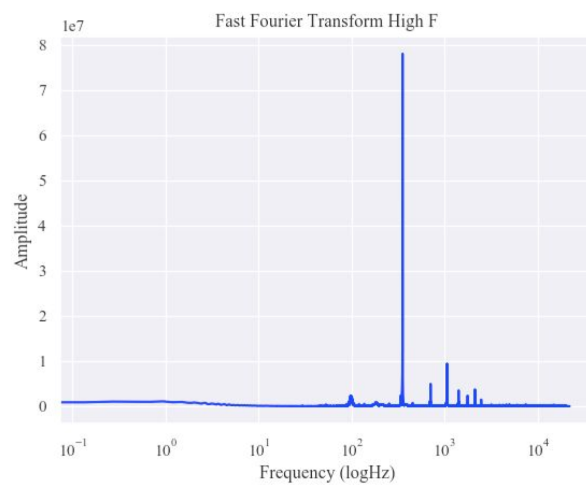


Figure 4: Example fast fourier transform plot. We record frequency in logHz from the top three peaks.

MEGA2560, the Adafruit electret microphone for recording guitar notes, the LCD for displaying information from the code that would run on the Arduino, the SD card breakout board for saving recorded guitar notes as .bin files, the

INA219 for current control from the attached battery pack (see Appendix A). At this stage, the PCB would be our working machine, and the breadboard with its various components would serve as a troubleshooting reference for the PCB.

3.3 3D Printed Case

The components listed above are sensitive to external vibrations and require protective casing. For this reason, we decided to house our PCB inside a 3D printed case. We decided to make adjustments to the case[6] designed by Professor Gollin. Figure 5 below shows the proposed lid to the case

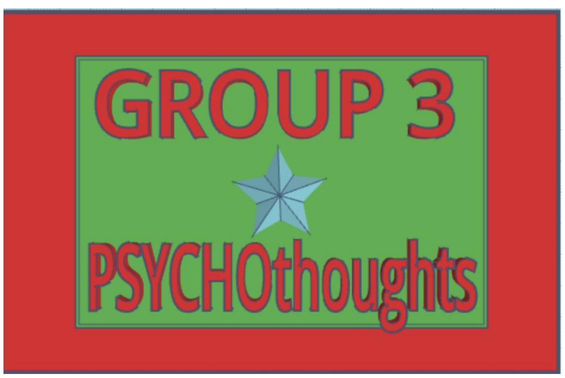


Figure 5: Proposed lid for 3D printed case.

Unfortunately, due to the outbreak of COVID-19, we were unable to print this case.

4. Data Analysis and Results

4.1 Data tables

Each participant gave us two sets of data, table 2 shows the results for the frequencies picked. In addition to this data we have information from the Google form. This gives us one set of quantitative and one set of qualitative data to analyze.

4.2 Musical Ability

Figure 6 shows graphed results for 24 participants. We used two different kinds of markers to denote the guesses by participants with self-defined musical abilities and lack thereof. Our criteria for identifying this was based on whether or subject could sing, play an instrument, or was a member in a band.

Figure 7 shows results within our musically talented people based on the type of instrument they play. The breakdown of our data suggested we group into 3 categories: piano, woodwind, and string. In the future, if more robust data is available there are many more instrument types that could be included.

Figure 8 further breaks down our results of Figure 6 by sorting our musically inclined subjects that play instruments by the tonality of their instruments.

We classified the types of instruments as piano, woodwind, and strings. The woodwind classification consists of the flute, saxophone, baritone saxophone, clarinet, and recorder. The string classification consists of ukulele, guitar and violin.

	FREQUENCIES (HZ)			
NOTES	Primary	Secondary	Tertiary	Subject Guesses
Middle A	222.9	443.9	887.6	195.605
High A	446.5	890.1	1,333.8	214.904
Middle C	266.7	532.6	798.6	290.414
High C	528.3	2,115.0	1,056.4	355.645
Middle F	176.3	353.8	N/A	252
High F	353.2	1,059.6	706.8	541.952

Table1:: Frequency data for all notes used in testing

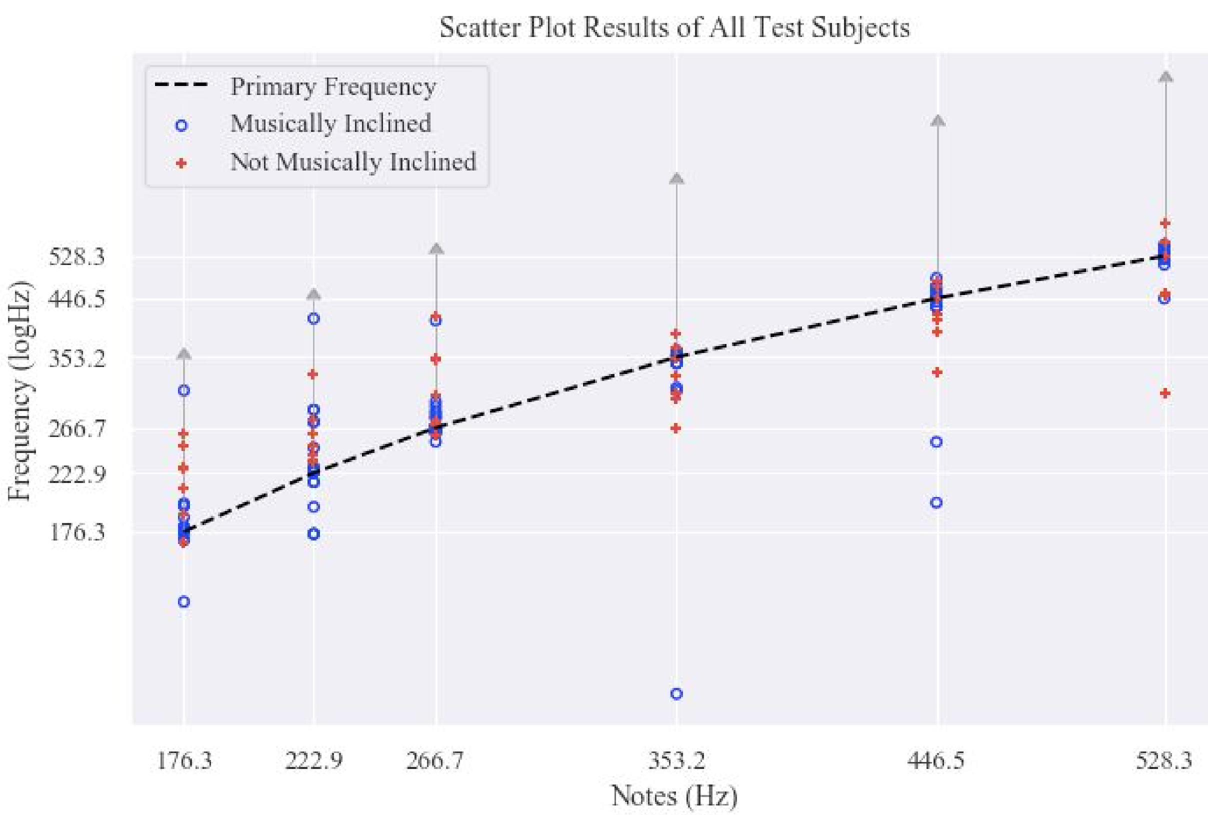


Figure 6: This plot graphically shows all 24 subjects' test results. Vertical gray arrows represent one octave

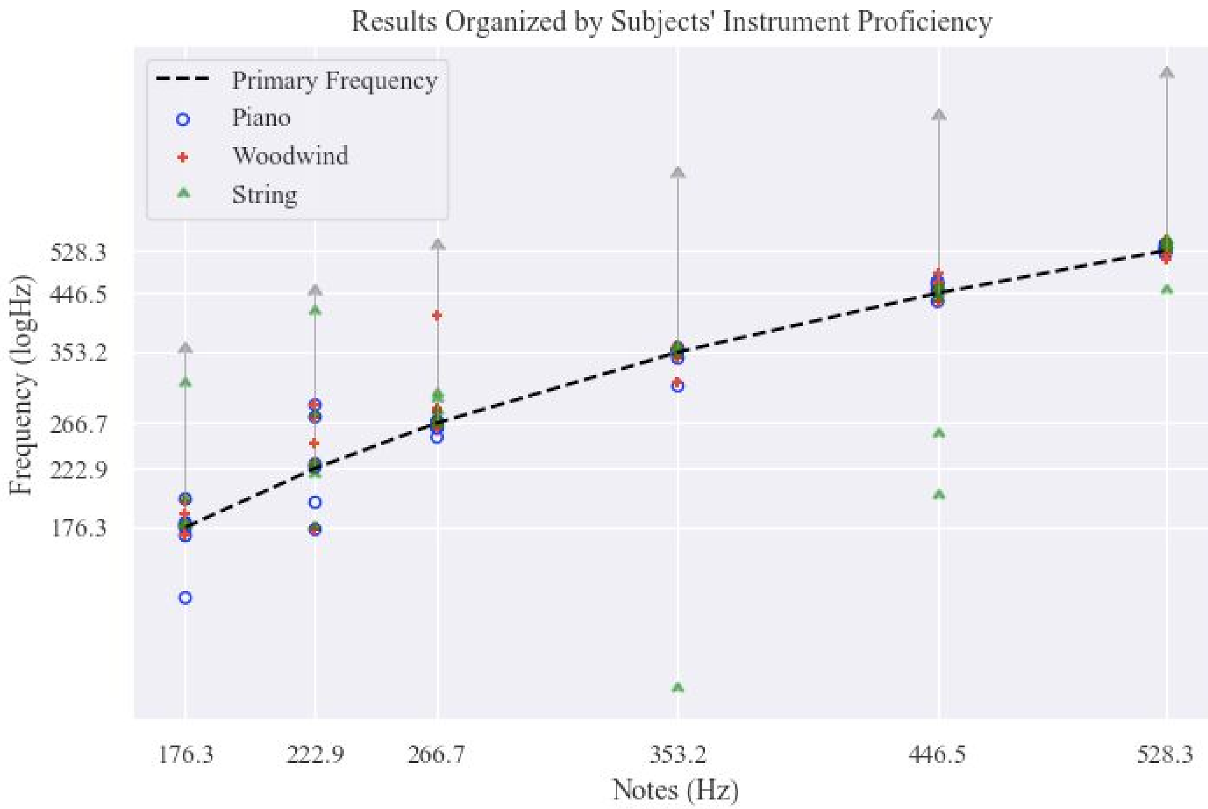


Figure 7: This plot graphically shows results for musically talented subjects based on the instrument they play. Vertical gray arrows represent one octave.

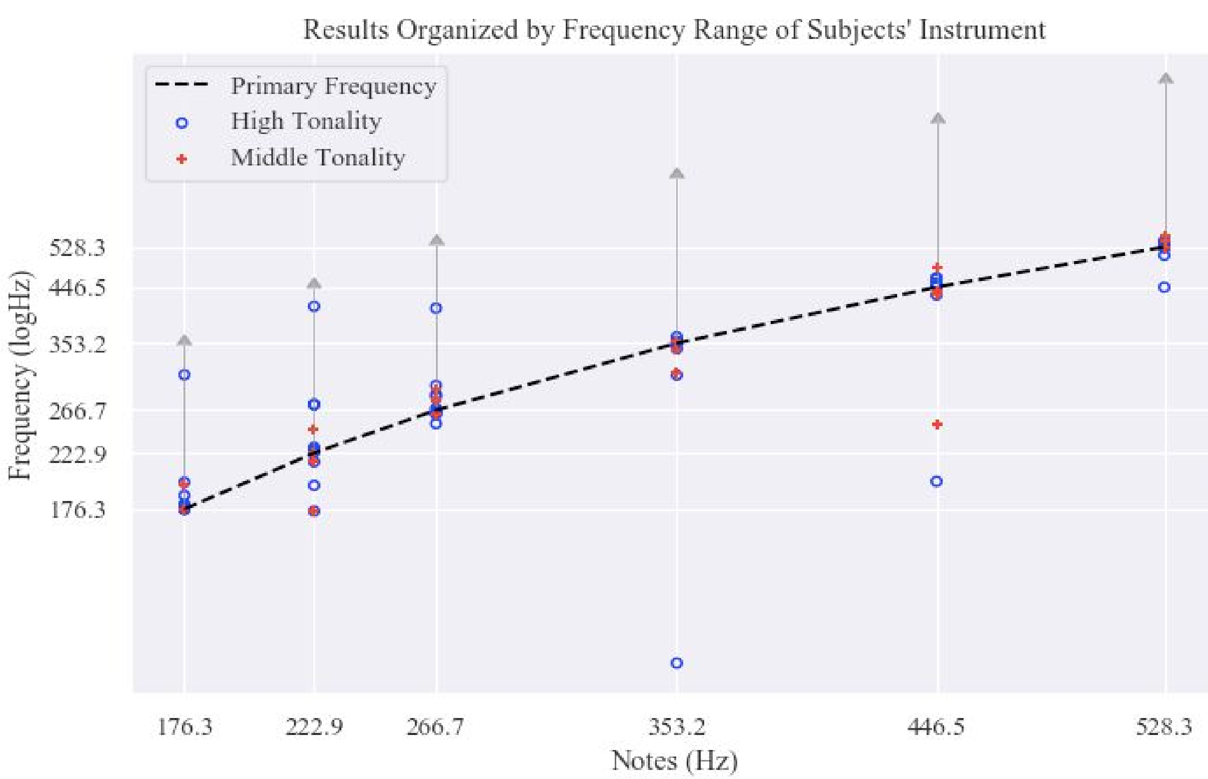


Figure 8: This plot graphically shows results for musically talented subjects based on the tonality of the instrument they play. Vertical gray arrows represent one octave.

The tonality of instruments were classified as either high tonality or middle tonality. We sorted the instruments played by subjects into these categories based on the range of frequencies the instruments cover. High tonality includes ukulele, flute, clarinet, and violin. Middle tonality includes piano, guitar, saxophone, and baritone saxophone.

Looking at Figure 7, there are multiple string instrument subjects who were far from the actual frequency. There was a much higher number of participants playing string instruments than woodwind, and of those subjects, one in particular was an outlier and consistently incorrect. If this study was a larger study, this error could have been reduced.

Moving on to Figure 8, the higher tonality points are far from the actual frequencies of the middle notes. We suspect this is because subjects who play these higher toned instruments are not used to working with lower frequency ranges.

4.4 Case Study One

One of our tested subjects was very quick to understand our testing procedure. The plot for just their results are shown in Figure 9. From the plot you can see that they had higher accuracy for every note but high A (446.5). On completion of this note, the subject remarked that they had simply "forgotten the note." by that point.

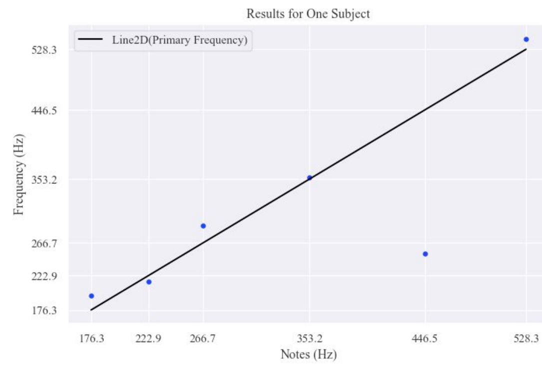


Figure 9. Guessed frequencies vs ground truth for a single subject.

5. Discussion

5.1 Pre-Pandemic goals

The experiment itself was meant to be run as follows. A tone would be played on a guitar live with the test subject. After listening to the tone played, the subject would then manipulate a dial or slider to adjust a playback tone in their attempt to match the frequency that the instrument produced. Due to the stay-at-home orders from the COVID-19 virus, we have had to change our method of data collection and testing.

5.2 Zoom

Zoom is an advanced video conferencing software, but comes with its fair share of setbacks. With this method of testing, we were unable to make each test completely uniform. Ideally our participants would be using the same device to optimize the audio quality throughout the trial. Unfortunately we could not control this factor among all of our test subjects. The speaker or headphone quality might influence the quality of the audio for each participant. There is also the issue of internet connection, which could lead to imperfections in the audio and video quality.

In the interests of time and feasibility, we tested less people than we wanted to. Ideally we would have liked to test at least 50 subjects, with a reach of over 100 to allow us to clearly see trends in the data. With less data, we want to ensure that our conclusions are accurate. We intend to put error bars on our plots to gauge uncertainty and precision in our estimates.

We wanted this experiment to be a test on pitch recall and reproduction, not memory, so we had to ensure that our test subjects could hear the audio sample an appropriate amount of times. This was difficult with the Zoom format, as we had to forego our original plan of giving the participant control over an adjustable slider. With all of the changes we've had to make to our procedure, we've decided to give participants the opportunity to provide feedback.

5.3 Privacy

Regarding our survey, we acknowledge that by using a Google product, we've had to surrender some of our privacy. With time and convenience in mind, we've elected to use Google Forms.

5.4 Future Changes and Limitations

We also would have liked to test using more instruments to explore differences in timbre. Since many of our participants played a variety of musical instruments, we are interested to see how one might perform better with the tone of one instrument compared to another. There also exists the factor of language, and how upbringing might influence the way one hears and identifies tonal differences.

5.1 Statistical Analysis

We calculated the mean, median, and standard deviation for each note. The median tended to be closest to each reported frequency. Our participants scored the best for the "High C" note.

Note	Middle A	Middle C	Middle F	High A	High C	High F
Mean	243.6	292.9	198.0	416.9	523.8	331.5
Median	228.3	276.9	179.6	450.0	531.3	355.6
Std Deviation	48.98	47.7	40.3	76.1	37.2	64.4

Table 2: Mean, median, and standard deviation for each note.

By adopting a Bayesian approach, we were able to build a likelihood function from our data. We hypothesized that musically inclined people would arrive at a frequency that was within one semitone of the note's actual primary frequency. After defining these parameters, we treated each trial as a case of success or failure, where success was the subject's ability to guess within this range. Thus, the likelihood function represents the probability of obtaining these results from our sample population within the selected parameters.

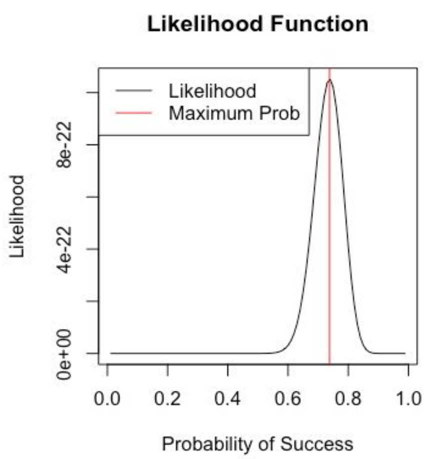


Figure 10: Plotted likelihood function.

We also plotted the natural log of the likelihood function, for simplification, as it is a monotonically increasing function. The probability which maximized our likelihood came out to be 0.738.

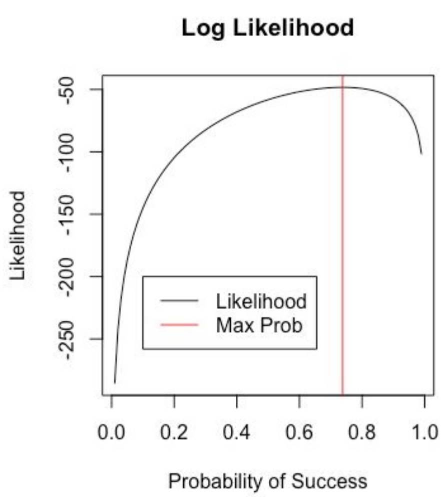


Figure 11: Plotted log-likelihood function.

6. Conclusion

Musically predisposed people were at an advantage in identifying the original tone played in this experiment. Through testing we noticed that people who had experience singing or playing instruments were able to recreate the original tone with hums or whistles, and were able to reproduce that hum or whistle until they arrived at a tone which closely matched the original. However it's worth noting that some of these musically inclined people were subject to identify the correct note, but the wrong octave, which explains some of the extremities in our plots. The not musically inclined subjects generally struggled. However, a few of them who performed well attributed their success to good memory and well-timed requests to reproduce the original tone.

During testing, it was noticed that the subjects soon forgot the original note after hearing the new changing note several times. On several occasions, when the subject thought they might be close, then heard the original note one more (and final) time, they realized how far from the original they actually were; like “walking in the dark, thinking you remembered the path, but realizing afterward that you had strayed far from it,” one person noted.

7. Acknowledgements

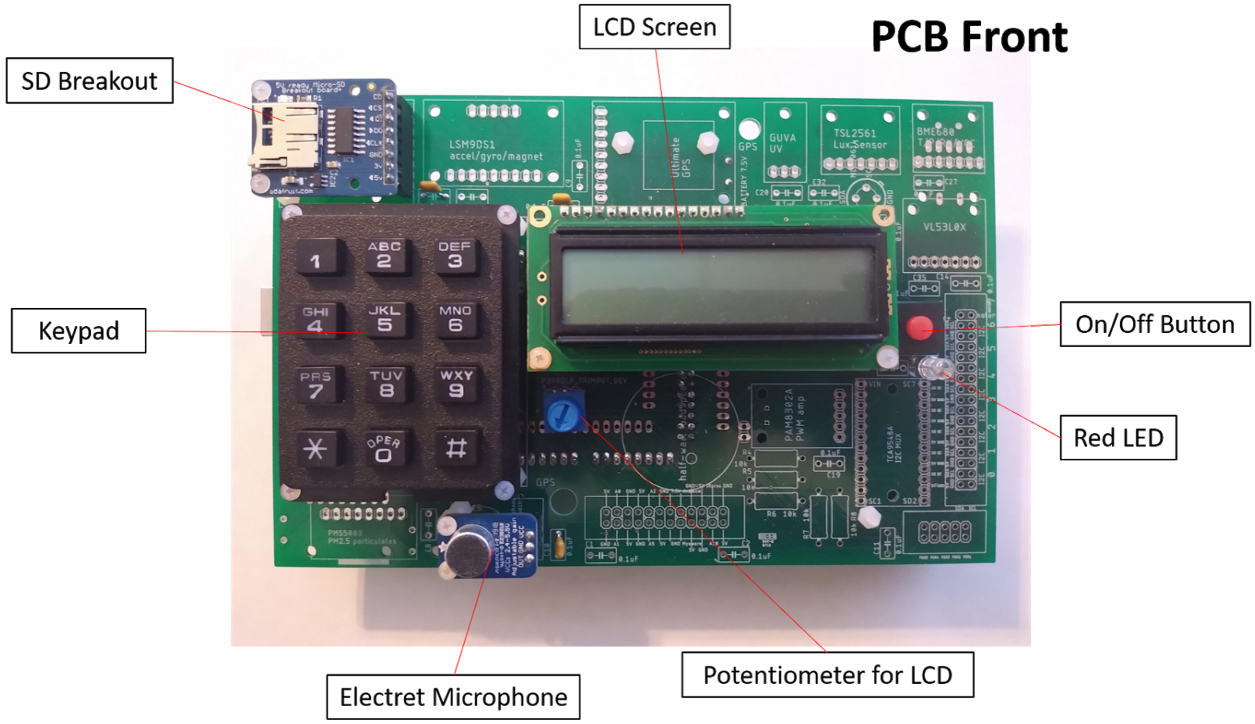
We acknowledge Professor George Gollin and the PHYS 398 class at the University of Illinois for the resources and source code provided in setting up our

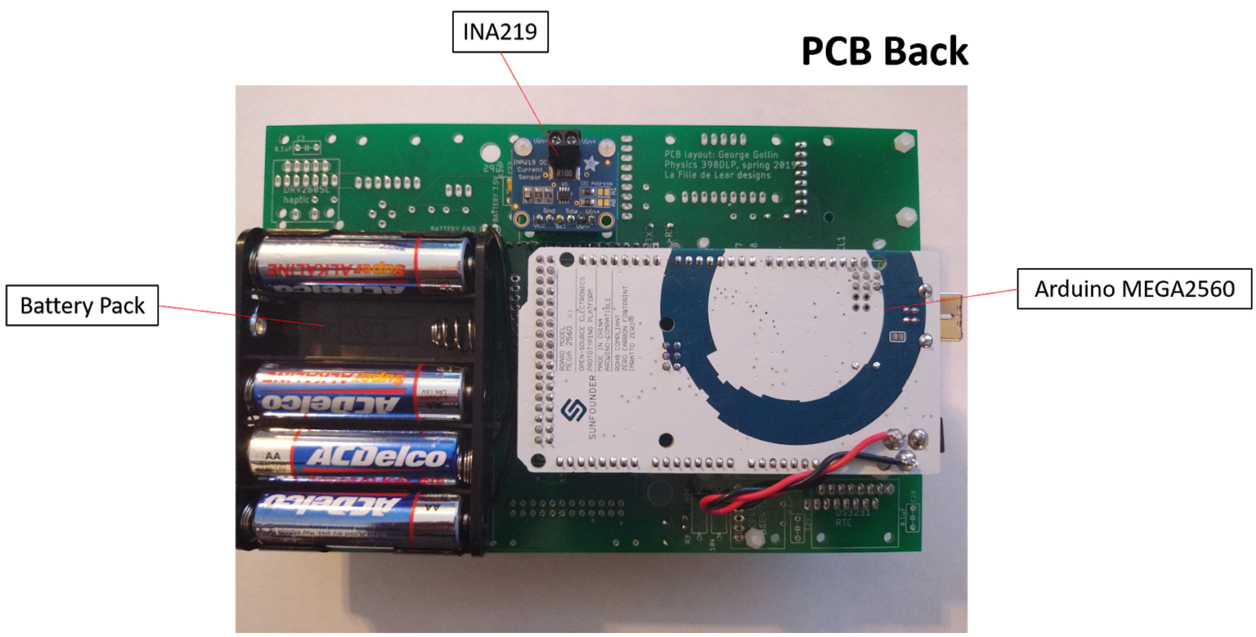
experiment. We additionally acknowledge software and services provided by Arduino, Google, Adafruit, Zoom and Audacity.

8. References

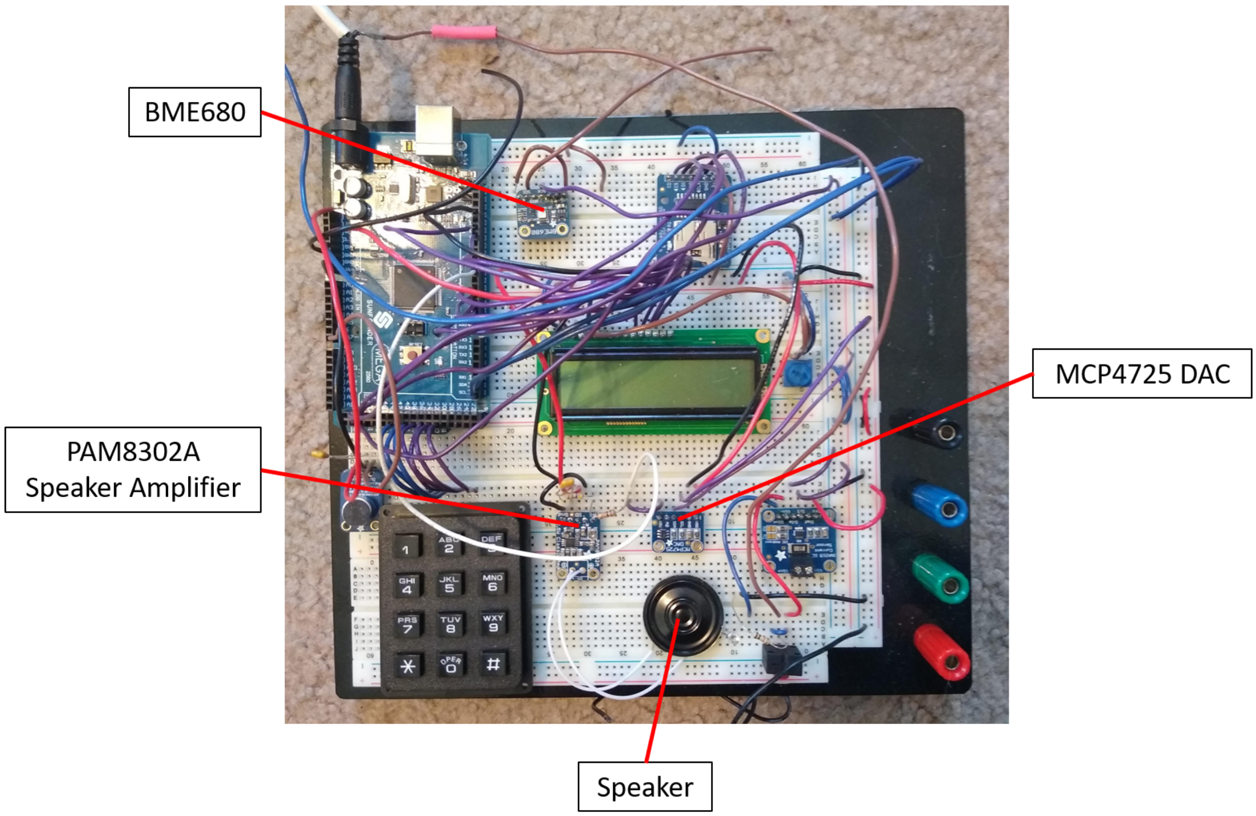
1. University of Virginia Health System. (2014, May 20). Pitch-detection secrets of the inner ear revealed by research. ScienceDaily. Retrieved April 2, 2020 from www.sciencedaily.com/releases/2014/05/140520115508.htm
2. Gates B, Takahashi, IK, Bush, GW. The use of Word in different situations. Microsoft Home Journal. 2007;13: 10-19.
3. Kliewer, Vernon (1975). "Melody: Linear Aspects of Twentieth-Century Music", Aspects of Twentieth-Century Music, pp. 270–301. Wittlich, Gary (ed.). Englewood Cliffs, New Jersey: Prentice-Hall. ISBN 0-13-049346-5
4. Tone and Semitone. (n.d.). Retrieved from https://www.simplifyingtheory.com/tone-semitone-sharp-flat/
5. Mathematics and Music - Study. (n.d.). Retrieved from http://www.simplifyingtheory.com/mathematics-and-music/
6. Design Like a Physicist. (n.d.). Retrieved from https://courses.physics.illinois.edu/phys398DLP/sp2020/code.asp

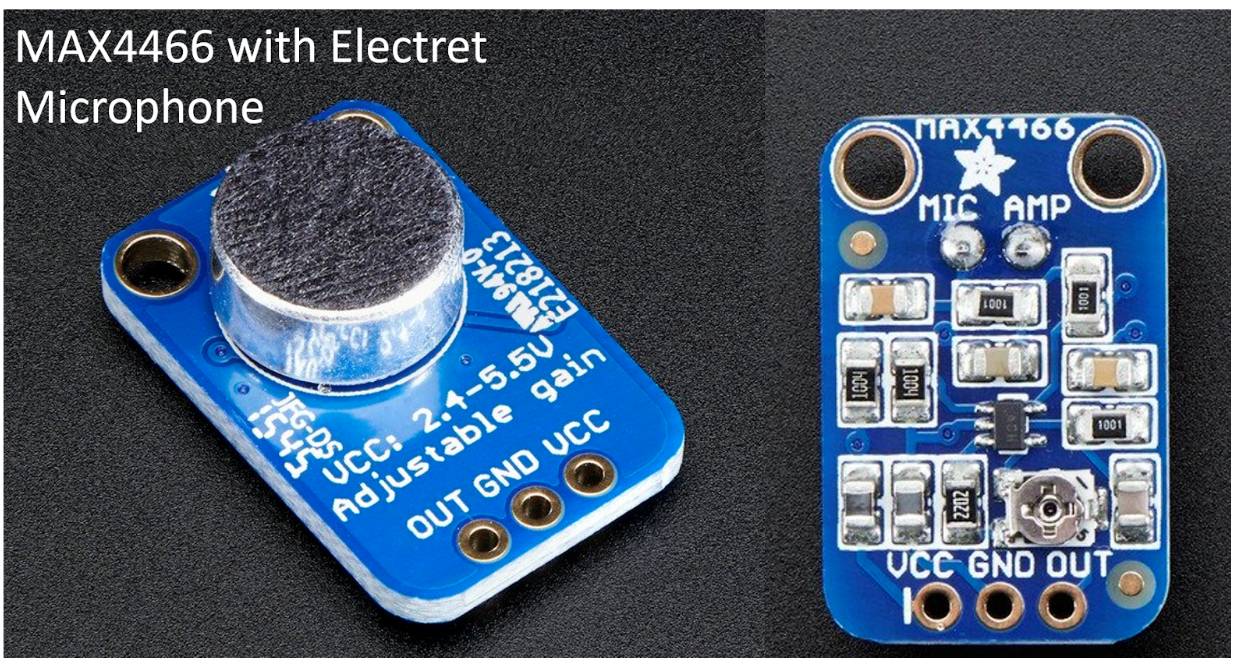
Appendix A: PCB Design and Layout





Breadboard


BME680
MCP4725 DAC
PAM8302A
Speaker Amplifier
Speaker


MAX4466 with Electret
Microphone
VCC: 2.4-5.5V
Adjustable gain
OUT GND VCC
MAX4466
MIC AMP
VCC GND OUT

Appendix B: Survey Questions

1. What is your age?
2. What is your major?
3. Will you be using noise-cancelling headphones during the experiment?
4. Do you sing?
5. If you answered yes, do you have formal training, or are you self-taught?
6. Do you play a musical instrument?
7. If yes, what instrument(s) do you play?
8. If yes, how long have you been playing this instrument?
9. If you answered yes, do you have formal training, or are you self-taught?
10. Do you have any hearing-related disabilities?
11. Do you sing along to songs as you listen to them?
12. What genre(s) of music do you listen to? Check all that apply.
13. Do you regularly watch musicals?
14. Do you play in a band?