The Effects of Auditory Feedback on Instrumentalists' Timbre Production

Madeline Huberth
University of Michigan, Ann Arbor
mhuberth@umich.edu

December, 2011

Abstract

Previous work has shown that the absence of auditory feedback in classical musicians has minimal effect on performance factors such as note accuracy, timing, and dynamics. This study explores the extent to which an instrumentalists' timbre is affected by changes in auditory feedback. Nine cellists were recorded playing scales as well as excerpts from the orchestral literature under various feedback conditions: with their hearing completely impeded, masked by noise, and masked by orchestral recordings of the excerpts, simulating orchestral playing. Data analysis focused on changes in the spectral energy distribution and relative harmonic amplitudes, tendencies within each individual, across participants, and across the selected notes of the excerpts. Changes in relative harmonic strengths across the three conditions were analyzed. Overall, the results show that auditory feedback condition has little significant effect on timbre production, with the majority of the comparisons of spectral energy distribution between conditions yielding non-significant results. However, in the scale experiment, 25% of the comparisons between conditions, (56 comparisons total - 8 participants, 7 notes each,) showed a significant increase in spectral energy in the low harmonics when feedback was removed (ρ: .004 –.04), while 16% of comparisons showed a significant decrease in spectral energy in the high harmonics (ρ: .004–.025). This result suggests that, when feedback is removed, cellists are playing further away from the bridge; however, further study is needed to confirm this conclusion.

ACKNOWLEDGEMENTS

I am so grateful to those people who made this research possible and to those who made my experience during my undergraduate one that I will always remember fondly.

My utmost gratitude goes to my advisor, Dr. Timothy McKay, whose expertise, enthusiasm, and inspiration have been truly invaluable. Throughout my thesis-writing period, he provided sound advice, encouragement, and a lot of good ideas. More so than anyone, he guided my academic education, and the things I have learned from him are immeasurable. It is difficult to overstate my gratitude to him. Thank you for pushing me.

I would like to thank all those who offered me technical support and advice during my research. Dr. James Beauchamp of the University of Illinois Urbana-Champagne kindly supported me as I learned to work with his software. I am deeply thankful to him. I also appreciate the programming assistance I received from Marshall Weir, Jay Wren, Scott Reed, and Adam Becker, who patiently dealt with all of my questions.

I thank Dr. Donna Wessel Walker, my academic advisor. She encouraged me to combine my interests in unconventional ways, and has suggested directions and opportunities to me that have shaped my experience at Michigan. Additionally, I am indebted to Richard Aaron, my cello teacher and friend. He helped me grow as a musician, and time and time again, instilled confidence in me.

To my friends at the Telluride House, my family away from home, thank you for your conversation, personality, and humanity. Finally, I would like to thank those closest to me, my parents, for their undying love and for fostering an environment in which pursuing my interests felt natural.

Contents

List of Figures

1 Introduction

It is widely thought that in order to deliver a successful performance, one where the music has all of the detail the musician intends it to have, a musician must hear his sound during the performance to evaluate and adjust it. Most may argue that altered or absent auditory feedback would disrupt the performance process, creating discrepancies in timing, relative dynamics, intonation, as well as expressivity and tone quality. However, musicians often perform in situations where their normal auditory feedback is altered by some external circumstance and as a result, hear their own sound with far less detail than they do when they play by themselves. For instance, a bassist in a modern symphonic orchestra plays in a section of six to ten basses, where he is expected to adjust his playing so that his sound blends with the rest of the ensemble. Consequently, he often hears the sound of the section above his own playing. Does this reduction in his ability to hear his own playing have the potential to disrupt his performance? Does the tone quality that he produces change as a result of altered auditory feedback?

The necessity of auditory feedback in musical performance is an area of recent research. Prior studies used pianists as participants and tested note accuracy, rhythm, and dynamic control when auditory feedback was removed. They concluded that the removal of auditory feedback in musicians minimally disrupts these aspects of their performance. Other studies altered auditory feedback by either preponing or delaying the feedback, and found that such alterations more significantly disrupt the musicians' performance.

However, the effects of auditory feedback removal on timbre have not been studied in great detail. This study seeks to expand upon previous research, and tests the hypothesis that aural feedback affects the timbre of a musician's performance, as measured by harmonic spectral energy distribution. Cellists are the participants, as they have more control over the timbre of their sound output than those playing on keyboards. In the study, 'feedback' and 'auditory feedback' will be used interchangeably, as will 'timbre' and 'quality'.

1.1 Literature Review

Research on the necessity of auditory feedback has shown that when feedback is present but altered, such as in delayed auditory feedback (DAF), the musical performance is adversely affected (see Palmer 2006; Yates, 1963). Pfordresher and Palmer (2006) tested the effects of DAF on pianists and found it to disrupt note accuracy significantly, but cause less disruption in timing, regardless of the "feedback distance," or the length of time of delay of the feedback. However, when pianists received feedback at such a delay where the metrically accented beats from the feedback synchronized with the unaccented beats they played, the weight with which the metrically accented beats were performed was disturbed. Studies have also evaluated the effects of the complete absence of auditory feedback. In several experiments, the lack of auditory feedback had no significant effect on musical performance in the cases of professional musicians (Finney, 1997; Finney & Palmer, 2003; Gates & Bradshaw, 1974; Pfordresher, 2005; Repp, 1999; Repp & Knoblich, 2004) and in untrained performers (Pfordresher, 2005). The focus of these experiments were on variables that have definitive correctness, such as note accuracy; however, Repp (1999) focused specifically on the change in expressivity of the performance when feedback was absent, using horizontal and vertical timing, horizontal and vertical dynamics, and pedaling as parameters. He reported small effects of auditory feedback absence on expressive parameters of production, where the only performance parameter that was seriously affected by feedback deprivation was pedal use – most pianists changed the pedal less often when they could not hear themselves. Any changes in the other parameters could be attributed to subtle impairment of motor control in the absence of feedback, or perhaps less motivation to play expressively in the absence of sound. After the data had been collected, a group of objective pianist listeners were asked to identify which of the performances was played without auditory feedback. They also stated their degree of surety, with the choices being "quite confident," "not so sure," and "just guessing," given values of 3, 2, and 1, respectively. The overall percentage of correct judgments was 63.5%, with the scores of the individual participants ranging from 52% to 77%. The average confidence rating was 2.2, but participants' confidence in correct judgements (2.22) was not significantly higher than their confidence in incorrect judgments (2.16), confirming the relative subtlety in the effects of feedback deprivation on performance.

There has been no documented research on how a lack of feedback affects a sound's timbre in performance. The majority of the aforementioned studies were carried out on MIDI electronic keyboards. There are significant advantages to such a tactic: auditory feedback can be easily and completely removed, and the originally silent performance can be easily recorded, analyzed, and even played back with sound. Although utilizing MIDI technology facilitates the analysis of timing and note accuracy as well as a rough study of dynamics, it does not allow for the study of timbre, since timbre is constant on a MIDI keyboard. When the timbre remains constant, a significant portion of the artistry - the creation of a personal tone quality - is lost. The lack of documented research on the effects on timbre of the absence of auditory feedback in musical performance motivates this study.

1.2 The Definition of Timbre

In his book "On the Sensations of Tone," or "Die Lehre von den Tonempfindungen," Hermann Helmholtz postulates that musical tones have three distinguishing qualities, with his original German being: *1. die Stärke, 2. die Tonhöhe,* and *3. die Klangfarbe*, which most literally translates to *1. force/intensity, 2. pitch,* and *3. tone color.* Force is described by the amplitude of a sound, and thus, its loudness, while a tone's pitch depends on the length of time, or the *period,* T, in which a single vibration is completed. The *frequency,* f, of a pitch is the inverse of its period, $f = 1/T$. Musical tones are said to be higher the greater their frequency, that is, the shorter their vibrational periods [9].

Tone color, or *timbre*, is more difficult to define. The American National Standards Institute (1973) defines timbre as "...that attribute of auditory sensation in terms of which a listener can judge that two sounds, similarly presented and having the same loudness and pitch, are different." Helmholtz states that, since the force of the sound depends on the amplitude of the vibration and the pitch depends on the vibrations length in time, the only possible hypothesis is that the quality of tone should depend on the sound content, or frequency content, within a single period of vibration [9]. Timbre can be qualitatively grasped by considering many different instruments all playing the same pitch at the same loudness; whether the pitch is played by a trumpet, a violin, a piano, or the human voice, the ear has no trouble distinguishing the instrument.

1.2.1 Vibrations of the Strings

Any vibrating structure's motion can be described by the superposition of a number of separate free vibrations. In the case of musical instruments, when the vibrations have definite pitch (such as the sound from a wind, brass, string, or keyboard instrument, as opposed to a struck percussion instrument such as the snare drum), the most prominent vibrational components typically hold the simple ratios 1:2:3:4 (etc). These separate vibrational components are called *harmonics*, where the second, third, and fourth (etc.) components are *overtones* of the first harmonic. The first harmonic is known as the *fundamental frequency*, while the second harmonic is known as the *first overtone*, the third harmonic is known as the *second overtone*, etc.

In the case of string instruments, harmonic vibration originates in the motion of the strings. In this section, we will consider the vibration of a string fixed at both ends. The displacement of the string from its original position, y, is a function of both time t and position x along the string. At any position x, the angle $\theta(x)$ between the string and the horizontal satisfies $\tan\theta(x) = \frac{\partial y}{\partial x}$ and, provided that $\theta(x)$ never becomes large, the motion of the string will essentially be determined by the wave equation:

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}, \tag{1}$$

where $c = \sqrt{T/\rho}$. T is the tension in the string (given in Newtons = kg m/s^2) and ρ is the linear density of the string (in kg/m). Mathemetician Jean-le-Rond d'Alembert (1717-1783) discovered a simple method for finding the general solution to the wave equation. Making the change of variables:

$$u = x + ct, \quad v = x - ct, \tag{2}$$

it follows that, by the multivariable form of the chain rule, we have:

$$\frac{\partial y}{\partial t} = \frac{\partial y}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial y}{\partial v}\frac{\partial v}{\partial t} = c\frac{\partial y}{\partial u} - c\frac{\partial y}{\partial v}. \tag{3}$$

Differentiating again yields

$$\begin{aligned}
\frac{\partial^2 y}{\partial t^2} &= \frac{\partial}{\partial u}\left(\frac{\partial y}{\partial t}\right)\frac{\partial u}{\partial t} + \frac{\partial}{\partial v}\left(\frac{\partial y}{\partial t}\right)\frac{\partial v}{\partial t} \\
&= c\left(c\frac{\partial^2 y}{\partial u^2} - c\frac{\partial^2 y}{\partial u \partial v}\right) - c\left(c\frac{\partial^2 y}{\partial v \partial u} - c\frac{\partial^2 y}{\partial v^2}\right) \\
&= c^2\left(\frac{\partial^2 y}{\partial u^2} - 2\frac{\partial^2 y}{\partial u \partial v} + \frac{\partial^2 y}{\partial v^2}\right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\frac{\partial y}{\partial x} &= \frac{\partial y}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial y}{\partial v}\frac{\partial v}{\partial x} = \frac{\partial y}{\partial u} + \frac{\partial u}{\partial x} \\
\frac{\partial^2 y}{\partial x^2} &= \frac{\partial^2 y}{\partial u^2} + 2\frac{\partial^2 y}{\partial u \partial v} + \frac{\partial^2 y}{\partial v^2}.
\end{aligned} \tag{4}$$

Then, the wave equation becomes

$$c^2\left(\frac{\partial^2 y}{\partial u^2} - 2\frac{\partial^2 y}{\partial u \partial v} + \frac{\partial^2 y}{\partial u^2}\right) = c^2\left(\frac{\partial^2 y}{\partial u^2} + 2\frac{\partial^2 y}{\partial u \partial v} + \frac{\partial^2 y}{\partial v^2}\right). \tag{5}$$

or

$$\frac{\partial^2 y}{\partial u \partial v} = 0. \tag{6}$$

Integrating this equation directly, the general solution is given by $y = f_1(u) + f_2(v)$. Substituting back, we obtain

$$y = f_1(x + ct) + f_2(x - ct) \tag{7}$$

where the function f_1 represents a wave traveling to the left and the function f_2 represents a wave traveling to the right, each with velocity c.

Since the left and right ends of the string are fixed, the boundary conditions are such that $y = 0$ at $x = 0$ and $x = l$. Requiring these conditions yields

$$0 = f_1(ct) + f_2(-ct) \tag{8}$$

for all t, so that

$$f_2(\lambda) = -f_1(-\lambda) \tag{9}$$

for any value of λ. Thus

$$y = f(x + ct) - f(ct - x). \tag{10}$$

The physical interpretation of this is that, at the boundary, the wave traveling to the left returns inverted as a wave traveling to the right. Substituting the other boundary condition, $x = l$, $y = 0$ gives $f(l+ct) = f(ct-l)$ for all t, so that

$$f(\lambda) = f(\lambda + 2\lambda) \tag{11}$$

for all values of λ. Equation 11 means that the function f appearing in d'Alembert's solution is periodic with period $2l$, so that f has a Fourier series expansion. So, for example, if only the fundamental frequency is present, then the function $f(x)$ takes the form $f(x) = Ccos((\pi x/l) + \phi)$. If only the nth harmonic is present, then we have $f(x) = C_n cos((n\pi x/l) + \phi)$,

$$y = C_n cos\left(\frac{n\pi(x + ct)}{l} + \phi\right) - C_n cos\left(\frac{n\pi(-x + ct)}{l} + \phi\right). \tag{12}$$

The theory of Fourier series allows us to write the general solution as a combination of the above harmonics. Using manipulations of trigonometric identities, we can rewrite the nth harmonic solution (Equation 12) as

$$y = 2C_n sin\left(\frac{n\pi x}{l}\right) sin\left(\frac{nct}{l} + \phi\right). \tag{13}$$

The frequency of the nth harmonic is given by $2\pi\nu = n\pi c/l$, or replacing c by its value $\sqrt{T/\rho}$,

$$\nu = (n/2l)\sqrt{T/\rho}. \tag{14}$$

Thus, the frequency of a stretched string is inversely proportional to its length, directly proportional to the square root of its tension, and inversely proportional to the square root of the linear density [10].

1.3 Fourier Analysis and Timbre

A branch of mathematics called "harmonic analysis" breaks up any curve, no matter how complicated it may be, into its constituent simple harmonic curves [14]. The theorems were developed by J. B. J. Fourier (1768-1830), and describe the original sample from a function in the time-domain to a function in the frequency-domain. The output is a representation of the frequencies contained in the original sample.

If $f(t)$ is a real of complex valued function of a real variable t, then its $Fourier\ transform$, $\hat{f}(\nu)$, is the function of a real variable ν defined by

$$\hat{f}(\nu) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i\nu t}\, dt \tag{15}$$

The *energy density* at a particular value if ν is defined to be proportional to the square of the amplitude. So,

$$Energy\ Density \propto |\hat{f}(\nu)|^2. \tag{16}$$

Integrating this quantity over an interval will measure the total energy corresponding to frequencies in this interval.

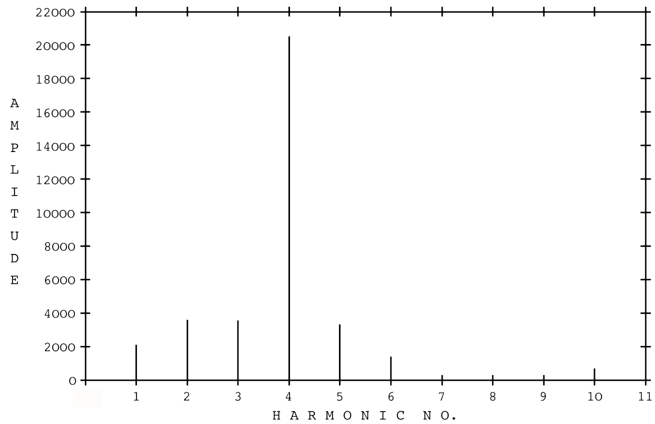
The distribution of energy as a function of frequency for a sound is called its *spectrum*, and determines the timbre of the emitted tone. Features of the spectrum that would affect timbre include the range between harmonic and noiselike components of the spectrum, the spectral envelope of the sound, and the time envelope of the sound (rise, duration, and decay) [12]. For pitched instruments that can produce a constant-amplitude tone though, (cello, flute, trombone,) the relative amplitude of harmonics in the spectrum is especially important; it is what distinguishes one instrument from another. See Figure 1 for a depiction of the relative harmonic strengths in the spectrums of the oboe, trumpet, and violin.

There exist several strategies to describe the spectrum of a sound, described in the Table 1 [13]. Our approach for spectral analysis was similar to that of the spectral centroid, in that we extracted the amplitudes of the first eleven harmonics and compared how the spectral energy was distributed between the harmonics. See Figure 2 for an example of a spectrogram, another widely accepted way of depicting of spectral energy distribution.

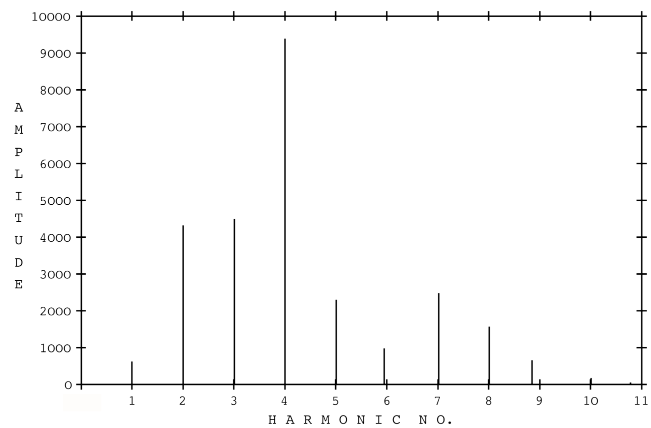
Name	Type	Physical Correlate	Perceptual Correlate	Description
Spectral centroid	Spectral	Energy concentration in low/high spectral area	Brightness/dullness	Balance of energy in spectrum
Irregularity	Spectral	Fluctuating energy between adjacent harmonics	Richness	Amplitude variation of adjacent components.
Roughness	Spectral	Beating of overlapping frequencies	Harshness/smoothness	Inharmonic and noise components in spectrum.
Harmonicity	Spectral	Harmonic/Inharmonic	Cohesive/Diffusive	Ratio of harmonic to inharmonic spectral components.
Attack/Decay times	Temporal	Slope of attack and decay	Instrument identification	Time taken to reach max. amp from 0 (attack).

Table 1: Possible approaches to timbre analysis. Courtesy of a lecture by Deirdre Bolger of CNRS-LMS, Paris, given at Universitt Graz, Austria, 23 November 2005.

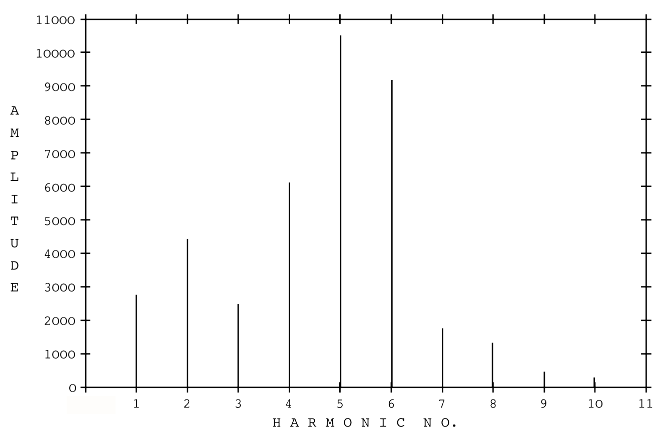
The difference in tone as a result of relative harmonic strength is evident in the construction of large organs, where the choir manual is built such that at least the first eight harmonics can be separately sounded. By combining these harmonics in various ways, different timbres are produced for the same pitch. With the addition of each harmonic above the first, a ‘clearness’ or ‘brilliance’ is added to the sound [14]. (For a thorough review of how relative harmonic strength, or spectral energy distribution, of a musical sound relates to timbre, see Grey 1977.)



(a) Oboe



(b) Trumpet



(c) Violin

Figure 1: Harmonic spectrum of the oboe, trumpet, and a violin sounding A440. The violin's warmth comes from a stronger presence of low harmonics and a more evenly distributed spectrum.

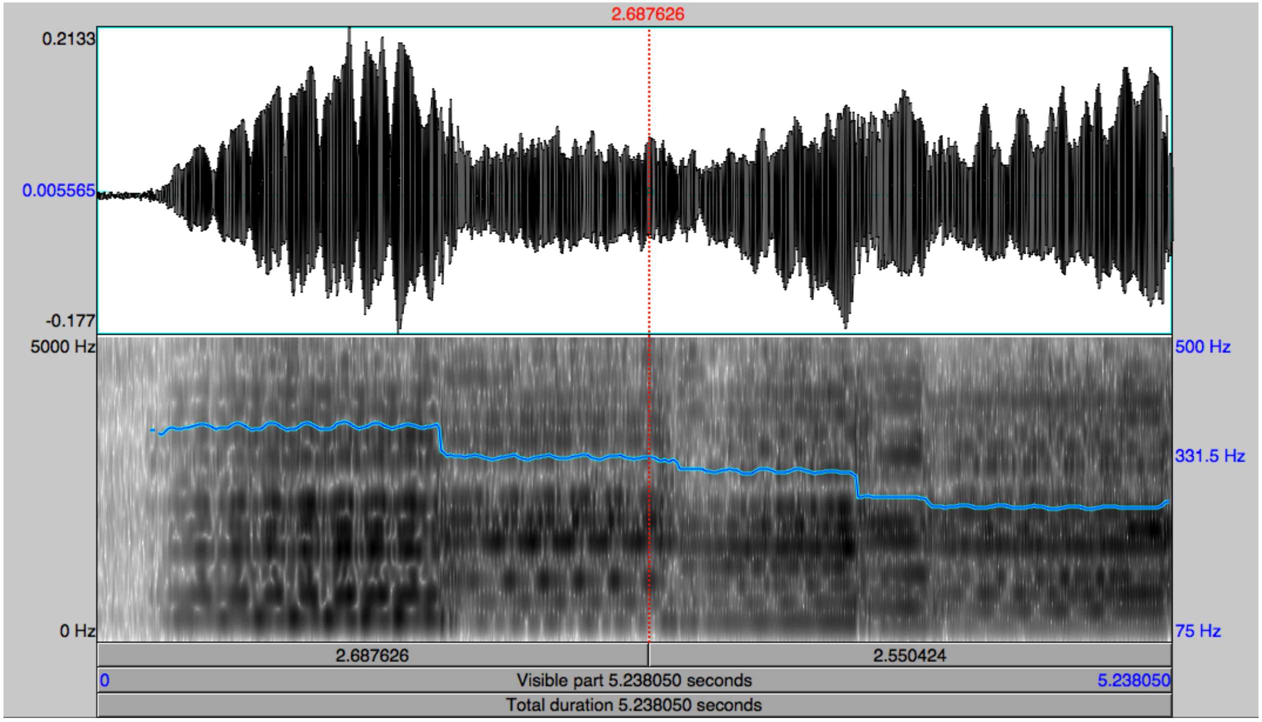


Figure 2: Spectrograms are two-dimensional (time-frequency) representations of audio signals that show how the spectral energy density of the signal varies over time. Dark bands in a spectrogram represent areas of high spectral energy. Depicted are the first five notes of the cello part of Brahms Symphony No. 2 (second movement), taken from a recording of one of the participants in this experiment. The line through the middle of the spectrogram tracks the falling pitch of the scale as the cellist plays. The spectrogram was generated using Praat [18].

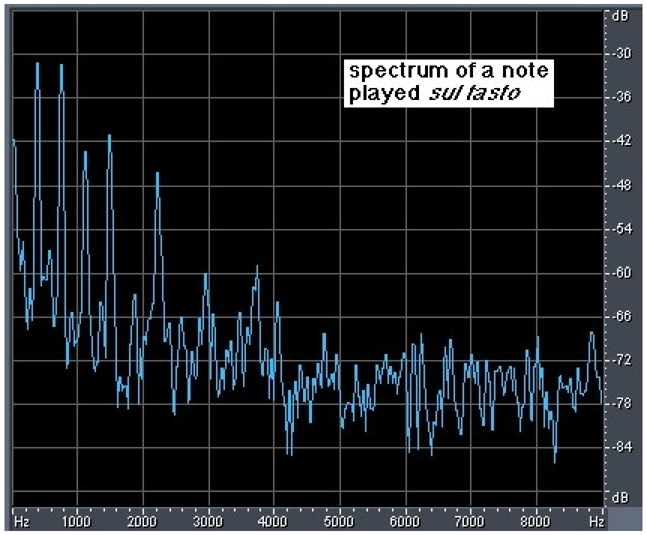
1.3.1 Use of the Bow: How String Instrumentalists Control Timbre

The strings of a string instrument begin the vibration of the entire body, and the frequency of their vibration control the pitch of the emitted tone. With one hand, instrumentalists press the string down at different locations, effectively changing its length. To produce sound, the string is excited with the other hand, either by plucking the string, or drawing the bow across it, setting it into periodic, vibrational motion. The shorter the length of the string, the higher the frequency of its vibration, and thus the higher the emitted pitch.

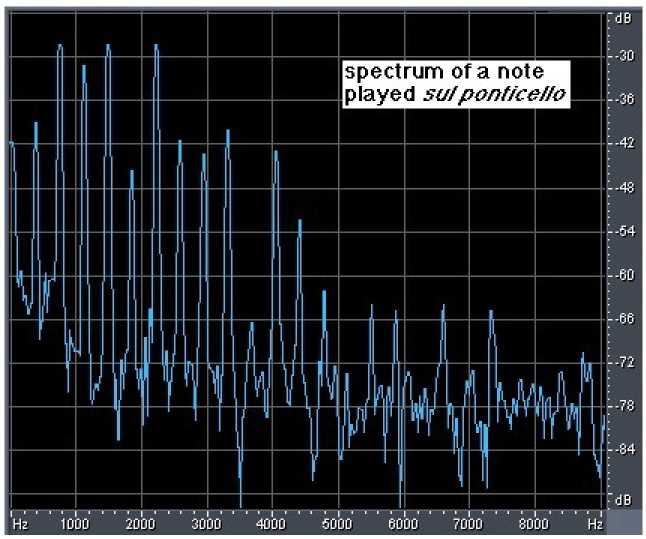
Much of the emitted timbre of the instrument is a function of the specific instrument's design; the body of the instrument has resonances that amplify some frequencies over others. In the case of the violin family, the components which contribute most significantly to these resonances are the top plate (the plate under the strings that the bridge stands on and which has the *f* holes in it), the back plate (the back of the instrument), and the air contained within the main body of the instrument [16]. Particularly, two acoustic resonances, the "air resonance," or the resonance of the air contained within the body of the instrument, and the "top resonance," or the main resonance of the top place, dominate the modification of the spectrum after the vibrations pass through the bridge and to the instrument body. While these parameters are out of the musician's control, he can alter the timbre of the sound through his bow and how he excites the string.

It is widely accepted among string players that the three most common factors that determine tone quality are *bow speed, bow pressure* and *distance from the bridge*. Different musical terms have explicit and implicit instructions regarding how the bow should be used. For instance, *flautando*, or 'flute-like,' implies using a fast bow with little pressure, drawing the bow further away from the bridge, (typically at the base of the fingerboard,) and achieves an airy, light sound. Similarly, *sul tasto*, which literally means 'at the fingerboard' also implies a light sound and little bow pressure. *Sul ponticello*, meaning 'at the bridge,' requires the instrumentalist to play as closely to the bridge as possible (with variable pressure or speed depending on what the music demands). Acoustically, produces a characteristic glassy sound, which emphasizes the higher harmonics at the expense of the fundamental. See Figure 3 for Fourier transforms of *sul tasto* and *sul ponticello* sounds.

This study seeks to determine the effect of the absence of auditory feedback on the timbre of cellists. If indeed a significant effect does exist, it might be explained by alterations in how the bow is used, be it intentional or unintentional.



(a) Sul tasto



(b) Sul ponticello

Figure 3: Spectrum (a) is the Fourier transform of a note played *sul tasto*. Notice in the spectra that the sound is relatively weak in high harmonics. In contrast, spectrum (b) is the transform of a note played *sul ponticello*. Note the weak fundamental and the strong harmonics in the middle and upper frequencies. Our study quantified the differences between similar spectra by grouping harmonic amplitude values (1-3, 4-6, and 7-11) and comparing how those amplitude values changed in the different auditory feedback conditions (see Section 2: Research Method). Spectra courtesy of Joe Wolfe, University of New South Wales [17].

2 Research Method

In the experiment, cellists were recorded performing a scale under two conditions: once with completely unobstructed hearing and once with reduced level of hearing. Hearing was reduced with the aid of ear plugs, noise-canceling headphones, and pink noise streamed through the headphones. Additionally, the cellists were recorded performing a pair of orchestral excerpts under the aforementioned conditions and under a third condition: with hearing masked by a full-orchestra recording of the excerpt.

2.1 Participants

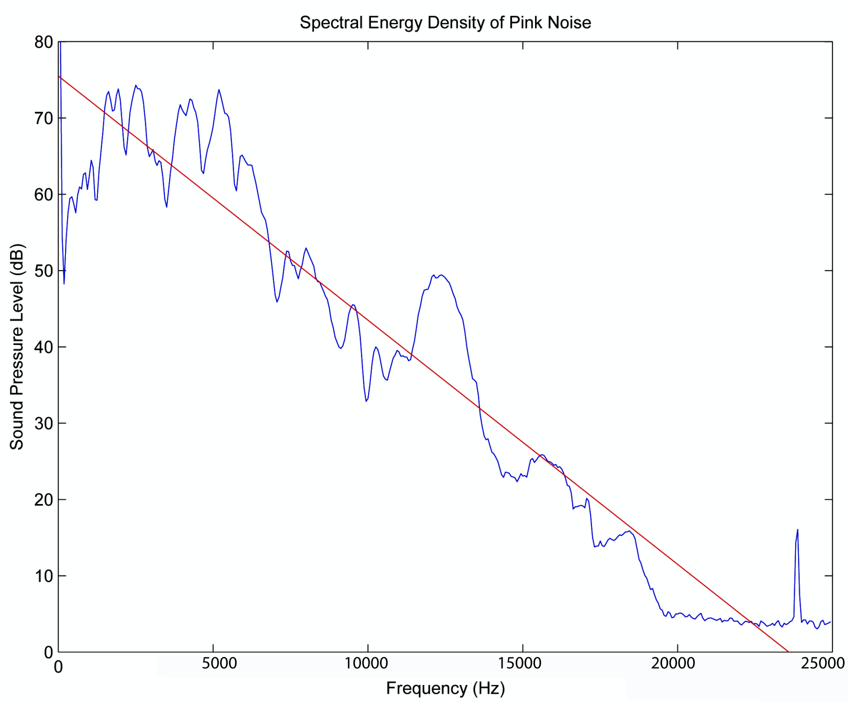
Nine cellists from the University of Michigan School of Music, Theatre, and Dance, agreed to participate in the experiment. They were all cello performance majors, and their ages ranged from 18 to 22 (average of 19.8). The cellists had an average of 7.6 years of orchestral experience (range of 4-10), and practiced for an average of 3.2 hours per day (range of 2.5-4). All were familiar with the orchestral excerpts before the start of the experiment. Seven intended to pursue cello professionally, two were unsure. Given the intensity of their study of cello performance, it would be fair to consider these subjects highly trained musicians.

2.2 Removal of Hearing

Several measures were taken to effectively remove the hearing of the participants. During both the scales and the orchestral excerpts, pink noise was played through the headphones at a maximum of 80 dB. During the excerpts, recordings of the orchestral excerpts were also played at a maximum of 80 dB. Additionally, the participants wore 32 dB earplugs and BOSE QuietComfort II Acoustic Noise Canceling headphones, which create antinoise to prevent the wearer from hearing any steady-amplitude external noise. Pink noise was chosen for its energy distribution of $1/f$, where each octave has an equal amount of power. Consequently, lower frequencies carry more power on average than high frequencies, masking the lower sounds of the cello. The frequency range of the noise extended above 20,000 Hz – the upper limit of the range of human hearing – and thus the entire range of human hearing is masked. The spectrum measurements were made with a Stanford Research Systems SRS760 FFT Spectrum Analyzer, Brel & Kjaer type 4134 microphone with type 2619 preamp, and type 2804 power supply. The microphone coupler was 6cc Type 1. Measurements made by Chris Ellinger of the Kresge Hearing Research Institute at the University of Michigan, Ann Arbor. See Figure 4 for an FFT of the pink noise used for the experiment.

After they completed their recordings, they were asked (in conversation) how effective they thought the noise, headphones, and earplugs were in removing their hearing. There was some variance in how well the participants felt they heard their own playing while pink noise was played through the headphones. Seven participants felt that although they could hear a bit of their playing, the level of feedback reduction was great enough that could not make corrective adjustments to their performance. Two felt they could hear their playing well enough to correct intonation, but not control timbre. The removal of feedback during the portion where orchestral recordings were played was more successful – many of the the participants stated that they could only hear themselves when they were out of tune or out of time with the recording, but otherwise could hear only the recording. Measures were taken to reduce the possibility of intonation and timing errors by using recordings from the performances of the Michigan's University Symphony Orchestra made in September and October, 2010 – the majority of the participants had recently performed the pieces and were playing along with their own performances.

Figure 4: The spectral energy density of the pink noise used to mask aural feedback. This figure plots log(1/f), and so the energy drop per octave is linear on this scale.



2.3 Materials and Apparatus

The acoustic sounds of the instrumentalists were converted to electric signals by a Shure SM81 Instrument Microphone, fed through a TASCAM mkII Audio Interface, and into SNDAN, a frequency analysis UNIX package [19] [20][21].

The Shure SM81 microphone was chosen for its flat frequency response for the entire range of its sensitivity (20-20,000 Hz). Its polar (pickup) sensitivity pattern is a cardioid: it is very sensitive to sound waves that approach the front of the microphone, moderately sensitive to sound waves that approach it from the side (6 dB less than in the front,) and rejects sound that is received from the rear [19]. The TASCAM us-122 mkII interface sends signal from the microphone into the computer at a 96 kHz sampling rate [20]. Given that our microphone has a sensitivity that reaches 20 kHz and our audio interface should sample at 96 kHz, audio quality was lost only when the sound waves were received by the microphone, not when processed through the audio interface and transmitted to the computer. SNDAN is a package of signal analysis, graphics, processing, and synthesis routines for UNIX. One group of programs converts signal files into "analysis files", which contain data for partial amplitudes, frequencies, and phases. Another group is used to view data, either in signal or analysis form. A third group includes programs to modify the analysis data. The fourth type is used to synthesize signal files from analysis files. The software was graciously provided by James W. Beauchamp of the University of Illinois Urbana-Champaign [21].

"Fast-Fourier transforms," or FFTs, are typically employed to reduce the amount of computer time necessary to calculate a discrete Fourier transform for a sampled function (see Cooley and Tukey, 1965). Our software, SNDAN, performs a pitch-synchronous short-time Fourier analysis on a sound file and outputs a harmonic analysis file. Each harmonic is characterized by its time-varying amplitude and frequency over the entire duration of the sound. The program requires that the user input the fundamental analysis frequency for analysis of the sound, but for each note of the scale and each extracted note of the orchestral excerpt experiment, this input process was automated through the use of Python scripts. Programming support was provided by Jay Wren of SRT Solutions (Ann Arbor, MI) and Marshall Weir of Mobiata (Ann Arbor, MI).

2.3.1 Calibration

To ensure our recording equipment and SNDAN analysis software accurately estimated each fundamental frequency and its respective harmonics of our recordings, we generated a succession of frequencies ranging from 70 Hz to 5000 Hz using MATLAB and compared the analysis frequency that SNDAN suggested with

SNDAN Analysis: Expected Values vs. Measured Values		
Expected Frequencies	Standard Dev.	Avg % Deviation
69.30	0.0151	0.0269
123.4	0.0843	0.1127
293.66	0.0079	0.0036
784.00	0.0523	0.0280
2217.46	7.6685	0.514
4698.64	9.5851	0.1994

Table 2: Calibration of recording equipment (microphone and interface) and SNDAN setup. Expected frequencies are measured in Hz.

the MATLAB input frequency. The MATLAB script used chromatic equal tempered frequencies, in which the width of a semitone - the smallest ratio between two successive classical pitches - was equal to the twelfth root of two: $r = \sqrt[12]{2} \approx 1.059463$ [10]. The frequencies used for this script were calculated assuming A440. Seven such wave files were recorded by the Shure SM81 microphone. The microphone output files were analyzed for fundamental frequency accuracy in SNDAN. Six pitches of the 75 equal-tempered recorded pitches were randomly selected for analysis, and the same pitches were analyzed in each of the seven trials. The fundamental frequency suggested by SNDAN was compared to the expected values from the MATLAB script.

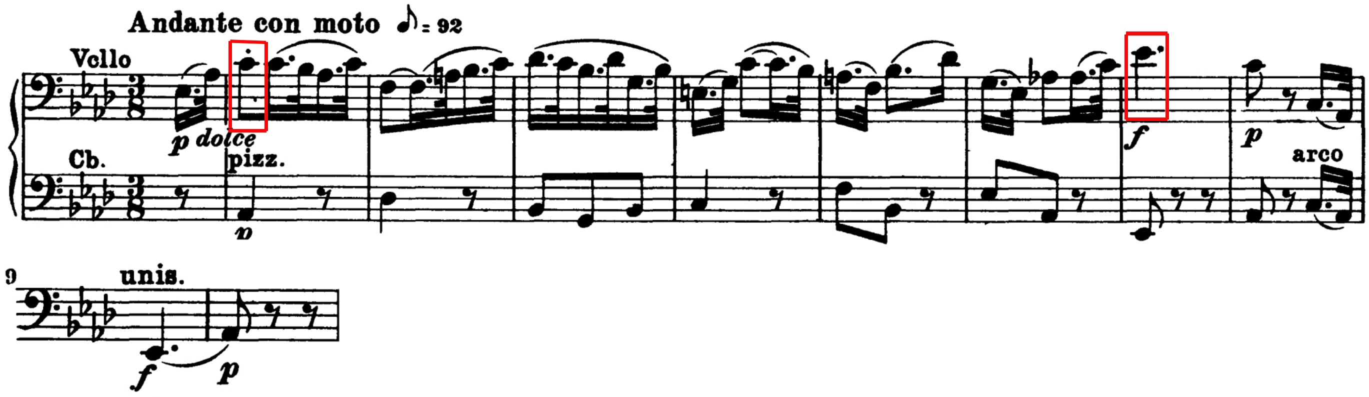
The accuracy of the recording/analysis sequence is presented in Table 2. Although the residuals tend to increase for higher frequencies, the fractional error remains small, with the highest deviation at .91% and a total average deviation of .21%.

Our recording and analysis equipment measures with a .11% error in the range of the fundamental frequencies that we are considering. The error increases to 1% in the range above 2000 Hz, which affects the upper half of the harmonics of the highest pitches of the scale, and only the last two harmonics of the pitches from the excerpts.

2.4 Design and Procedure

The participants each tuned to A440, and were asked to warm up for three minutes with whatever exercises felt most natural. Each participant then performed twelve repetitions of a B major scale from B3 to B4 (six repetitions without headphones, and six repetitions with pink noise streaming through the headphones), twelve repetitions of the first six measures of the second movement, Adagio non troppo, of Brahms' Symphony No. 2 in D major, Op. 73 (four repetitions without headphones, four repetitions with pink noise streaming through the headphones, and four repetitions playing along with the orchestral recording), and twelve repetitions of the first ten measures of the second movement, Andante con moto, of Beethoven's

(a) Brahms Symphony No. 2 in D major, Op. 73, II. Adagio non troppo



(b) Beethoven Symphony No. 5 in c minor, Op. 67, II. Andante con moto

Figure 5: The cello parts of the selected orchestral excerpts.

Symphony No. 5 in c minor, Op. 67 (with the same repetition procedure as the Brahms). The orchestral excerpts are notoriously difficult cello soli and are frequently requested at professional auditions. All cellists were familiar with the music ahead of time - they had practiced the music for their university's orchestral auditions or performances in the past. The orchestral recordings were taken from the University Symphony Orchestra's performances of Brahms Symphony No. 2 (Sept 2010) and Beethoven Symphony No. 5 (Nov 2010). Many of the participants performed in the orchestra when these recordings were made, and thus did not have to adjust their style of play or timing from what they had rehearsed for performances of these works. For the cellists who did not play in these concerts, the excerpts were played twice before their takes were recorded. These orchestral excerpts were chosen for the same reason they are used at professional auditions - the challenge of the excerpts is the creation of beautiful sound quality; thus, it was expected that the participants had previously worked on the excerpts with a focus on timbre.

2.5 Data Analysis

The harmonic data from SNDAN were imported into SPSS for statistical analysis. For each pitch selected, the frequencies and amplitudes of the first eleven harmonics were extracted, as per Shackleton and Carlton

(1994), who found that in the region of 1 kHz, numbered harmonics up to the eleventh convey enough spectral information to create a pitch percept even when they are not separately audible. The harmonic amplitudes and their ratios to the fundamental were individually analyzed, and also separated into groups (harmonics 1-3, 4-6, and 7-11) for further analysis, where the combined strength of each group of harmonics was calculated and compared to the total spectral energy. The harmonics from the first seven pitches of the scale (from B3 - A4) were analyzed, while four select pitches from the excerpt were analyzed: the B natural on beat 3 of the second measure of the Brahms, the E natural on beat 3 of the fifth measure of the Brahms, the C natural on beat one of the first full measure of the Beethoven, and the E flat in the seventh measure of the Beethoven. These notes were selected because SNDAN was able to extract the entirety of a pitch when the pitch rose above or below a certain amplitude threshold, and based on the way cellists typically play these excerpts, phrase breaks and dynamic changes allowed for easy extraction of these four pitches. Due to time constraints, participant 9 only completed the excerpt portion of the experiment. There were 1,104 notes in total that were analyzed, providing 12,144 harmonic amplitude values to work with.

The statistical tests utilized during analysis were the Mann-Whitney U test (for comparing two groups) and the Kruskal-Wallis test (for comparing more than two unpaired groups). The Mann-Whitney U test is used to test the hypothesis that two population distributions are identical. The test statistic is based on the ranks of the observations rather than on their numerical values, and so no assumptions are made regarding the shape of the populations [24].

The logic and procedure of the Kruskal-Wallis test is very similar to that of the Mann-Whitney test, with the main difference being that the Kruskal-Wallis compares three or more groups of data while the Mann-Whitney test is limited to two. Here as well, the null hypothesis is that there is no significant difference among the groups in the ranking scores. The test statistic H evaluates the discrepancy of the ranking scores among the three groups. The larger the H score, the larger the discrepancy, making it more likely that the three groups are difference, resulting in a smaller p-value [25].

3 Results

3.1 Experiment 1: Scale Analysis

The notes of the scale were analyzed individually with respect to the weight of each group of harmonics compared to the total energy of the combined harmonics. The average energy distribution of each group of harmonics is summarized in Figures 6, 7, and 8. Figures 6 and 7 shows an analysis with combined data from all participants. Figure 8 depicts the changes in spectral energy distribution with respect to each individual player.

Possible correlations in the data of the scale experiment include timbre change tendencies within each player, the tendencies of the players as a collective, and consistencies in timbre changes from note to note of the scales. A Mann-Whitney test was conducted to evaluate differences between the two conditions – with and without feedback – across all players and all notes. The test, which was corrected for tied ranks, was highly significant in the case of the low harmonics (ρ =.001), significant for the case of the middle harmonics (ρ = .035), and non significant in the case of the high harmonics (ρ = .056). However, this test was run simultaneously on all participants and all notes, and follow-up tests that analyzed individual players and individual notes returned largely non-significant results.

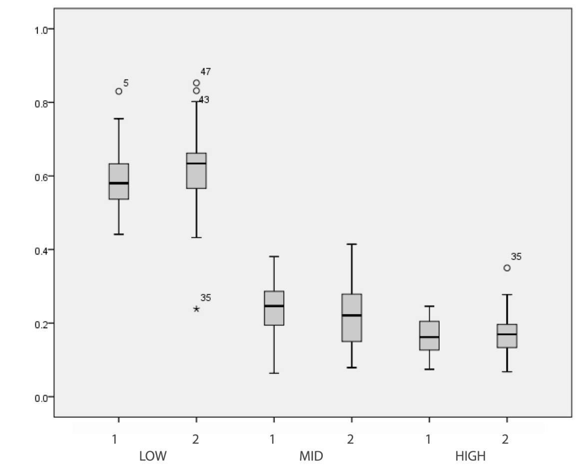
Examining Participant 1 more closely and separating the analysis to one note at a time, we find that a Kruskal-Wallis test across three harmonic groupings for each note (21 comparisons in total) yielded only one instance of significance, where the middle harmonics in the fourth note of the scale (E4) proved to differ significantly between conditions (U = 4, Z = -2.242, ρ = .025). Comparatively, combining the notes of the scale under one analysis yielded ρ values of $<.001$ in each of the three harmonic groupings. Considering that it is standard performance practice to try and maintain an even tone from one note to the next, and since no audible inconsistencies or large dynamic fluctuations were made by the performers during the scales, it is likely that harmonic changes between notes can be attributed to the varying resonances of the cello, where the resonances that amplify one group of harmonics over another depending on what note is played. From this, for the rest of the analysis, only notes of the same pitch were compared, and no groupings were made across notes.

Each of the participants' data were analyzed with the same procedure as participant 1. With 8 participants, 3 energy groupings per participant, and 7 notes to analyze, there were 168 comparisons in total. By and large, the different notes of the scales and the harmonic groupings did not show substantial changes between conditions. In 15 comparisons, the low harmonics varied significantly between conditions; in 13 comparisons, the middle harmonics; and in 11 comparisons, the high harmonics. The variability across notes also showed no recognizable pattern, and no particular note had more than 10 cases of significant change between conditions.

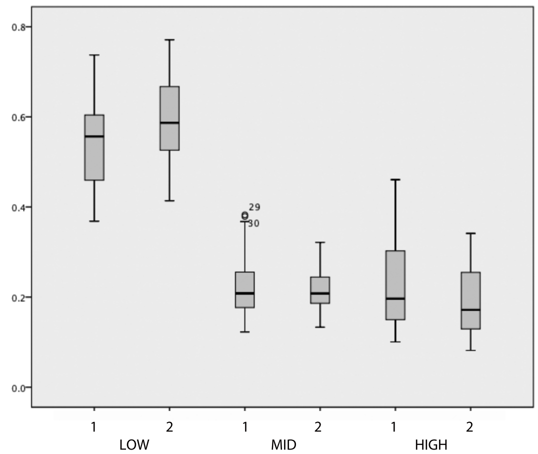
Regarding the individual players' tendencies, participant 3 and participant 1 showed hardly any tendency for change between conditions - they showed no measurable change in their sound for all but one of group or harmonics in one note. Participant 2 was similar, only showing changes in the middle harmonics in note 2

($\rho = .004$) and in the low harmonics of note 7 (ρ =.037). The rest of the participants, although they showed a starker contrast between conditions, played with no significant change in their sound for over half of the experiment. (Out of 24 comparisons for each participant, participants 4 and 5 each had 5 comparisons of significance, participant 6 had 8, participant 7 had 10, and participant 8 had 9.)

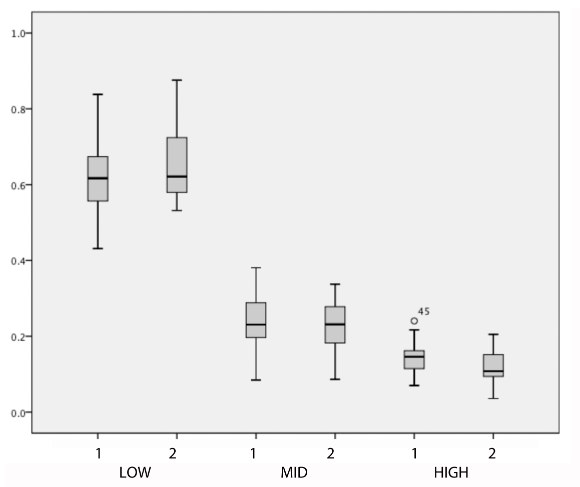
Figure 8 illustrates how each player's harmonic energy tended to shift between feedback conditions. The figure depicts a relatively consistent pattern: for 12 out of 15 cases where there was a measurable change in the low harmonics, spectral energy concentration was higher during the condition where feedback was removed. In 11 out of 13 cases for the middle harmonics, spectral energy concentration decreased. The high harmonics did not show as consistent of a pattern – in 7 out of 11 cases, spectral energy concentration decreased in the high harmonics.



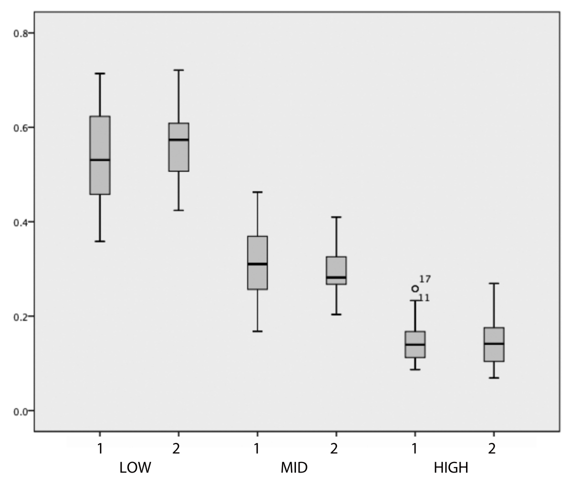
(a) Note 1



(b) Note 2

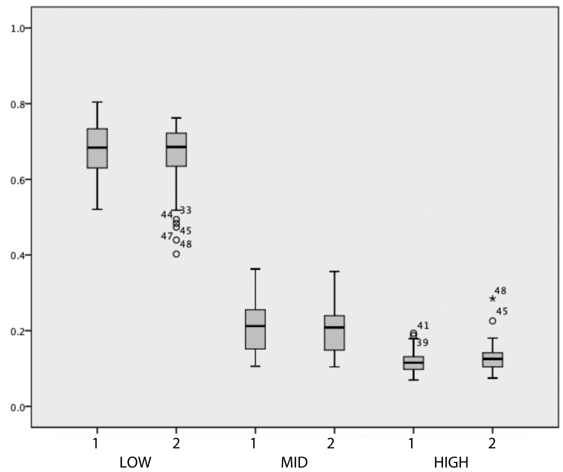


(c) Note 3

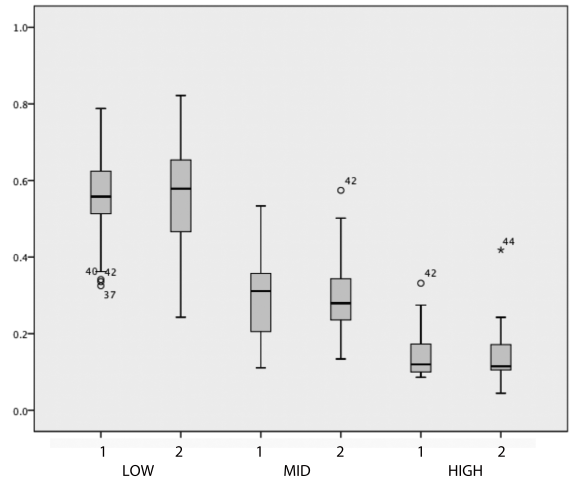


(d) Note 4

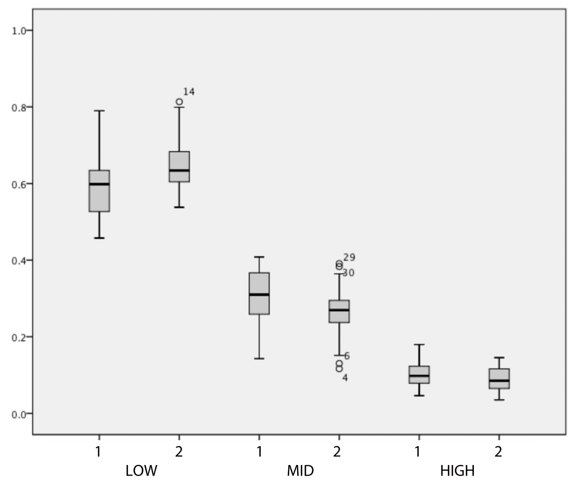
Figure 6: Boxplots showing the distribution of energy concentrations in each note of the scale, across all participants. The x-axis organizes the energy distributions into conditions: LOW = Harmonics 1-3, MID = Harmonics 4-6, HIGH = Harmonics 7-11, and into conditions: 1 = no noise, 2 = noise. Each box depicts a five-number summary: the box itself spans from the 25th to the 75th percentile, the lower whisker extends to the 25th percentile minus 1.5*IQR (IQR=interquartile range, the difference between the 75th and 25th percentiles of the data), and the upper whisker extends to the 75th percentile plus 1.5*IQR. Outliers are also depicted. Continued on page 21.



(a) Note 5

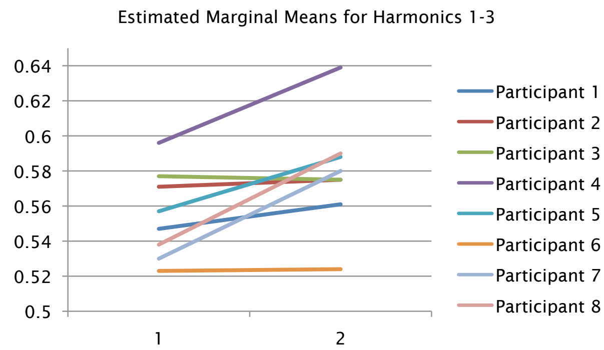


(b) Note 6

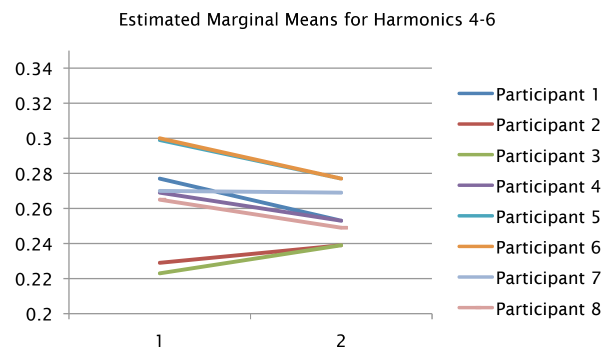


(c) Note 7

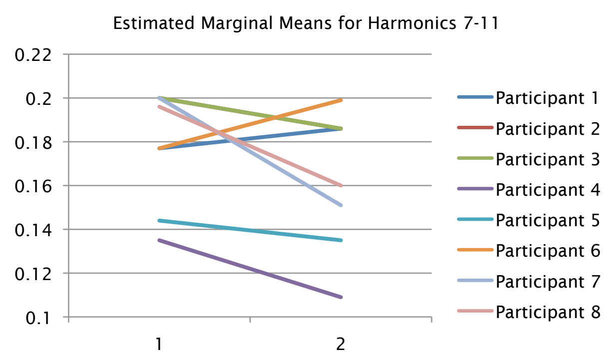
Figure 7: Continuation of boxplots showing the distribution of energy concentrations in each note of the scale, across all participants. Notes 5-7 depicted.



(a)



(b)



(c)

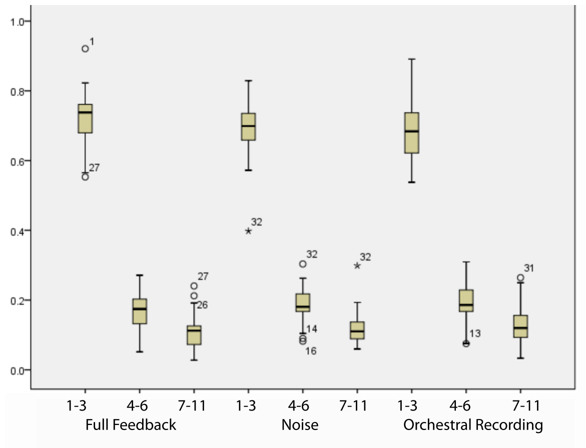
Figure 8: The energy distribution changes in groups of harmonics between conditions. The x-axis conditions: 1 - noise off, 2 - noise on. The total energy of each sound file was normalized to 1; the y-axis is the energy distribution such that the total energy, across all three subfigures, equals 1. All notes of the scale were combined in this analysis.

3.2 Experiment 2: Excerpt Analysis

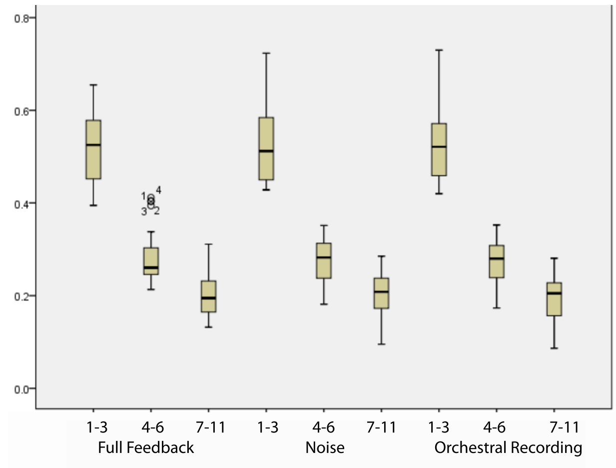
For the excerpt experiment, a Kruskal-Wallis test compared effects of the feedback conditions within each participant for each note as well as across all participants for each note. Henceforth, condition 1 refers to the condition where feedback is present, condition 2 refers to the condition with pink noise, and condition 3 refers to the condition where the orchestral recording is played. With a total of 108 comparisons (9 participants, 4 notes each, 3 harmonic groupings), 31 showed a significant difference between the three feedback conditions. Although this number may seem high, correlations between spectral energy distribution and feedback condition are not readily apparent. In Table 3, the test statistics of a Kruskal-Wallis test show that no one feedback condition has a significantly higher or lower concentration of spectral energy than the others. Figure 8 shows a representation of the energy concentration in each note of the excerpts, and Figure 9 represents the tendencies of the harmonic energy distribution in each participant across all notes of the excerpts.

With regard to individual players, participant 4 played the most consistently between conditions, with only one comparison of significance, during note 1 of the Beethoven, in the middle harmonics ($H = 6.269$, $\rho =.04$). Participant 2 and participant 5 also showed minimal change between conditions, each having only two comparisons of significant change, both during the first analyzed note of the Brahms excerpt – participant 2: low harmonics ($\rho = .025$) and mid harmonics ($\rho = .02$), participant 5 low harmonics ($\rho = .004$) and high harmonics ($\rho = .044$). Regarding the playing of the rest of the participants, there was no tendency in any participant for the spectral energy distribution to shift from one set of harmonics during another when changing between conditions.

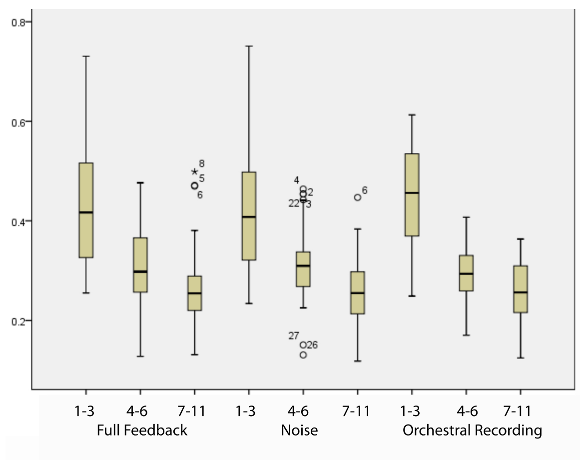
There was a surprising distribution of these significant comparisons among the four notes tested. The first note of the Brahms contained 18 out of the 31 comparisons, whereas only 6 were during the second note of the Brahms, 4 were during the second note of the Beethoven, and 3 during the first note of the Beethoven. However, there was no consistency as to which condition contained the most spectral energy in each of the individual notes.



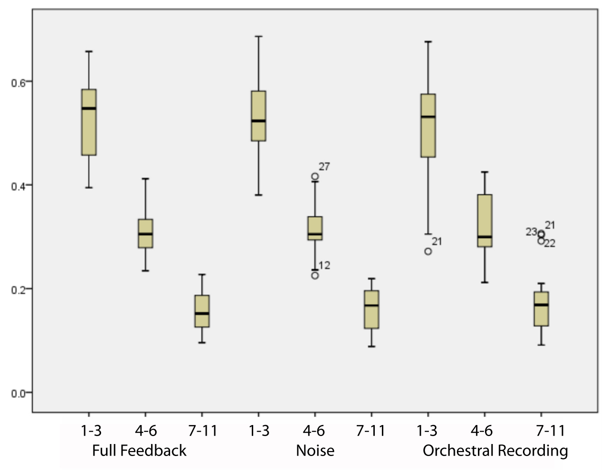
(a) Note 1 from Beethoven



(b) Note 2 from Beethoven



(c) Note 1 from Brahms



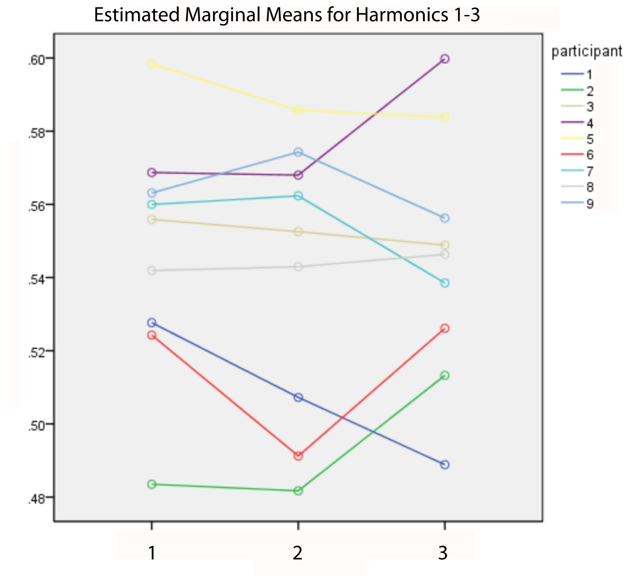
(d) Note 2 from Brahms

Figure 9: Boxplots representing the distribution of energy concentrations in each note of the excerpts. The x-axis organizes the energy distributions into harmonic number groupings and conditions. The y-axis is given as a ratio of energy to the whole. All participants are represented in this analysis.

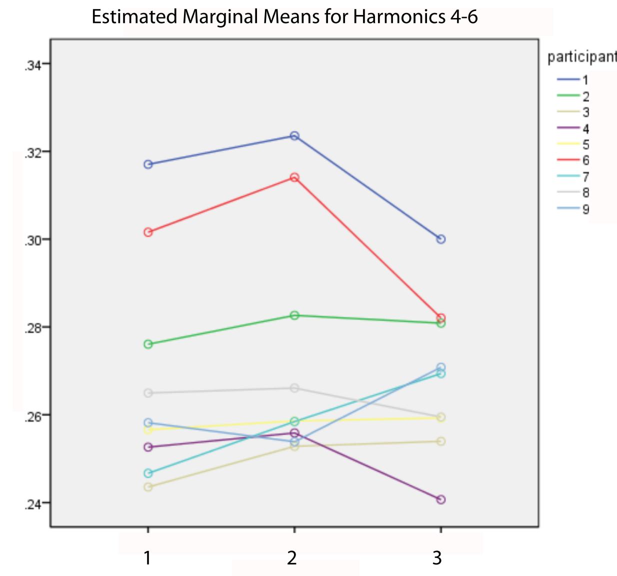
Table 3: Kruskal-Wallis Test Mean Ranks for Extracted Notes from Excerpts

Note number:	1			2			3			4		
Harmonics:	1-3	4-6	7-11	1-3	4-6	7-11	1-3	4-6	7-11	1-3	4-6	7-11
With feedback	63.0	48.3	47.2	53.0	53.0	54.8	51.7	52.6	53.7	55.5	52.0	52.6
Pink noise	53.6	56.6	55.1	53.2	56.0	56.6	50.5	58.4	53.6	53.8	54.6	53.0
Orchestra recording	47.0	58.5	61.2	57.3	54.4	52.1	60.0	51.0	54.6	51.3	53.9	55.0
Chi-square	4.760	2.172	3.615	.443	.165	.370	1.955	1.116	.017	.328	.135	.108
Asymp. Sig	.093	.338	.164	.801	.921	.831	.376	.572	.992	.849	.935	.947

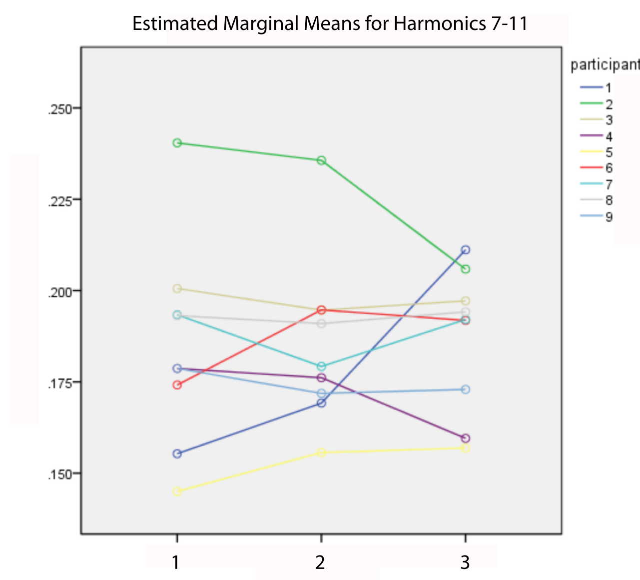
Note 1: first extracted note from Beethoven excerpt.
Note 2: second extracted note from Beethoven excerpt.
Note 3: first extracted note from Brahms excerpt.
Note 4: second extracted note from Brahms excerpt.



(a)



(b)



(c)

Figure 10: The energy distribution changes in groups of harmonics between conditions. The x-axis correlates to the three feedback conditions: condition 1, the full feedback condition; condition 2, the headphone/pink noise condition; and condition 3, the orchestral recording condition. The total energy of each sound file was normalized to 1; the y-axis is the energy distribution such that the total energy, across all three subfigures, equals 1. All analyzed notes for the excerpt were combined in this analysis.

4 General Discussion and Conclusions

This experiment represents a first attempt to assess the effect that feedback alteration has on a musicians' sound quality, specifically harmonic spectral energy distribution, in cellists. The data provides support for the hypothesis that feedback alteration has minimal effect on timbre, and complements the results of the previous literature, whereby removing feedback altogether has little effect on the player's ability to play consistently. This hypothesis is more strongly supported by the excerpt experiment, where a few players (participants 2, 4, and 5), showed very little change between conditions. The scale experiment supports the hypothesis as well; however, if the player *is* affected by feedback condition, it is likely that the spectral energy concentration will be higher in the lower harmonics when feedback is removed. This slight increase in spectral energy concentration in the lower harmonics in some cases during the feedback removal condition might be attributed to bowing placement, where the players, by playing closer to the fingerboard and thus further away from the place where the string is anchored by the bridge, are exciting fewer higher harmonics. A follow-up test regarding bowing placement's effects on the spectral energy distribution, in which a cellist (the PI) played 40 repetitions of G3 at distances 1, 2, 3, and 4 inches from the bridge (10 repetitions each). The spectral centroid averages of the trials were 585, 542, 443, and 434 Hz, respectively, and differed significantly ($\rho = .005$), showing a strong correlation between bowing placement and spectral energy distribution.

From the cellists' perspective, it is reassuring to know that the removal of one's ability to hear oneself does not imply large changes in sound quality. A lot of stock is placed in an orchestral musicians' ability to produce a controlled timbre. Naturally, there are other factors that musicians need to consider regarding consistent playing – rhythm and pitch accuracy being the most salient, but experimental data showing the attainability of consistency in playing regardless of feedback condition is a testament to the high control the musicians have over their own sound quality.

However, included in the survey distributed to the participants before the start of the experiment was the question: "Do you think solo playing requires a different quality of sound than orchestral playing?" All participants answered "yes." Considering the lack of variation in sound quality between the feedback conditions in the excerpts was so minimal, this unanimous response begs several questions. Were they playing with their desired orchestral sound, with the intention of blending with the recording, during the orchestral excerpt experiment? If they were in fact playing with their envisioned kind of sound quality and making a substantial difference in sound quality to their ears, are our data analysis methods sensitive enough to detect the changes in their sound? Musicians are frequently credited with having an unusually keen sense of hearing – the ability to discriminate changes in pitch is less than 1/250 semitone was found amongst professional musicians from the Royal Opera in Vienna [26]. Additionally, research has shown that when modifying the acoustical properties of a violin, perceptual thresholds are significantly lower in musically trained participants than for non-trained participants [27].

How can these questions be addressed? Future experiments might use alternate data collection techniques that record the actions of the bow of the player in more detail to pinpoint which factor of bow technique is primarily responsible for the timbre changes (if any). The question of the intention of the player during the style of music selected for the experiment could have an effect on the sound quality. If the only musical examples used are orchestral excerpts, the players might be aiming to play, throughout the entirety of the experiment, with an 'orchestral sound', yielding little difference between playing it with full feedback and along with an orchestral recording. Future experiments might consider using different selections for different parts of the experiments – a solo piece to be used during the full feedback condition and orchestral excerpts to be used during the reduced feedback condition – so that the musicians approach the experiment fully cognizant of the comparisons between styles of play. Furthermore, experiments might outline instructions for the participants in greater specificity, so that, if the aim of the experiment is to account for the adjustments players make while blending their sounds – as opposed to an account of feedback removal alone – they are playing with the intention of blending with the orchestral recording.

This experiment raises interesting questions in the Western orchestral performance tradition that warrant further study from both the physical and musicological perspectives. The results of the present study support the hypothesis that musicians' timbre is largely unaffected by aural feedback conditions. The tendency to play such that there is less harmonic spectral energy content when feedback is removed was present during the scale experiment but not the excerpt experiment; however, refining the design of the orchestral excerpt experiment could elucidate findings not evident from this study.

Possible future experiments include an exploration of which bow techniques orchestral players utilize most frequently to change their timbre, and with the addition of piezoelectric sensors, the physical details of the player's bow action – pressure, velocity, and distance from the bridge – during different feedback conditions. Additionally, timbre is very multidimensional, and the tests that explore its other components (outlined in Section 1.3,) including irregularity (fluctuating energy between adjacent partials), roughness (beating of overlapping partials), harmonicity (the ratio between harmonic and inharmonic sounds), as well as attack/decay times, can provide a comprehensive account of timbre changes.

This study does not deny that normal auditory feedback is unnecessary to refine the quality of sound during practice. Indeed, aural feedback during all stages of learning a piece is important. However, once the process of becoming technically proficient on one's instrument and knowing the music that one is performing is complete, timbre production can move largely in a feed-forward manner, as evident in the results of this study.

References

[1] Pfordresher, Peter Q., and Caroline Palmer (2006). *Effects of hearing the past, present, or future during music performance.* Perception and Psychophysics. 68.3: 362–376.

[2] Yates, Aubrey I. (1963). *Delayed Auditory Feedback.* Psychological Bulletin 60.3: 213-232.

[3] Finney, Steven A. (1997). *Auditory Feedback and Musical Keyboard Performance.* Music Perception. 15.2: 153–174.

[4] Finney, Steven A., and Caroline Palmer (2003). *Auditory feedback and memory for music performance: Sound evidence for an encoding effect.* Memory & Cognition. 31.1: 51–64.

[5] Gates, A., and Bradshaw, J. L. (1974). *Effects of auditory feedback on a musical performance task.* Perception and Psychophysics. 15: 191–200.

[6] Pfordresher, P.Q. (2005). *Auditory feedback in music performance: The role of melodic structure and musical skill.* Journal of Experimental Psychology: Human Perception and Performance. 31: 1331–1345.

[7] Repp, B.H. (1999). *Effects of auditory feedback deprivation on expressive piano performance.* Music Perception. 16: 409–438.

[8] Repp, B. H., & Knoblich, G. (2004). *Perceiving action identity: How pianists recognize their own performances.* Psychological Science. 15: 604–609.

[9] Helmholtz, Hermann (1885). *On the Sensations of Tone.* New York: Dover Publications, Inc. 1954. 10–32.

[10] . Benson, Dave. *Music: a Mathematical Offering.* Cambridge University Press: Cambridge, UK. 2008.

[11] Benson, Donald C. (2003). *A Smoother Pebble: Mathematical Explorations* (pp 56). New York: Oxford University Press.

[12] Schouten, J.F. (1968). *The perception of timbre.* Reports of the 6th International Congress on Acoustics.

[13] Bolger, Deirdre. *An Exploration of timbre: its perception, analysis and representation.* Institut fr Musikwissenschaft, Universitt Graz. Graz, Austria, 23 November 2005.

[14] Jeans, James (1937) *Science & Music.* New York: Dover Publications, Inc. 1968.

[15] Grey, J.M. (1977). *Multidimensional perceptual scaling of musical timbres.* Journal of the Acoustical Society of America 54: 1496–1516.

[16] Howard, David M., and Jamie Angus. *Acoustics and Psychoacoustics. Amsterdam.* Focal, 2009. 170-181.

[17] Wolfe, Joe. *Articulation and Vibrato on the Violin.* School of Physics at UNSW, Sydney, Australia. Web. 14 Nov. 2011. http://www.phys.unsw.edu.au/.

[18] Boersma, P., & Weenink, D. (2005). Praat: doing phonetics by computer (Version 4.3.01) [Computer program]. Retrieved from http://www.praat.org/.

[19] SHURE Global. *SM57 Instrument Microphone*. http://www.shure.com/proaudio 03.07.2010.

[20] TASCAM; Teac Professional. *US-122mkII* http://www.tascam.com/ 03.07.201.

[21] Beauchamp, James W. (2007). "Analysis and Synthesis of Musical Instrument Sounds", in Analysis, Synthesis, and Perception of Musical Sounds: Sound of Music, J. W. Beauchamp, Ed., Springer, N. Y., pp. 1 – 89.

[22] Cooley, James W. & Tukey, John W. (1965). *An algorithm for the machine calculation of complex Fourier series*. Math. Comp. 19 (90): 297-301.

[23] Shackleton, Trevor M. and Robert P. Carlyon (1994). *The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination*. Journal of the Acoustical Society of America. 95.6: 3529–3540.

[24] Kirk, Roger E. *Statistics: an Introduction*. Belmont: Thomson/Wadsworth, 2008. 502–505.

[25] Vaughan, Liwen. *Statistical Methods for the Information Professional: a Practical, Painless Approach to Understanding, Using, and Interpreting Statistics.* Medford, NJ: Information Today, 2001. 149– 151.

[26] Seashore, Carl E. *Psychology of music*. Dover Publications: New York, 1967.

[27] Fritz, Claudia; Cross, Ian; Moore, Brian, C.J.; & Woodhouse, Jim (2007). *Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin*. Journal of the Acoustical Society of America. 122.6: 3640 – 3650.