2012

Direction of mistuning, magnitude of cent deviation, and timbre as factors in musicians' pitch discrimination in simultaneous and sequential listening conditions

Norman Alan Clark
Louisiana State University and Agricultural and Mechanical College

DIRECTION OF MISTUNING, MAGNITUDE OF CENT DEVIATION, AND TIMBRE AS FACTORS IN MUSICIANS' PITCH DISCRIMINATION IN SIMULTANEOUS AND SEQUENTIAL LISTENING CONDITIONS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
In partial fulfillment of the
Requirements for the degree of
Doctor of Philosophy

In

The College of Music and Dramatic Arts

by
Norman Alan Clark
B.M.E., University of Florida, 1976
M.F.A., University of Florida, 1981
M.S.I.R., Troy State University, 2004
August 2012

ACKNOWLEDGEMENTS

There are many people without whom this document would have never existed. I owe them all a tremendous thank you. First of all my committee; My Major Professor Dr. James Byo for teaching me how to think, how to edit, and how to persist; My Minor Professor Dr. Steven Beck for greatly improving my counterpoint, teaching me how to use ChucK and introducing me to the wonderful world of Psychoacoustics; Dr Jane Cassidy for many hours of patient tutoring in statistics, her wealth of knowledge about intonation, and her always positive encouragement; Dr. Evelyn Orman for teaching me to keep my facts straight and for opening my eyes to the teaching potential in music technology and the power of music sequencers and digital editing software. I would also like to thank Dr. Linda Moorhouse for her friendship, advice, and encouragement; Dr. Steven Morrison for asking me just the right "sharp questions" at just the right time; Mr. Roy King and Mr. Eric Melley for encouraging members of their college ensembles to participate in this study and for helping me keep a sense of humor.

Numerous graduate students also contributed greatly to the cause; Mr. Jeff Albert, Mr. Cory Knoll, and Miss Lindsey Jacob for helping me learn to use ChucK; Dr. Elizabeth Menard, Dr. Amanda Schlegel, Mrs. Rebecca Carter Belello, and Mr. Neil Nelson for being fantastic office mates, friends, and guinea pigs; Miss Katy Strickland for being a superb pilot testing volunteer and a great friend. Several band directors were fabulous as well; Mr. Eddie Hirst and Miss Kristie Smith for their unselfish help by making their high school and middle school students available for pilot testing; Miss Patti Roussel for graciously allowing me to take her students out of class on many occasions for the good of advancing the profession.

And finally thank you to my family; Joseph and Jeffrey for being patient and supportive when dad wasn’t always there, and Colleen for loving me, believing in me, and straightening me out when I needed it.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS..ii

LIST OF TABLES...vi

LIST OF FIGURES……………………………………………………………………..… vii

ABSTRACT………………………………………………………………………………viii

CHAPTER

1 INTRODUCTION AND STATEMENT OF THE PROBLEM……………….. 1
Introduction………………………………………………………………… 1
Statement of the Problem……………………………………………….7

2 REVIEW OF LITERATURE………………………………….……………. 10
Intensity of Sound……………………………………...………………. 12
Timbre……………………………………………………...……………. 15
Temperament ……………………………………………....…….……. 26
Age and Musical Experience………………………………………..…. 29
Direction of Approach……….………………….………………………... 34
Pitch Discrimination……….…………….………………………………. 37
Simultaneous and Sequential Presentation…………………………….. 44
Precursors……………………………………………………………….44
Purpose and Research Questions ……………………………………… 49

3 METHODOLOGY AND PROCEDURES ……………………………………51
Participants 51
Variables…………………………………………………………….. 52
Development of Reference and Test Tones…………………………….58
Development of Pitch Discrimination Test……………………………..59
Pilot Testing……………………………………………………………64
Procedures ………………………………………………………………...65
Reliability ……………………………………………………………….66

4 RESULTS……………………………………………………………………. 67

5 DISCUSSION………………………………………………………….......76
Response to Pitch ………………………………………….…...………76
Response to Timbre…………………………………………………… 77
Response to Mode of Presentation...78
Response to Cent Deviation..80
Interactions 84
Age and Experience.. 88

Fatigue... 89
Conclusions.. 90
Future Research... 93

REFERENCES..96

APPENDICES
A IRB Request for Exemption..104
B School Administrator Consent Letter..................................... 105
C Teacher Consent Form...106
D Parental Consent Form..108
E Adult Participant Consent Form.. 110
F Participant Data Form.. 112
G Script for Pitch Discrimination Testing Session 113
H Pitch Matching Test Answer Sheet...115

VITA.. 119

LIST OF TABLES

3.1 Demographics 52

3.2 Music Instruction (in Years) – High School Participants 53

3.3 Music Instruction (in Years) – University Participants 54

3.4 Cent Deviation Frequencies 55

3.5 Table of Random Orders for Test Tones 61

3.6 Test Administration Order Matrix 63

4.1 Four-Way ANOVA with Repeated Measures for Pitch, Timbre, Presentation, and Cent Deviation 68

4.2 Scheffé Post-Hoc Analysis of the Effect of Cent Deviation (Probabilities) 70

4.3 Levels of Mistuning: Percent Correct and Ranges 74

4.4 Pre-break versus Post-break Percent Correct 75

LIST OF FIGURES

4.1 Mean Scores for Cent Deviation....69

4.2 Interaction Between Pitch and Cent Deviation....71

4.3 Interaction Between Timbre and Cent Deviation.... 71

4.4 Interaction Between B-flat and E Reference Pitches with Same and Different Timbres....72

4.5 Interaction Among B-flat Pitch and Timbre and Presentation and Cent Deviation.... 73

4.6 Interaction Among E Pitch and Timbre and Presentation and Cent Deviation.... 73

ABSTRACT

The main purpose of this study was to investigate high school and college wind instrumentalists' pitch discrimination when judging pitch pairs separated by 0, 5, 7.5, and 10 cents. Participants listened via headphones to a pre-recorded two section perception test; each section (one sequential and one simultaneous) containing 56 tone pairs. Each pair consisted of an in-tune reference tone followed by a test tone of the same pitch (B-flat$_4$ or E$_4$), which was either identical in tuning or altered to one of six mistunings. Tones also varied in timbre (square or sawtooth wave) with the reference and test tones being either the same or different in timbre. Participants circled on an answer sheet whether test tones were lower, the same, or higher than their paired reference tones.

The main effects of pitch, timbre, presentation order, and cent deviation were significant ($p < .05$). Participants were significantly more accurate identifying mistunings at the 10 and 7.5 cent levels than at the 5 and 0 cent levels. Responses were least accurate when stimuli were in-tune. Different timbre pairs resulted in more correct responses than same timbre pairs and participants correctly identified the tunings and mistunings for the B-flat pitch pairs significantly more often than the E pitch pairs. Simultaneously presented pairs resulted in more accurate responses than sequentially presented pairs. University students responded more accurately than high school students at all levels of mistuning.

In the timbre and cent deviation interaction, the different timbre pairs were correctly identified at a higher rate than were the same timbre pairs, except at 0 cent deviation where the reverse occurred. The pitch by cent deviation interaction produced the largest effect size of all (partial $\eta^2 = .66$). Participants responded more accurately to E when it was flat than when it was sharp and more accurately to B-flat when it was sharp than when it was flat, a finding that is

inconsistent with listeners' general tendency to discriminate flat better than sharp in previous research.

CHAPTER 1
INTRODUCTION AND STATEMENT OF THE PROBLEM

Introduction

Intonation is a fundamental element that contributes to quality in music performance. One's ability to hear, judge, and adjust pitch is widely regarded as a major criterion of skillful musicianship. Evidence of its position as both fundamental element of and roadblock to quality music making is found in Battisti's words as he spoke about the wind band; "I know of no other organization that buys so much tuning equipment, and spends so much time on tuning, yet plays so out of tune" **(**Battisti, 1998, p. 4).

Agreement is lacking among music practitioners as to how to approach tuning in the large ensemble rehearsal. That fact alone indicates that this is an area that is in need of further study. Two approaches dominate in band settings. In one, the conductor/teacher uses an electronic tuner to determine whether the musicians are 'in-tune' by either showing them the readout on the tuner while they play a predetermined pitch, or by looking at the readout him/herself and instructing students how to adjust their tuning apparatus. Many conductors and players approach tuning and intonation from this surface level perspective only. It however does not challenge the musicians aurally.

Another approach to tuning depends on performers and teachers making aural judgments. In this performer-centered setting the musicians are called on to decide when they are in-tune based on a purely aural process; essentially when it sounds right it is 'in-tune.' Two tones are judged to be in-tune when they fall within some aurally determined acceptable range, even though they may actually be one or more cents apart when measured electronically. The

conductor/teacher who guides students using this approach to tuning exposes them to different considerations as compared to the visual output of an electronic device. When the ear decides, rather than the tuner deciding, it faces the whole of music—pitch, timbre, tessitura, presentation mode, instrumentation, dynamics, rhythm, etc.

In musicians' vernacular, intonation and tuning are often used interchangeably, but *The New Grove Dictionary of Music and Musicians* differentiates between the two by defining intonation as the "inflections of pitch that form an inherent part of the performance itself," and tuning as "the adjustment, generally made before a musical performance, of the intervals or the overall pitch level of an instrument" (Lindley, 2001c, p. 884). Kennedy and Kennedy (2007) blur the distinction by defining intonation as the "act of singing or playing in-tune" (p. 373).

Whether intonation or tuning, each can be conceived aurally. Nichols (1947) posited that hearing and listening are separate yet related elements of a larger process, with hearing dealing only with the perception of sound, and listening defined as "the attachment of meaning to aural symbols" (p. 83). Hearing as a physical process is defined by Wagner (2009) as a sensory process whereby the ears are excited into vibration by sound waves, and these vibrations are translated by the cochlea, the auditory nerve, and the brain into recognizable sounds.

The lowest pitch distinguishable by the average listener was reported by Seashore (1938) to be approximately 16 Hertz (Hz) (cycles per second), and Olson (1967) identified $C_0 = 16.35$ Hz as the lowest musical pitch. Olson also designated the upper limit for pitch recognition to be around C_{10} which equals 16.74 kilohertz (kHz) or 16,744 Hz. Wagner (2009) designated the upper threshold for human hearing to be approximately 20 kHz for younger subjects. These two different limits may not be contradictory if the upper limit of human hearing is higher than the upper limit for pitch recognition as defined by Olson. Neither source clearly differentiates

between the ability to hear a high frequency sound and the ability to identify it as a specific pitch. The Western musical scale is based on 12 half steps per octave and a span of ten octaves between C_0 and C_{10}. The ranges of the wind instruments used in the modern band fall well within the piano's range of A_0 to C_8. Within this range the normal ear recognizes 88 discrete pitches, each separated by a half step. Olson (1967) stated that the physiological sensitivity of the average ear is far more capable and can detect differences between as many as 1,400 specific frequencies within the ten octaves from C_0 to C_{10} (120 half-steps) when tones are sounded simultaneously under laboratory conditions and beats are present.

The separate process of listening is often discussed and written about by conductors and music teachers and has long been held as one of the important skills necessary for good ensemble intonation. "Although all people of normal hearing can perceive sounds, listening requires active attending to information" (Lehman, Sloboda, & Woody, 2007, p. 205). Conductor Allan McMurray (1998) stated, "intonation begins with listening" (p. 59). Frank Battisti (1993), another respected university conductor, emphasized the importance of listening by suggesting that listening skills need to be taught.

Research into human sensitivity of perception was extensively examined in the 19th century by Ernst Weber, Gustav Fechner and others in the branch of psychology now known as psychophysics. Weber's research into the ability of subjects to detect the slightest changes in touch, vision, and loudness resulted in the phenomenon called just noticeable difference (jnd) — the slightest difference between two stimuli that can be detected (Weber, 1834). Fechner (1860), building on Weber's work, assigned a value of 0 to the point where observers could no longer detect a stimulus and devised a numeric scale for jnd values above that point.

We have seen that the ear is physically capable of perceiving slight differences in pitch. But where perception involves becoming aware of something via the senses, discrimination involves making fine distinctions between things perceived. When subjects focus their attention on listening for tuning discrepancies, they do not always accurately discriminate whether a second pitch is lower or higher than its paired reference pitch (Geringer & Witt, 1985; Rodman, 1981).

Non-musicians tend to be satisfied when the tuning of two pitches is merely close. However, performing musicians are held to a higher standard than the average person. In intonometry, the study of the measurement of pitch, pitch is considered to be the "essential perceptual parameter of tone," and pitch is in-turn "governed by the principle of categorical perception" (Fyk, 1995, p. 27). In speech related research the term *categorical perception* refers to the human tendency to group similar sounds into zones or categories and perceive similar sounds to have the same meaning, while ignoring sounds that do not contribute to that meaning (Siegel & Siegel, 1977). For example, a listener will categorize vowel and consonant sounds (i.e. *ay, ee, eye, bee, cee, dee*) spoken by one person as the same when they are spoken by another person, even if their voices differ in pitch, octave, and timbre. This ability to categorize sounds allows humans with different voice qualities, and even different accents, to communicate via language. Interestingly, musicians have been found to perceive intervals in a similar manner, categorically identifying two intervals as the same even when they differ by as much as a quarter-tone (Burns & Ward, 1978). In fact Siegel and Siegel (1977) found college musicians unable to tell the difference between sharp and flat test intervals. These student musicians were adept at identifying the intervals, but were not adept at determining whether they were in tune with reference tones.

Unlike the rigid, mathematics based tuning of Pythagoras and his followers, some contemporary researchers approach intonation as a system where single pitches interact as intervals either horizontally or vertically and are defined in context (Zanette, 2008), with musical context comprising some mix of the elements of rhythm, intensity, duration, tempo, pitch, timbre, envelope, articulation, melodic contour, tonality, and harmonic progression. Musicians must learn to attend to and adjust the pitch of their instruments or voices to match the intonation of other performers. They may do this in a solo setting or while listening to various combinations of instrument timbres in simultaneous or sequential listening contexts. Researchers have asked subjects to make decisions under many different conditions in an attempt to understand what happens in real-life music making contexts. Essentially the melodic or harmonic role of a tone has been shown to affect its tuning, with tones tending to lead to adjacent tones in melodic settings (Sogin, 1989; Swaffield, 1974; Yarbrough, Karrick, & Morrison, 1995).

Research concerning intonation may be organized into two broad categories, pitch discrimination and pitch matching (or performance). Discrimination in this context is defined as the "perception of quantitative or qualitative differences; the detection of similarities and differences" (Price, 1986). Pitch discrimination encompasses the process of hearing a pitch and deciding whether or not it is in tune with a reference pitch, hearing a pitch as an interval member, or hearing pitch in an implied tonality. Fyk (1995) stated that these decisions occur during three main stages of tone production; the initial stage, the quasi steady state, and the later stage. When musicians, while in the act of performing, focus their listening on intonation they are required to make judgments about direction and magnitude of mistuning, and whether they are hearing timbre discrepancies, pitch discrepancies, loudness discrepancies, or some combination of these.

Pitch matching begins with pitch discrimination, but extends to the skill of producing a pitch that matches a given reference. Accurate decision-making is at the heart of both processes.

Researchers have tested listeners' ability to perceive mistuned pitches in both non-musical and musical contexts (e.g., listening to pairs of computer generated tones in laboratory settings vs. listening to tones within instrumental recordings). They have asked listeners to identify mistuned pitches that have been altered in timbre, duration and tempo. The effects of intensity, octave, and direction of mistuning have been examined, as have the possible effects of different reference pitches, frequencies, and octaves. Researchers have also examined pitch discrimination by asking listeners to make decisions as to pitch relationships – same/different, lower/higher, or lower/same/higher.

Listeners' pitch matching abilities have been studied under performance conditions to include performing on brass, woodwind, and stringed instruments, as well as vocal performance. Performance tests have included the simple vocal and instrumental matching of reference tones as well as the performance of single line melodies or various intervals. Performance tests have also asked subjects to manipulate knobs and dials on tape players, computers, and keyboards in order to match computer-generated or pre-recorded pitches.

From Delezenne (1826) to Byo, Schlegel, and Clark (2011) many studies have included the examination of the pitch perception and performance abilities of listeners from various age groups and experience levels. Listeners have varied in age and experience from elementary school, middle school, and high school students, to college students and adults. Amateur and professional musicians and non-musicians have been tested at all levels and more experienced musicians have been shown to make more frequent, and better, tuning decisions than those with less experience (Duke, 1985; Madsen, Edmonson, & Madsen, 1969).

The human ear is amazing in its ability to register many different frequencies, timbres, and intensities simultaneously and translate these stimuli into electronic pulses which are then transformed by the brain into what we perceive as sound. The ear is able to recognize many different timbres by combining numerous simultaneously occurring overtones into individual tones, each with its own specific timbre. When perceiving spoken language, the ear is able to overlook slight differences. We automatically categorize the sounds of language so that spoken communication by one person is equal in meaning to spoken communication by another. This categorical perception is critical for the understanding of language, but it does not allow for the minute levels of differentiation necessary for satisfactory musical performance. These finer discriminations are something that we do not as yet fully understand.

Statement of the Problem

Only a few studies have included measurement of subjects' ability to detect differences between reference and test tones in both sharp and flat directions at specific levels of mistuning. Seashore (1938) reported a randomly selected group of adults to be 80% accurate at an average mistuning of 12 cents when indicating which of two presented tones was higher. Bentley (1966) reported that young children could not reliably discriminate between two sequentially presented tones at difference levels smaller than four cents. Madsen, Edmonson, and Madsen (1969) found that elementary through college aged subjects perceived changes in a modulating frequency most accurately during the first ten cents of change. Parker (1983) found student trombonists and violinists were accurate to “about” 20 cents when comparing paired sinusoidal reference tones with test tones altered from 10 to 100 cents sharp. Based on a synthesis of related literature Karrick (1998) set six cents as the threshold in his study, categorizing responses that varied six cents or less from equal temperament as in-tune (p. 120). Byo, Schlegel, and Clark (2011) used

a five cent threshold to determine the number of sharp, flat, and in-tune responses in tuning performance. If a listener's response was five cents sharp or flat, it was considered to be in tune. It was assumed that there is some range around 0 cent deviation that should count as in-tune, given the limitations of human pitch perception. Morrison (2000) took a different approach and did not allow for any range of deviation. In his study, for a response to be in-tune frequencies had to match exactly.

Relative to aural discrimination, researchers and active musicians often distinguish between objective reality and practical reality without knowing precisely what practical reality is. "Two cents sharp" indicates a real, empirically-measured difference between two pitches, but it may not be consistently perceived that way by trained musicians, even those with "good ears." In fact, a two cent difference may be perceived as no difference. In this example, the question centers on the performance of the human ear, scalar and harmonic implications aside. In practical terms, is 0 cent deviation in-tune, or is there a range of difference around 0 cent deviation that constitutes in-tune? To date research has yet to establish a range, with multiple studies providing conflicting results (e.g., Bentley, 1966; Byo, Schlegel, & Clark, 2011; Karrick, 1998; Madsen, Edmonson, & Madsen, 1969; Parker, 1983; Seashore, 1938).

More research is needed to determine the range within which wind instrumentalists accurately perceive pitch difference so that music teachers and conductors have a better understanding of what they can expect aurally from performers. Moreover, researchers in aural discrimination are making methodological decisions about in-tune performance that cannot be grounded in a research base when that base consists of few studies whose collective results are inconclusive. Therefore this study will examine the following questions:

How precisely do college and high school wind players perceive test tone pitches when they are presented both above and below a given reference pitch? Is musical experience a factor; do college musicians perform in this task differently than high school musicians? Are there differences when the pitches are sounded together versus one after the other? Does the actual pitch make a difference? How will the responses to these conditions compare condition to condition and high school to college? Finally, do pitch and/or timbre affect pitch perception between these groups and conditions?

CHAPTER 2
REVIEW OF LITERATURE

Tuning and intonation, the territory of music performers and teachers, resides within the larger category of acoustics, the territory of scientists. Proponents of basic or pure scientific research and proponents of applied science research approach the discipline with different motivations. Scientists conduct research to obtain and extend knowledge with muted concern for practical application, while performers and teachers conduct research with an eye toward useful application (Madsen & Madsen, 1997). Research results in both areas provide important context for the present study.

Scientists have been investigating pitch perception since the 19th century and their research has revealed much about human hearing and pitch discrimination. Given that musically expressive performance involves both a listen-and-decide element, and a sound production element, quantitative research examining music perception and performance dominates the literature. Often, and by necessity, research focuses on isolated elements in controlled settings, i.e. nonmusical contexts. The complexity of factors in musical listening makes it necessary to remove some of what makes music whole (and in-context) in order to isolate and examine single variables. Essentially the accurate measurement of any one of the many component elements of music may be confounded by the presence of others. This is true for both music performance and music perception tasks. As more sophisticated scientific instruments were and continue to be invented, scientists have devised more advanced and focused experiments designed to test the aural discrimination abilities of human and sometimes non-human subjects (e.g., Bernstein & Oxenham, 2003; Cramer & Zeitlin, 1955; Geringer, 1991; Wever & Bray, 1930). Results from single variable studies provide researchers with empirical information which they can consider

when developing more informed and more complex hypotheses. The process, over time, of combining the results of single variable studies into more complex multiple variable studies allows researchers to discover more of what is true and real in music.

Research into perception of pitch and tone has historically employed one of two measurement scales. Those who examined physiological questions about how humans hear and perceive pitch (Shower & Biddulph, 1931; Spiegel & Watson, 1984), timbre (Cramer & Zeitlin, 1955; Henning & Grosberg, 1968) and loudness (Wever & Bray, 1930, 1936; Sundberg & Lindqvist, 1973) reported their findings in the *Journal of the Acoustical Society of America* (JASA) using the cycles per second (cps) or Hertz (Hz) designations. In 1960 The General Conference on Weights and Measures officially replaced cps with Hz in honor of the 19^{th} century German physicist Heinrich Hertz. In contrast, research reported in the *Journal of Research in Music Education* (JRME) has tended to measure pitch using cents (100 cents = one semitone) as the standard, or convert results originally obtained in Hz to cents (e.g., Geringer, 1976, 1978, 1983). The "cps" and "Hz" terms are used in connection with physical measurement, in keeping with a pure research motivation. The term "cents" is used in connection with perceptual measurement, in keeping with an applied research motivation. Since this document reviews research from the physiological and perceptual approaches to studying pitch perception and discrimination, the Hz, cps, and cent designations all appear, depending on the terminologies used in source writings.

How well performers are able to notice intonation discrepancies and how well they are able to adjust pitch accordingly have been investigated from a number of different perspectives. An attempt to organize the literature chronologically proved unsatisfactory because research examining musical elements such as intensity, timbre, temperament, and cent deviation as well

as pitch discrimination and the effects of age and musical experience has not always been conducted in a coordinated and linear manner over the decades. An attempt at an organization based on researchers' lines of research was also unsuccessful because, while some lines exist, they are neither long enough nor connected enough to adequately cover most published research. An examination of the organization of other literature reviews revealed a third option, topical organization. A topical format was found to allow the flexibility needed to include studies from both pure and applied scientific research sources and to organize studies chronologically within topics when appropriate. Therefore this review is organized under eight topical headings; Intensity of Sound, Timbre, Temperament, Age and Musical Experience, Direction of Approach, Pitch Discrimination, Simultaneous and Sequential Presentation, and Precursors.

Intensity of Sound

Intensity was one of the first elements of musical performance to be examined using quantitative scientific methodology and equipment. Early experimentation in this area provided important historical foundations for research in tuning and intonation. Two important predecessors of modern scientific research into human perception were Ernst Heinrich Weber (1795–1878) and Gustav Theodor Fechner (1801-1887). Weber, often referred to as the founder of psychophysics, conducted numerous experiments examining visual and tactile sensory perception (Weber, 1834). In one notable experiment he tested human subjects by asking them to compare the difference between weights lifted in one hand with those lifted in the other. Results revealed that the difference in sensation was more accurately described as a ratio than as an absolute difference. If the differences were small, the subjects could not detect a difference. Weber's work in this area is an important foundation to discrimination research and it led to

Fechner's codification of the concept of *just noticeable difference* (jnd); a measure of the resolving power of a sensory system.

Fechner (1860) developed the jnd concept as a result of experimentation which included the following aural experiment. He first presented a stimulus tone of a specific intensity level to several subjects individually. He then reduced the intensity level of that tone until the point where each subject began to hear a change. He notated the amount of reduction in intensity for each subject and computed an average, calling that average one jnd. He then presented a tone one jnd softer than the first tone to each subject and repeated the measurement process, calling the resulting average two jnds. He continued this stepwise lowering of the intensity level of the tones until the final tone for each participant became inaudible. The average number of steps it took until a tone became inaudible equaled the number of jnds of that tone. He later devised a mathematical logarithmic formula to calculate the relationship between the physical and psychological magnitude of sensory stimuli which he named *Weber's Law* [$S = K \log I$] (Fechner, 1860). Advances in the creation and development of scientific instruments in the 20th century allowed for even more focused experimentation examining human perception, including the various functions of the ear and auditory nerve.

Wever and Bray (1930) were among the early pioneers who employed empirical research techniques using electrical test equipment to investigate the relationship between the electrical/physical processes that occur in the cochlea and acoustic nerve, and the frequency and intensity of sound. They surgically connected electrodes to the auditory nerves of anesthetized cats and then exposed the cats to sound stimuli of different intensities. The resultant electrical signals detected by the electrodes were amplified and played back to the researchers through speakers. The researchers were able to hear sounds as soft as a whisper, and variations in

intensity level were also detectable. These experiments established empirical evidence of a positive correlation between the intensity level of the sound stimuli presented and the strength of electrical signals detected in auditory nerves.

Wever and Bray continued their research and published results in 1936 describing the relationship between sound intensity and the magnitude of responses at different frequency levels within the cochlea of guinea pigs. They chose these animals due to similarities between their cochlea and those of humans. Test results indicated that as stimulus intensity was increased, cochlear response also increased. However, graphed measurements for stimulus tones of higher frequency tended to curve and reach a maximum level of response at lower intensity levels than those of lower frequency for all subjects. Essentially the cochlea reacted differently to stimulus intensity levels at different frequency levels.

Several early tests designed to examine parameters of aural music perception used techniques "borrowed" from audiometric studies which favored the use of headphones and sine tones versus complex tone stimuli. In one such experiment Cohen (1961) asked musicians (N = 10) to take a paired comparison listening test using sine tones of seven different frequencies (between 50 and 6000 Hz) at four different loudness levels and indicate whether each test tone was perceived as lower, equal to, or higher in pitch than its paired reference tone. Results showed louder sine tones to be judged lower in pitch than softer tones of the same frequency and wave form at lower frequencies, no apparent change at 1500 Hz, and a slight rise in pitch with increasing loudness for tones at 6000 Hz.

Fastl and Zwicker (1999) described a phenomenon called *pitch shift* where different sound pressure levels, measured in decibels (dB), directly affect the perception of the pitch of pure (sine) tones (p. 113).   Louder pure tones at lower frequencies were perceived differently

than they were at higher frequencies. For example, a 200 Hz sine tone presented at 80 dB sounded lower than a 200 Hz tone presented at 40 dB.

Although studies using pure sine tones provided an important foundation for research into musicians' perception, later experiments using more complex tones more closely approximated the sounds to which musicians regularly attend. Sundberg and Lindqvist (1973) provided evidence of a marked difference between results from experiments using complex tone stimuli and those using pure sine tones. In one experiment using complex tones they found that differences in reference tone intensity resulted in different perception of tones an octave apart. Subjects ($N = 4$) tended to expand the octave above a 2:1 ratio in octave generation tasks using complex tones when dB levels were increased. The pitch of the test tones was perceived to rise from an intensity level of 65 dB to about 80 dB.

Timbre

Timbre is an important element of musical performance which may influence subjects' ability to perceive and judge differences in pitch. Since timbre may be described as "the attribute that distinguishes sounds that are equivalent in pitch, duration, and loudness" a review of research into its effect on intonation is warranted (Thompson & Schellenberg, 2004, p. 428). The timbres of the different instruments give them their unique sonic identities. According to Pierce (1999), when "thinking about and experimenting with the pitch of musical tones, we must distinguish the sense of pitch from a sense of brightness or dullness" (p. 57). Research into the effect of timbre on tuning perception and performance can be organized into four main categories: simple versus complex tones, effects of different harmonic tunings, juxtaposing timbres, and bright and dark timbres.

Simple versus Complex Tones

Research has revealed that electronically produced complex tones such as those built from square, sawtooth, and triangular waves elicit more accurate tuning responses than those built from pure sinusoidal waves (e.g., Cramer & Zeitlin, 1955; Henning & Grossberg, 1968). Researchers have suggested that this is due to the presence of harmonics in complex tones, which affect listeners' ability to perceive pitch (e.g. Plomp and Mimpen, 1968; Spiegel and Watson, 1984). It is important to consider timbre's effect on perception because bands and orchestras comprise instruments which produce a number of different complex wave forms to which musicians must attend during tuning, rehearsal, and performance. Cramer and Zeitlin (1955) presented subjects (*N* = 15) with multiple sequentially presented tone pairs consisting of a simple or complex reference tone, a simple or complex test tone, and a repeat of the reference tone. Subjects were asked whether the test tones were the same or different in pitch from their paired reference tones. They discovered that subjects were significantly better at recognizing pitch differences in complex tones (rectangular pulse waves) than in pure (sine wave) tones. The context of this experiment was decidedly non-musical in that it was designed to test American soldiers' sensitivity to different types of sound displays being considered for use with land mine detectors. However, it is historically important because it is one of the early empirical studies that combined simple and complex reference and test tones at different pitch levels with a forced choice format.

Henning and Grosberg (1968) attempted to measure the ability of listeners to discriminate differences in frequency between complex tones (rectangular pulse tones) and simple tones (sine tones). Subjects (*N* = 2) were tested while wearing headphones and indicated on an answer sheet which of two sequentially presented tones was of higher frequency. Subjects were found to perform four times as well when detecting differences between tones around 250

Hz for complex (pulse train) tones than for sinusoidal tones of the same frequency. Performance at the 1000 and 2000 Hz levels was also better for the complex tone stimuli, but at the 4000 Hz level there was no significant difference in performance. This led Henning and Grosberg to conclude that since the harmonics of the very high frequency tones are above the range of normal hearing, subjects were no longer able to benefit from the presence of harmonics in the complex tone stimuli.

Spiegel and Watson (1984) tested subjects (*N* = 60) for their ability to identify both matching and mistuned sine and square wave tone pairs presented through speakers in a large auditorium. Results showed subjects were significantly better at detecting differences in the more complex square wave sounds. They suggested that one possible reason for this superior detection ability was that "listeners may be listening to many harmonic components simultaneously and combining information" (p. 1692). Although the sound pressure level was reported to be 75 +/- 5 dB throughout the auditorium, Sergeant (1973) argues that the different acoustic properties of different spots throughout a room may result in significantly different dB levels. Sergeant found differences in stimulus levels as much as 25 dB from subject to subject across a testing room and posited that such differences could affect the accuracy of test results.

The research examined so far was designed to test subjects' pitch perception under various conditions. Platt and Racine (1985) designed a performance-based study which tested for possible effects due to tone complexity by having string players attempt to accurately tune test tones in a paired comparison format. Subjects (*N* = 12) adjusted a dial in order to match a series of simple or complex test tones paired with simple reference tones. Subjects were found to be more accurate when tuning reference and test tones of the same timbre. They heard

complex tones as sharp relative to simple tones. In a second similar experiment, complex reference tones were found to produce more accurate results overall.

In order to test for possible effects of performance instrument, reference tone timbre, and reference tone octave on subjects' ability to perform in tune, Cassidy (1989) tested high school instrumentalists' (*N* = 24) ability to match reference pitches presented in three different octaves (below, same, and above), at eight different pitch levels, using sine, square, and sawtooth wave reference tones. The subjects (12 flutists and 12 clarinetists) listened via headphones and were recorded while matching the various reference tones. Analysis revealed a significant interaction between timbre and octave placement variables. Sine wave and square wave stimuli provided the most accurate tuning results when presented an octave below the tuning note, while sawtooth waves provided the most accurate results when presented in the same octave as the tuning note. In contrast to the findings of Greer (1970) no interaction was found between the timbre of the stimulus tone and the timbre of the instrument performed.

Geringer (1991) asked musician (*n* = 54) and non-musician (*n* = 54) subjects to listen to a series of pre-recorded electronic tones, synthesizer keyboard performances, and recorded instrumental performances and indicate whether each example increased, decreased, or stayed the same in intensity. Although this study examined the effect of timbre on the perception of intensity, it is important because Geringer found that timbre mattered. Both musicians and non-musicians were able to discriminate changes in intensity in the electronic tones more quickly than either the recorded musical performances or the synthesizer performances.

Effects of Different Harmonic Tunings

A review of research that examines tones with different harmonic structures is important here because harmonic structure determines timbre and timbre has been shown to affect subjects'

perception of intonation. The timbre of each instrument is determined by the presence, absence, and relative strengths of the harmonics present in each tone. Specifically, musical tones sound brighter when the area along their frequency range where partials are strongest (their spectral centroid) is higher than for darker tones of the same pitch (Schubert, Wolfe, & Tarnopolsky, 2004). Essentially the presence and relative strengths of partials above a given tone's fundamental determine its brightness or darkness and may affect how subjects perceive its pitch. Studies controlling for the presence, absence, and/or strength of selected partials have produced different intonation perception outcomes (Hartmann, McAdams, & Smith, 1990; Plomp & Mimpen, 1968).

Some researchers have theorized that the tuning of individual harmonics may affect subjects' ability to discriminate between tuned and mistuned tones. Plomp and Mimpen (1968) investigated the limits of subjects' (N = 6) ability to analyze frequency by devising an experiment that measured how many harmonics in a complex tone subjects could hear separately. While listening to a series of reference tones via earphones, subjects used a selector switch to indicate which of two sine tones paired with each reference tone exactly matched a specific harmonic of that reference tone. Results revealed that subjects could only accurately match tones at the level of the first five to seven harmonics. Plomp and Mimpen suggest that results from this experiment support the critical-band theory. "In a complex tone, the critical bandwidth corresponds to the smallest frequency difference between two partials such that each can still be heard separately" (Truax, 1999). Along the length of the basilar membrane each critical band is approximately 1.2 mm in length and each contains about 1300 receptor cells. According to Truax the basilar membrane contains 24 such critical bands.

In order to test the utility of the critical band theory Sergeant (1973) designed an experiment which tested subjects' (*N* = 56) ability to discriminate between tuned and mistuned tone pairs. He developed a listening test comprising two sets of tone pairs. One set pairing tuned or mistuned square wave tones and the other tuned or mistuned grand piano tones. The difference between subjects' mean scores of 17.45 for the electronic square-wave test and 17.1 for the piano tone test proved statistically non-significant. Sergeant therefore rejected Plomp and Mimpen's (1968) premise that subjects more accurately judge the pitch of more complex tones due to enhanced basilar response. He argued that the piano tones used in his study were more complex than the square wave tones, yet the results from both sets of tones were almost identical.

Hartmann, McAdams, and Smith (1990) conducted a series of experiments testing their own (*N* = 3) abilities to correctly identify mistuned harmonics within a complex tone. While listening through headphones they used a dial to match the pitch of a sine tone to a single intentionally mistuned harmonic presented within a complex tone. Only lower harmonics ranging from 12 harmonics for the 800 Hz fundamental tone to 16 for the lower 200 and 400 Hz fundamentals were tested. Results showed subjects to be more accurate detecting mistuned lower harmonics than they were detecting mistunings at 12 harmonics and above. The authors suggested "that the individual [perceptual] differences observed in our experiments are not due to random error but represent genuine idiosyncratic perceptual effects" (p. 1716). Further experimentation with a larger *N* would improve the likelihood that their results truly reflect the perception of the larger population.

Although the current study only used diotic presentation, it is important to acknowledge research that has been conducted investigating potential differences between diotic (the same stimuli presented to both ears simultaneously) and dichotic (stimuli presented to one ear only, or

different stimuli presented to each ear) perceptive abilities. Bernstein and Oxenham (2003) tested musically trained subjects (*N* = 4) to see if they would more accurately identify mistuned harmonics above the tenth harmonic diotically or dichotically (in this dichotic case even numbered harmonics were presented to one ear and odd numbered harmonics presented to the other ear.) In one experiment test tones comprising pure sine tones were gated on and off during the sustained presentation of complex tones at frequencies of 100 Hz and 200 Hz and containing the first 40 harmonics. Subjects were presented with gated test tones at various harmonic levels and asked to determine whether the test tones or the sustaining reference tones were higher in frequency. Results showed that subjects were able to identify tunings with an accuracy rate of 75% for pairs as high as the twentieth harmonic in the dichotic condition. This result is significantly higher than the five to eight harmonic level reported by Plomp (1964) and Plomp and Mimpen (1968), or the twelve harmonics reported by Hartmann, McAdams, and Smith (1990) for diotic listening tests, and indicates a significant difference between diotic and dichotic perception in the subjects tested.

<u>Juxtaposing Timbres</u>

In ensemble performance settings musicians regularly perform simultaneously with instruments of like timbre as well as instruments of different timbre. Investigation into the possible effect of different timbre combinations on subjects' pitch perception is important to a clear understanding of pitch perception in general. Researchers have examined this problem in several different ways.

In one performance-based study examining the effect of timbre on intonation, Greer (1970) recorded brass players (*N* = 32) while they performed twelve-tone duets along with recordings of organ (set on the flute stop), piano, oscillator, and their own instrument. Each tone

of each subject's resultant recording was analyzed for cent deviation from the reference and results revealed a significant effect due to timbre. The least deviation occurred when matching their own instrument tones and the most deviation (flat) was obtained when matching oscillator and organ tones. Reference tones with more complex timbres resulted in performances with more accurate pitch.

Madsen and Geringer's (1976) perception-based study examined subjects' ability to discriminate between poor tone quality and poor intonation as well as the effect of mistuned accompaniments. College music students (N = 50) first listened to two recorded unaccompanied trumpet solos and rated them for tone quality. They then listened to eight different sets, each containing three different combinations of sharp, flat, and in-tune accompaniments combined with trumpet solos with either good or bad tones. The subjects were able to accurately discriminate between good and bad tone quality in unaccompanied recordings, but not in accompanied settings. They preferred sharp and in-tune accompaniments significantly more than flat accompaniments, and noticed intonation problems significantly more than tone quality problems in the accompanied examples.

Wapnick and Freeman (1980) also examined possible effects of timbre on the perception of intonation and found that college music students confused poor tone quality with poor intonation. They found that subjects (N = 50) associated dark with flat and bright with sharp. Subjects made less accurate pitch decisions when test tone timbres had been altered, and more accurately identified flat tones than unaltered or sharp tones when reference and test timbres were the same. The researchers presented undergraduate music majors (N = 50) with a paired comparison task containing 48 pairs of clarinet tones; 24 based on A = 440 Hz and 24 based on A = 880 Hz. Both tones of each pair were electronically altered to either a bright or dark

condition and were presented via loudspeakers. The second tone in each pair (test tone) was also altered 12 cents sharp, unaltered, or altered 12 cents flat. Subjects indicated on an answer sheet whether the test tones were lower, the same, or higher than their paired reference tones. Results showed subjects to make significantly more errors when tones were of different timbres than when they were the same timbre. Subjects also incorrectly answered "flat" significantly more often when the tones were presented in the bright-dark order, and incorrectly answered "sharp" when the order was dark-bright.

Madsen and Geringer (1981) also found that college music students confused intonation and tone quality problems. Subjects (N = 480) incorrectly perceived intonation errors more often than tone quality errors while rating the intonation and tone quality of 24 duet recordings by professional oboe and flute performers. Each recording contained two examples of each of the 12 possible combinations of good and bad flute and oboe tones and intonation. Subjects listened to all examples in small groups, as a pilot study showed no significant difference between this setting and testing with individual headphones. Results showed music majors providing significantly more correct responses than non-majors. Sixty-two percent of responses by all subjects indicated "flat" even though flat notes were not recorded. The authors listed subjects' ability to detect that a problem existed, but inability to identify it as either a tone or a tuning problem as the most important finding of this study.

The following study is included here because, although it was primarily an error detection study, it examined the variables of timbre and context; important variables in the present study. Byo (1993) investigated the combined effects of timbre and context on the detection of performance errors by asking graduate and undergraduate music majors to listen to musical excerpts that included two different error types (pitch and rhythm), the placement of errors in

four different voices (soprano, alto, tenor, and bass), and textures comprising from one to four voices. Twenty taped excerpts were constructed, 16 of which included errors. Participants (*N* = 60) listened to the excerpts and responded by circling perceived errors in music scores. Results showed significant main effects for both timbre and error type. Participants correctly identified more rhythm errors than pitch errors and were more accurate in the single timbre setting. Analysis also revealed a significant three-way interaction between texture, timbre, and error type. Results lead Byo to suggest that musical context is a major contributing factor to participants' error detection ability.

Bright and Dark Timbres

The tendency for subjects to confuse bright timbres with sharpness and dark timbres with flatness as reported by Wapnick and Freeman (1980) was followed by later studies designed to focus even more specifically on this phenomenon. Ely (1992) tested undergraduate and graduate student instrumentalists' tuning ability using both performance and listening tasks in separate sessions. Subjects (*N* = 27) were recorded while playing along with melodies recorded by flute, clarinet, and saxophone faculty members. They later listened to the resulting unison duets and indicated whether they were in-tune or not. Interestingly, these subjects could detect intonation problems in dissimilar timbre combinations better than in similar combinations and they demonstrated different levels of accuracy with various timbral combinations. Subjects were able to detect intonation discrepancies best when listening to complex timbres.

Geringer and Worthy (1999) tested the effect of timbre on pitch perception by asking university (*n* = 72) and high school instrumentalists (*n* = 44) to compare a recorded test tone to recorded clarinet, trumpet, and trombone reference tones for timbre (dark, unaltered, bright) and pitch (flatter, the same, sharper) using a five step Likert-type scale for each. "Bright" test tones

were prepared by raising the amplitude of selected harmonics above the fundamental frequency of each sample tone 12 dB while leaving the fundamental frequencies unchanged. "Dark" test tones were prepared by lowering selected harmonics in the same manner. Results showed participants associating bright timbres with sharp intonation and dark timbres with flat.

In his experiment the following year Worthy (2000) prepared bright and dark tones in the same manner as for the Geringer and Worthy (1999) study and then added a performance component. As with Geringer and Worthy (1999), high school (n = 32) and college (n = 32) instrumentalists perceived bright timbre to be linked with sharpness and dark timbre with flatness. For the performance task subjects were instructed to either maintain optimum tone quality, or maintain the best intonation possible while matching a stimulus tone. Performance results mirrored perception results, but here the magnitude of flat response to dark was significantly more extreme than that of sharp response to bright.

Byo, Schlegel, and Clark (2011) examined how changes in timbre and octave affect college and high school wind instrumentalists' tuning ability. The researchers tested participants' (N = 72) ability to match tuning notes of different timbres by having them tune to four different reference tones; a B-flat$_4$ recorded by professional flute, oboe, and clarinet players and a B-flat$_2$ recorded by a professional tubist. Participants tuned more accurately to flute, clarinet, and oboe tones in octave four than they did to the tuba tone in octave two, even though 82% of them reported tuning to the tuba as the standard tuning method employed by their school bands. These results contrast Cassidy's (1989) results which showed subjects to be most accurate when sine (flute-like) and square (clarinet-like) wave stimuli were presented an octave below, and lease accurate when these stimuli were presented an octave above. She also found responses to sawtooth (oboe-like) stimuli to be most accurate when presented in the same octave.

These seemingly contradictory results may be due to factors other than octave. Also, similarities between tuning results obtained using wind instrument timbres and those obtained using synthetically generated tones cannot be assumed to be equivalent.

Temperament

A well informed decision as to which tuning system to use as a reference standard should precede any research into the perception of intonation. The three tuning systems most commonly referenced in the literature are Pythagorean tuning, just intonation, and equal temperament, although other systems are historically important. The system of equal temperament is “widely regarded as the normal tuning of the Western 12-note chromatic scale” (Lindley, 2001a, p. 275). This system is based on a cycle of 12 identical 5ths and an octave which is divided into 12 semitones (half-steps) of equal size. One of its predecessors, just intonation, is based on the consistent use of harmonic intervals “tuned so pure that they do not beat, and of melodic intervals derived from such an arrangement” (Lindley, 2001b, p. 290). A third tuning system, Pythagorean tuning, is based on 5ths and 4ths which are tuned to their pure mathematical ratios of 3:2 and 4:3 respectively. Mean tone tuning is similarly constructed, but the 5ths are slightly narrower than the perfect 3:2 ratios used for Pythagorean 5ths in order to allow for more agreeably tuned thirds. Since these systems derive their pitches in different ways, the scale tones, other than the octaves, are slightly different in frequency.

The following studies examined the perception and performance tuning accuracy of subjects in tuning systems currently or previously employed in Western music. Vos (1988) recorded a series of 24 pitch-altered melodic and harmonic fragments using computer generated complex tones which were altered in pitch to conform to one of five tuning systems; equal temperament, Pythagorean tuning, mean-tone tuning, Silbermann tuning, and Salinas tuning.

(Silbermann tuning, and Salinas tuning are mathematical variants of mean-tone tuning.) In the first of two perception-based experiments Vos asked musicians (*N* = 24) to listen via headphones and subjectively rate fragments played in equal temperament, Pythagorean, and mean-tone tuning for musical acceptability. He reported that subjects' average ratings showed "no differences worth mentioning" between the musical acceptability of fragments played in these three tuning systems (p. 2390). In a second experiment beats were removed from the harmonic fragments by deleting selected overtones from certain tones. Subjects' average acceptability ratings were significantly higher across all tuning systems for fragments which had been altered versus those that were unaltered. In both experiments analysis predicted higher overall acceptability ratings for fragments with perfectly tuned harmonic fifths and major thirds, suggesting that the tunings of fifths and thirds may contribute more to subjects' judgment of pitch than the tunings of specific overtones.

Rakowski (1990) looked for possible correlations between musicians' tuning of isolated intervals in the equal tempered, just, and Pythagorean tuning systems. He required music students (*N* = 4) to tune a variable frequency oscillator to obtain various melodic intervals above and below 12 different stimulus frequencies presented within one octave. These stimulus frequencies were presented via headphones as square, triangular, and sinusoidal wave tones. Results revealed that musicians tended to slightly compress the tuning of smaller intervals and slightly expand the tuning of larger intervals compared to their equally tempered values. He found no correlation between subjects' interval tunings and the Pythagorean and Just scales. He suggests, however, that this lack of correlation may not apply to harmonic settings. When subjects perform simultaneously with reference tones, the presence of beats may serve as a cue to intonation problems.

Loosen (1993) asked professional violinists (N = 8) to accurately record a major scale three octaves ascending and descending very slowly without vibrato in order to determine which tuning system, equal temperament, Pythagorean, or just intonation, violinists most closely approach when performing unaccompanied diatonic scales. He then drew several frequency samples from each recorded scale tone and computed means. Calculations comparing the mean absolute differences between the recorded scale tones and their theoretical values in the three tuning systems revealed that subjects performed most closely to both equal temperament and Pythagorean tuning and that observed interval sizes were not significantly different ($p > .45$) between the two systems. Subjects consistently performed less closely to just intonation.

Karrick (1998) conducted two performance-based experiments in an effort to determine whether any one of three tuning systems (equal temperament, just intonation, or Pythagorean tuning) yielded more accurate tuning results with experienced musicians. In the first experiment professional (N = 8) and college (N = 8) instrumentalists' recorded harmonic intervals which were compared to equal tempered, just, and Pythagorean tuning references. Here subjects were found to deviate least from equal temperament and most from just intonation. In a second experiment Karrick used an arbitrarily established "in-tune" threshold of six cents to determine the accuracy of deviation from equal temperament. Results from this experiment showed both groups to produce more sharp responses in melodic context when performing below the stimulus than when performing above. It may be that the more accurate results obtained from the system of equal temperament are due to subjects' familiarity with this system.

An examination of different tuning systems suggests that subjects were more aware of mistunings in harmonic contexts than in melodic contexts (Vos, 1988). Subjects were not bothered by tuning differences in melodic fragments played in different tuning systems, and they

were more accepting of harmonic stimuli from different systems when beats were eliminated by removing certain harmonics. Subjects were also found to deviate most from just intonation and least from the Pythagorean and equal temperament tuning systems while performing unaccompanied tasks.   Also, an increased sensitivity to harmonic mistunings (when beats are present), versus melodic mistunings, may be a significant factor in subjects' abilities to identify intonation discrepancies.

Age and Musical Experience

The potential effects of musical experience, age, and training on subjects' tuning perception and performance have been considered in several studies. Experiments researching intensity of sound, timbre, temperament, musical context, pitch discrimination, and pitch direction have examined the influence of age, experience, and training on subjects' various performance and perception abilities. A better understanding of the possible effects of these experience-related influences may lead to more focused and effective instruction in tuning and intonation.

In one early study examining age, experience, and training Mason (1960) investigated whether wind players of different experience levels—members of a faculty woodwind quintet and a student woodwind quintet—tended to perform closest to the Pythagorean, just, or equi-tempered tuning systems. Mason recorded each member's part from a transcription of Ravel's "Pavane pour une infante defunte" in both solo and ensemble settings. He then extracted and analyzed each of the tones of the G major scale for its deviation from equal temperament, just intonation, and Pythagorean tuning. In both solo and ensemble settings the members of the more experienced faculty quintet deviated least from equal temperament while members of the student

quintet performed closest to Pythagorean tuning. Both groups deviated most from just intonation.

Madsen, Edmonson, and Madsen (1969) tested subjects (N = 200) from several different age groups to see whether auditory discrimination ability increases with age and with musical training. Eight different groups of 25 subjects were drawn from the second, fifth, eighth, and eleventh grades along with college junior music and non-music majors, graduate music students, and music faculty. When presented with electronically produced F-sharp (369.99 Hz) stimulus tones which modulated up or down over time, younger subjects made more incorrect responses and answered "sharp" more frequently, while older subjects answered more accurately but with more of a tendency to identify tones as flat. These results suggest that older subjects may be less accepting of flat mistunings than sharp.

Madsen (1974) sought to determine whether there were consistent patterns of sharp and flat deviation in scalar vocal performance of grade school (n = 8), high school (n = 8), and undergraduate vocal (n = 8), violin (n = 8), and piano majors (n = 8). Subjects (N = 40) were recorded while singing one of eight possible combinations of the C and D major scales in either the down-up or up-down directional pattern. The less experienced subjects sang flat both before and after mid-test treatment sessions, while the more experienced subjects sang sharp, with the vocal majors performing the sharpest. These results also suggest that more experienced subjects tend to err in the direction of sharpness.

In order to examine possible differences in perception due to musical experience one group (n = 5) of adult musicians with experience tuning stringed instruments and one group (n = 7) without experience were tested for their ability to accurately tune computer generated test tones in a paired comparison format (Platt & Racine, 1985) using a potentiometer. Subjects with

tuning experience were found to be more accurate tuners. A follow-on experiment revealed that practice with both auditory and visual feedback slightly improved test results for complex tones. Essentially subjects (n = 22) with musical experience, not just string tuning experience, were better tuners than those without experience (n = 10).

Duke (1985) examined the consistency with which subjects performed melodic and harmonic intervals relative to equal temperament and across different age groups. Junior high (n = 16), senior high (n = 16), and college wind instrumentalists (n = 16) first recorded melodic intervals, four above and four below reference tones presented via headphones. They were then recorded playing harmonic intervals above and below their previous melodic recordings. Analysis of their recorded tones revealed significant differences in deviations from equal tempered reference tones, with the junior high and high school students erring sharp and college students erring flat. No significant difference in tuning accuracy was attributed to headphones, nor did verbal inducement significantly affect results. Melodic intervals slightly expanded while descending, and contracted while ascending across age groups. Interestingly college subjects' tuning of intervals tended to be flatter while their tuning of tones in melodic context tended to be sharper. Duke was careful to warn that statistically significant differences may not equate to musically significant differences; an important point to consider in intonation research.

Musicians have been shown to be more accurate that non-musicians at identifying pitch discrepancies when different timbres are involved. Geringer and Worthy's (1999) research previously discussed under *Bright and Dark Timbres* showed college music majors (n = 36) to be more accurate in their responses, and apparently less distracted by timbre variations, than non-music majors (n = 36) and high school musicians (n = 44). Worthy (2000) found a significant interaction between education level and timbre in both the perception and performance

components of his follow-on investigation with high school (*n* = 32) and college students (*n* = 32) apparently linking bright with sharp and dark with flat.

In order to determine whether trained musicians' pitch discrimination perception was more accurate than non-musicians when judging complex tones versus pure tones Sergeant (1973) presented two groups of subjects with a pair of two alternative forced-choice listening tests. The groups comprised musicians (*n* = 25) and non-musicians (*n* = 21). Test one included pairs of tuned and mistuned sine tones while test two was similarly constructed but used square wave tones instead. Results showed that the musicians scored significantly higher than the non-musicians for both the pure and complex tone conditions. Sergeant also argues that pitch discrimination ability cannot be based solely on basilar response patterns within the inner ear since improvement in discrimination has also been positively linked to training.

Spiegel and Watson (1984) tested both adult members of a major symphony orchestra (*n* = 30), and a group of adult non-musicians (*n* = 30) for their ability to identify both matching and mistuned sine and square wave tone pairs. Results showed musicians to score higher than nonmusicians as a group.

Yarbrough, Karrick, and Morrison (1995) asked elementary, middle school, and junior high school wind players (*N* = 197) to match a recorded tone (perception) by manipulating a variable-pitch keyboard, and then tune their instruments to match the same recorded tone (performance). First year players (*n* = 50) tended to tune flat for both the perception and performance tasks, while fourth year players (*n* = 26) tuned sharp for both. However, no correlation was found between their performance and perception scores. Years of instruction did not affect the direction of error for the perception task, and the mean absolute cent deviation from the reference tone for fourth year players was 14 cents.

Both of Morrison's (2000) context-based experiments, discussed in the next section, revealed that younger and older subjects were able to tune isolated pitches better than those within melodies. Both age groups performed significantly more sharp than flat responses, and the more experienced subjects tended to err in the direction of sharpness.

In summary, results have shown that older, more experienced subjects notice flat more than sharp mistunings. We have also seen that more experienced musicians err in the direction of sharpness in both performance and perception tasks. Musicians more accurately perceived intonation discrepancies whether timbres were altered or unaltered. And, while musicians are more accurate performing melodic versus intervallic tuning tasks, they outperform non-musicians in both areas.

Quantitative research examining tuning and intonation can be viewed from the perspective of the contexts in which researchers ask subjects to perform. In order to isolate and examine different elements that may affect pitch discrimination, researchers have designed experiments requiring subjects to perform tasks which may not be considered wholly musical. For example, presenting pairs of sine tones to one ear of selected subjects through headphones in a sound controlled booth (Cohen, 1961) may enable researchers to better isolate and measure subjects' perceptions of intensity, but the process removes much of the context present in musical performance. Asking musicians to match pitch while listening to reference tones (Byo, Schlegel & Clark, 2011) increases the demand on researchers' ability to isolate a variable or variables by including more of the context of live music. But testing subjects' perceptive abilities in actual musical performance settings remains a difficult and complex undertaking. The following sections include reviews of studies that procedurally involve subjects in various musical contexts, the first of which is direction of approach.

Direction of Approach

One of Madsen's (1966) early studies examined the intonation of groups of elementary school students (n = 8), high school students (n = 8), and undergraduate music majors (n = 24) while singing ascending and descending scales. They were recorded in two sessions singing the C and D major scales unaccompanied in both ascending and descending formats. A brief period of either practice, verbal instruction, or distraction was inserted between sessions to examine whether training would affect performance. Madsen's most important finding here showed a marked difference in results due to scale direction. The total absolute cent deviation of all subjects in the ascending scale was approximately four times that of the descending scale.

Madsen (1974) returned to the examination of intonation ability in context when he repeated his 1966 study with alterations, this time looking for possible patterns of sharpness and flatness in the subjects' (N = 40) singing of the C and D scales. This time each scale tone was measured for possible sharp or flat deviation from an in-tune reference scale. No significant difference was found between sharp and flat deviations or between the various pitches within the scales. There also was no evidence of progressive sharping or flatting as the subjects sang through the scales.

Swaffield (1974) examined tuning within the context of scales by testing undergraduate music students' (N = 25) perception of intonation in ascending scale fragments by requiring them to tune pre-recorded flute, clarinet, horn, and violin test tones to the implied tonics of paired scale fragments by adjusting the speed of a tape player. The test tones were two seconds long, were tuned to begin at either 20 cents flat, in-tune with, or 20 cents sharp relative to A = 440 Hz, and were presented in various timbres, note values, and intensity levels. Swaffield found that subjects tuned the test tones flat or sharp consistent with how they were initially presented. Also,

as reported in earlier research (Cohen, 1961; Fastl & Zwicker, 1999), in-tune tones were perceived as lower in pitch when their intensity levels were increased.

In order to investigate possible effects of both scale direction and accompaniment on tuning performance, Geringer (1978) recorded and measured undergraduate and graduate music student instrumentalists' (*N* = 96) intonation accuracy while performing both accompanied and unaccompanied ascending scale patterns. After making an initial recording of both patterns, one-half of the subjects were arbitrarily instructed to correct for sharp performance while the other half were asked to perform as in tune as possible. All subjects then either recorded the scale pattern again, or adjusted the recording of their first performance via the speed control of the tape player. Subjects' perception of intonation on accompanied scales (as measured by their adjustment of tape player playback speed) was found to be significantly better than on the unaccompanied scales, and verbal suggestion did not significantly affect either performance or perception results. Results also indicated a tendency toward sharp intonation with perception results significantly sharper than performance results. College students' apparent tendency to perform in the direction of sharpness (or possibly their distaste for flat intonation) supports the findings of Madsen (1974), Geringer (1976), and others.

Sogin (1989) measured college and professional string instrumentalists' (*N* = 48) intonation deviation by examining selected tones from both ascending and descending pitch sets, performed with and without vibrato. Results revealed that subject tones rose in pitch, with test tones in descending sets significantly sharper than in ascending sets and both ascending and descending test tones consistently sharp. Significant interactions with "direction by tone" and "direction by location by tone" indicate that direction and tone were interdependent. Subjects were only slightly sharper with vibrato. This study also supports earlier findings (Geringer &

Witt, 1985; Madsen, Edmonson & Madsen, 1969); older subjects tend to perform sharp more often than flat.

In order to see if the direction of approach to target affects intonation, Yarbrough, Karrick, and Morrison (1995) asked first (n = 50), second (n = 61), third (n = 60), and fourth (n = 26) year brass and woodwind instrumentalists to both adjust a variable-pitch keyboard, and to tune their wind instrument to match given F and B-flat concert reference pitches. The two stimulus pitches were presented to the subjects via headphones in the same octave as their normal tuning note. One half of the stimulus tones were presented sharp and one half were presented flat relative to the tuning standard. Subjects tended to perform sharp when approaching a target note from above and flat when approaching from below. "Years of instruction" was found to significantly affect subjects' direction of error with the elementary players answering "flat" more often for both perception and performance tasks and the middle school players answering "sharp" more often for both. A significant improvement in both perception and performance ability was found between first and third year subjects.

Geringer (1976) combined research into both perception and performance in musical context by examining how accurately college music majors' (N = 60) adjusted recordings of orchestral excerpts which had been intentionally mistuned either sharp or flat. Subjects tuned the excerpts to satisfy their own tuning preferences, with no reference pitch, by varying the playback speed of the tape player. Both direction and magnitude of mistuning were measured via cent deviation from the A = 440 Hz standard. Subjects tended to tune the test excerpts sharp significantly more frequently than flat, and the magnitude of mistuning was also much greater towards sharp. Sharp responses were found to be the most sharp with higher initial presentation pitch while flat responses were found to be most flat with lower initial pitch.

Pitch Discrimination

Among the earliest pitch discrimination experiments for which we have written documentation are Delezenne's studies measuring the perception of musical intervals. In 1826 he published results from experiments that he conducted using a monochord (Delezenne, 1826; Pickler, 1966) which examined the ability of subjects to detect tuned and mistuned intervals at various frequencies. Delezenne reported that subjects, both with and without musical training, were able to detect unisons, which had been mistuned at a level of approximately six cents (Shackford, 1962).

Moving forward to the early 20th century, two researchers at the Bell Telephone laboratory, Shower and Biddulph (1931), published results of an experiment which measured subjects' differential pitch sensitivity (the minimum change in frequency subjects could detect.) Stimuli were presented between 31 to 11700 Hz at sensation levels between 5 dB above threshold to "the maximum their subjects could tolerate" (Shower & Biddulph, 1931, p. 275). The researchers invented a rotary condenser that allowed them to test for thresholds of audibility at various dB levels. Using this device they presented a group of adult male subjects (N = 5) between the ages of twenty and thirty with a sequential series of three short sinusoidal tones in a different paired comparison design; an initial tone followed by a second tone of different frequency followed by a repeat of the original tone. They found that subjects were more sensitive to changes as frequency increased from about 62 Hz to 1000 Hz and their sensitivity remained essentially constant from there to around 8000 Hz where it began to decrease.

In an attempt to codify subjective perception of differences in pitch, Stevens, Volkmann, and Newman (1937) conducted research that required a small group of adults' (N = 5) to evaluate multiple pitch pairs. The subjects were asked to listen to 60 dB stimulus tones, and then adjust

test tones with a crank until they sounded half as high as their paired stimulus tones. Frequencies chosen for each test tone were recorded and averaged across all subjects. The average test tone perceived as "half as high" as the 1000 Hz stimulus was assigned the value of 500 mels. The 500 mel tone was actually 558 Hz. Subsequent steps on what became known as the *mel scale* were attained by repeating the halving process downward step by step.

Madsen, Edmonson, and Madsen (1969) tested eight different groups of subjects ranging from the second grade to college music faculty to determine whether the ability to hear changes in the pitch of a modulated frequency test tone changes with age. Modulated frequency F-sharp (369.99 Hz) electronic stimulus tones were presented to the subjects in three categories; without frequency alteration, with ascending frequency, and with descending frequency. Subjects were tested individually using headphones, an on-off response switch, and an answer sheet. Results indicate an improvement in auditory discrimination ability with increased age and increased musical training. Younger subjects made more incorrect responses and answered "sharp" more frequently, while older subjects answered more accurately and with more of a tendency towards flat responses. Overall, subjects incorrectly chose more "same" responses (427) than either of the other two (312 flat and 237 sharp). This finding is consistent with results reported in Geringer and Witt (1985) and Rodman (1981). Subjects were most accurate perceiving frequency change during the first five seconds of each tone (10 cents of pitch change) with 84.2% correct responses. Responses became consistently less accurate over time; the 25-30 second segment had the most responses, but a large number of them were incorrect.

Miles (1972) conducted research to determine how well beginning band students could learn to perceive beats between tones, and play in-tune with others in unisons, major thirds, perfect fifths, and triads. Subjects (N = 118) were taught to perceive and eliminate beats by

using an intonation trainer, a device comprising two amplified adjustable variable oscillators, and by playing along with the researcher. Since there was no time limit for each test session, the majority of subjects achieved perfect scores and all of the subjects were able to recognize beats. Most (95%) of the subjects were able to play a perfect fifth free of beats, and 88% were able to tune a major third free of beats. Results of this study suggest the potential for beat elimination as a teaching strategy and support the positive effect of focused intonation training.

Geringer (1983) examined preschool (*n* = 72) and fourth grade (*n* = 72) students pitch discrimination and pitch matching abilities with a listening task which required them to determine whether the second of two electronically produced pitches matched the first, and a performance task which required them to sing the final note of a simple melody. Significant differences were found due to age for the pitch-matching test, but not for the discrimination test. Geringer suggests that pitch matching ability might be somewhat affected by physical development and that pitch discrimination may be more the result of acquired skill.

Parker (1983) examined the difference between the frequency perception of university student musicians who played variable pitch instruments (trombonists (*n* = 20) and violinists (*n* = 20) and those who played a fixed pitch instrument (piano (*n* = 20). Stimuli consisted of 70 tone pairs which were made up of a fixed frequency sinusoidal reference tone and an altered frequency sinusoidal test tone (ranging up to 100 cents above fixed frequencies in 10 cent increments) presented in random order via tape through headphones. Stimulus tones were presented for one second followed by two seconds of silence followed by a four second break between pairs. Subjects were asked to indicate whether the two tones were heard as the same or different. Results showed no significant difference among trombone, violin, and piano players in

their ability to detect changes in pitch. The points at which all subjects began to accurately discern two different pitches were found to be approximately 20 cents.

In a study investigating possible differences between string player's intonation perception and performance, Geringer and Witt (1985) tested to see whether high school, college, and professional string players performed and perceived intonation at the same accuracy level. Two groups of string players, one of high school students (n = 60), and another comprising both college and professional (n = 60) players were instructed to tune to an oboe stimulus which was either 25 cents sharp, 15 cents flat, or in-tune; or to ignore the stimulus if it seemed incorrect. They were then asked whether the stimulus was sharp, flat, or in-tune and their perceptions were compared to their performances. Professional and college performers as a group tuned significantly sharp to all stimuli while high school subjects tuned flat to the in-tune stimuli. In the perception task only twelve of forty in-tune stimuli were judged correctly with more stimuli perceived to be flat than sharp. The college and professional subjects tended to play sharper than the high school performers suggesting either less tolerance for flatness, or a tendency of the more experienced performers to over-compensate in the direction of sharpness.

Fyk (1985) examined the effect of pitch, dynamic level, and duration on the vocal pitch matching accuracy of ten year old nonmusicians. Male (n = 13) and female (n = 15) subjects were asked to vocally match rectangular pulse wave reference tones. Reference tones were presented at three pitch levels (250, 440, and 500 Hz), at two different dynamic levels (37 and 54 dB), and at 20 different durations ranging from 6 to 2000 milliseconds (ms). Fyk found that pitch matching abilities varied greatly among the subjects depending on the duration of the reference tones. Louder sounds produced significantly more correct responses for reference tones of less than 200 ms duration, but were less important for reference tones of longer

durations. Subjects also tended to perform more accurately and respond more quickly to reference tones within their vocal ranges.

The effect of reference pitch length, pitch matching experience, and ear training were all examined in two pitch matching experiments conducted by Fyk (1987). First year (*n* = 12) and fourth year (*n* = 12) music education students attending a university in Poland participated, with first year students serving as the experimental group and fourth year students serving as the control. Subjects, one group with training and one without, matched electronic test tones to stimulus tones using a pitch control knob. Subjects were overall less accurate identifying discrepancies in pairs presented in a lower octave (A_2) and more successful judging pairs presented in a higher octave (C_6). Training was found to significantly affect the accuracy of pitch matching and discrepancies in pairs of lower pitch took longer to accurately identify than those of higher pitch. After training, subjects were 50% faster and more precise as stimulus tone frequency increased. Subject accuracy also increased with the duration of the stimulus tones.

Duke, Geringer, and Madsen (1988) tested music (*n* = 200) and non-music (*n* = 200) majors to see whether they could perceive tempo changes more accurately than pitch changes in recorded music. Subjects listened to ten different pairs of orchestral excerpts in which the second excerpt had been altered either in tempo, pitch, or both and indicated on an answer sheet whether the tempo and pitch had either increased, decreased, or remained the same. Analysis revealed that subjects perceived tempo changes more accurately than pitch changes and that music majors were better able to identify tempo changes in excerpts with slower subdivisions. Frequency alterations had little effect on perception of tempo, but tempo changes did influence subjects' perceptions of pitch changes.

Yarbrough, Morrison, and Karrick (1997) investigated the effect of experience, private lessons, and awareness of intentional mistunings on the ability of high school wind players to tune accurately. Subjects (*N* = 113) were assigned to one of three groups; the first group knew that both their instruments and a variable pitch keyboard were mistuned sharp, the second group knew that both were mistuned flat, and the third was not informed of mistunings. They then tuned their instrument (performance) and the keyboard (perception). Results showed no significant difference in performance or perception due to years of experience or treatment group. Subjects performed the perception task more accurately and those with private lessons tuned significantly better. Subjects performed in the sharp direction significantly more often than flat, but showed no propensity for sharp or flat in the perception tasks.

Morrison (2000) investigated the possible relationship between instrumentalists' ability to accurately tune isolated pitches and their ability to tune pitches in melodic context in two linked experiments. In the first experiment, instrumentalists in their first (*n* = 20), second (*n* = 51), third (*n* = 35), or fourth (*n* = 31) year of band experience were recorded matching a synthesized B-flat concert reference. They were then recorded while attempting to play in tune with a short synthesized melody. Both their recorded B-flat pitch and four recorded target pitches (Gs) taken from the melody were analyzed and compared to recorded standards. In a second experiment high school instrumentalists in their fifth (*n* = 41), sixth (*n* = 57), or at least seventh (*n* = 69) year of instrumental instruction were randomly assigned to one of three groups and either tuned as in experiment one, were asked to play as in-tune as possible without tuning, or simply recorded with the stimulus with no preparation. Unlike the findings of Sogin (1989) and Yarbrough, Karrick, and Morrison (1995) the direction from which pitches were approached was not found to be a significant factor with either age group in this instance. Morrison

suggested that the concert band practice of tuning regularly to B-flat may cause an automatic response that confounds our understanding of how musicians actually play in tune.

Ballard (2006) examined correlations between participants' perception and performance of stimuli tuned to equal temperament, Pythagorean, and just intonation and presented in several contexts. More specifically he compared their accuracy of vocal and instrumental performance with their ability to detect intentionally mistuned notes within a melody and within harmonic intervals. He found significant differences between undergraduate wind instrument majors' (N = 60) intonation in instrumental performance and pitch perception. Participants matched digitally produced recorded melodic stimuli significantly better than harmonic in instrumental performance, vocal performance, and perception. They also perceived and performed repeated pitches within the melody more accurately than changing pitches, and performed more accurately instrumentally than vocally. No significant correlation was found between the ability to perceive intonation discrepancies and the ability to perform in tune.

Hayes (2009) tested the ability of middle school instrumentalists (N = 87) to tune to a series of chromatically derived reference pitches produced by professionals performing on the same instrument. Students heard each reference pitch for three seconds and then produced a test pitch for approximately the same length of time. Participants were allowed to repeat the tuning process if dissatisfied with their sound production. After playing each test pitch, participants rated the intonation of their performance using a 5-point Likert scale. Participants' recorded pitches were converted to .wav files and analyzed for fundamental frequency using the Praat software program. Their intonation performance did not differ by grade level, but did differ by instrument, with flute players performing significantly worse than clarinet, alto saxophone, or trumpet players. The most in-tune pitches for all subjects were G_3 and A_3 and the least in-tune

pitches were A_4 and $A\#_4$. Correlations between participants and expert judges' evaluations were not significant.

Simultaneous and Sequential Presentation

Reference and test tones have been presented in either simultaneous or sequential formats to facilitate investigation of a number of different research questions in numerous studies. Research examining tuning in the simultaneous setting has analyzed tones performed in unisons, octaves, or intervals in isolation (e.g., Ballard, 2006; Byo, Schlegel, & Clark, 2011; Cassidy, 1989) and in melodic contexts (e.g., Morrison, 2000; Sogin, 1989; Yarbrough, Karrick, & Morrison, 1995). Subjects have tuned by adjusting the length of their instruments or by adjusting a tuning knob on an electronic device (Fyk, 1987; Yarbrough, Karrick, & Morrison, 1995). The only qualitative study reviewed (Miles, 1972) examined the presence or absence of beats as an aid to tuning accuracy in the simultaneous setting. Researchers have conducted experiments examining subjects' abilities to detect differences in frequency between sequentially presented tones (e.g., Cramer & Zeitlin, 1955; Schellenberg, 2002). In the simultaneous condition the presence or absence of beats has been reported to aid subjects in the tuning process (Hall & Hess, 1984; Olson, 1967; Vos, 1988). In the sequential condition, the presence of silence between pitches is a factor (Fastl & Zwicker, 1999; Harris, 1948). This presents an important question: which condition, if either, sets up listeners (players and teacher/conductors) to be more rather than less discriminating? None of the research reviewed has used both sequential and simultaneous presentation methods in one perception-based study.

Precursors

The following studies are important precursors to the present study in that they either share one or more component with the present study, or they pose one or more of the questions

upon which this research is based. Carl Seashore's *Measures of Musical Talent* (MMT) (1919) represents an important early contribution to empirical research in pitch discrimination. Seashore developed a series of tests which were administered using phonograph recordings and answer sheets and which measured subjects' ability to discriminate differences in pitch, intensity, time, and consonance, as well as measuring their tonal memory. The test battery was designed to be administered to public school students in the fifth grade, with Seashore suggesting that "this is the earliest age at which group measurements can be made satisfactorily" (p. 3). The pitch discrimination portion of the test, an important precursor to the present study given its purpose and methodology, presents recorded pairs of pure tones separated by sequentially diminishing frequency differences. The large span of mistunings used by Seashore, 30, 23, 17, 12, 8, 5, 3, 2, 1, and 0.5 Hz, convert to values of 115, 89, 66, 47, 32, 20, 12, 8, 4, and 2 cents respectively when calculated with A = 435 Hz as the tuning standard. It is interesting to note that the 115 cent level of mistuning results in test tones more than a semitone away from their paired reference tones. Subjects are asked to indicate on an answer sheet whether the second of the two tones is higher or lower than the first tone presented. Seashore (1938) later reported results of pitch discrimination tests which showed the "average threshold for an unselected group of adults" to be approximately 3 Hz (11.9 cents) at the A = 435 pitch level (p. 56). According to Seashore, measurements were averaged across "thousands of trials" (p. 56), but he did not provide data for pitch discrimination testing or for different age groups or levels of mistuning.

Like Seashore (1919), Bentley (1966) also designed a test to examine the level, measured in cents, at which subjects could accurately identify mistuned test tones in a paired comparison listening format. The first task in the pitch discrimination portion of his *Measures of Musical Abilities* (MMA) test began with a one second sinusoidal reference tone of A_4 (440 Hz)

presented via speakers, followed immediately by a test tone of equal length, one semitone (26 Hz, or 50 cents) above. Following a six second break the A_4 reference tone was again sounded and was followed by a test tone one semitone below. The length of test and reference tones, and the time between pairs was an important consideration in the design of the present study as well. The test progressed through pairs of progressively smaller intervals of 18, 12, 10, 8, 5, and 4 Hz, presented both above and below their paired reference tones in the same format. It is noteworthy that the differences in cents between reference and test tones decreased as the test progressed. However, the difference between the MMA's A_4 reference tones at 440 Hz and the smallest increments (4 Hz) presented in any of the tone pairs is 16 cents; a larger interval than several of the smaller intervals appearing in Seashore's MMT. Pairs of tones which were exact duplicates of one another were interspersed between some of the different test pairs. Subjects were asked to mark on an answer sheet whether the second tone was the same or whether it moved up or down. Unlike Seashore, Bentley inserted a no-change "same" response option between the down and up options on his answer sheet (Young, 1973). This lower-same-higher format is used in several more recent studies including the present research. Bentley limited the smallest difference between reference and test tones to 4 Hz (about 16 cents) because he suspected that factors such as room noise, room acoustics, electronic playback limitations, and subject head movements might interfere with discriminations of smaller differences (3, 2, and 1 Hz or 12, 8, and 4 cents). Whereas Seashore targeted his MMT at subjects no younger than fifth grade, Bentley asserted that his MMA could be administered to even younger elementary aged children. He reported that subjects' pitch discrimination mean scores steadily improved from age 7 to 14 (*N* not reported) and that mean scores for errors increased as interval sizes decreased.

Rodman's (1981) dissertation was discovered early in the research process and is important for several reasons. He posed similar research questions concerning pitch discrimination, the possibility of age and experience affecting discrimination, and the possible effects of sharp and flat mistunings. He also employed similar methodology, using a format which incorporated randomly ordered matching and mistuned pairs of reference and test tones. Finally, this dissertation was an important starting point for the researcher, providing important bibliographic information concerning previous research in this area. Rodman presented junior high (n = 622), high school (n = 671), and adult musicians' (n = 54) with a listening test designed to determine at what level(s), measured in cents, they were able to accurately identify mistuned test tones. Pitch perception ability was tested using forced choice, paired comparison listening tests with tuned and mistuned test tones presented in random order following their paired reference tones. Stimuli spanned five octaves and were presented at 2, 5, 10, or 15 cent increments above or below the reference, or with no-change in pitch. All tones were two seconds in length and were produced in triangular wave forms. In analysis Rodman relied primarily on descriptive statistics and, as such, results are not generalizable to a larger population. He reported for example that high school students responded 40% "lower" and 36% "same" when judging test pitches which had been mistuned ten cents flat. When judging test pitches which had been mistuned ten cents sharp, these same students responded "higher" 34% of the time and "same" 39% of the time. Adult subjects tended to respond "lower" to test tones that were mistuned by only 2 or 5 cents, and also tended to indicate a mistuning when there was none. The elementary subjects actually chose the "same" response for most test items. Neither adults nor students were able to discern two cent variances at better than a 33% accuracy rate. Consistent with previous research, Rodman's results indicated that in 18 of the 20

mistuning/octave pairs tested, subjects were more accurate identifying flat variances than they were sharp variances. Rodman suggests that "studies to define pitch discrimination parameters are needed" (p. 155).

Questions that arose while gathering data for Byo, Schlegel, and Clark's (2011) study about the mass tuning of the concert band motivated the researcher to question the exact limits of pitch discrimination for high school and college instrumentalists. By defining in-tune as 0 cents deviated from a tuning standard, Morrison (2000) chose a literal definition of in-tune rather than one that considers the capability of the human ear. The fact, however, that only 19 of Morrison's 685 total responses were judged to be in-tune suggests that accurate perception at 0 cent deviation in either perception or performance tasks may be an unrealistic expectation for human listeners. Wapnick and Freeman (1980) chose a 12 cent level because it fell roughly in the middle of the range of subjects' scores obtained during pilot testing. During pilot testing, subjects consistently responded reliably to stimuli which had been mistuned by 15 or more cents, and unreliably to stimuli mistuned at or below the 8 cent level. Karrick (1998) decided on an arbitrary threshold of six cents after reviewing the literature and finding that "detectable differences could range between 2 and 20 cents" (p. 119). Madsen, Edmonson, and Madsen (1969) found that subjects perceived changes in modulated frequency tones most accurately during the first five seconds of change, during which time the tones modulated ten cents. Byo, Schlegel, and Clark chose a five cent limit based on the research of Rodman (1981) who concluded that it was the "minimum distinguishable variance" for high school musicians (p. 140).

Researchers' selection of various pitch discrimination limits for their studies highlights the need for an empirically derived level, measured in cents, at which participants can be

expected to accurately perceive intonation discrepancies. Research is needed to answer questions about the capability of the human ear. At what point are two tones different, but so close, that musicians are not able to reliably detect that difference? When is it reasonable to say that a musician's ear is "good?" What level of pitch discrimination should teachers and conductors expect of wind instrumentalists of different ages and levels of experience?

Purpose and Research Questions

Therefore, the main purpose of this study was to investigate high school and college wind instrumentalists' pitch discrimination when judging pitch pairs separated by 0, 5, 7.5, and 10 cents. Specific research questions were as follows:

(1) Do participants demonstrate differences in their ability to accurately identify mistunings among cent deviation levels (plus and minus 10, plus and minus 7.5, plus or minus 5, and 0)? If so, what is the nature of these differences?

(2) Does timbre (same and different) affect participants' accurate identification of mistuned tones?

(3) Does pitch [B-flat$_4$ (466.165 Hz) and E_4 (329.628 Hz)] affect participants' accurate identification of mistuned tones?

(4) Does listening condition (simultaneous and sequential) affect participants' accurate identification of mistuned tones?

(5) Do age and experience affect participants' accurate identification of mistuned tones?

To answer the research questions, high school and college wind musicians (N = 128) participated in a thirty-minute process which included completing a 13-item participant data form and performing a discrimination task by listening, through headphones, to a pre-recorded

listening test. Participants listened to a reference tone and decided whether the test tone in each pair was lower than, the same as, or higher than its reference tone.

CHAPTER 3
METHODOLOGY AND PROCEDURES

Participants

The 128 musicians participating in this study included 64 from the high school level and 64 from the university level. High school participants were wind instrumentalists recruited from a successful high school band program in the southern United States, with success defined by consistent superior ratings at state-sanctioned large ensemble adjudicated events. Undergraduate university students were wind instrumentalists recruited from a major university school of music in the South. University participants comprised both music and non-music majors.

The sample was one of convenience and comprised volunteers. An attempt to balance woodwind and brass instruments yielded 31 woodwind players and 33 brass players in the university group and 37 woodwind and 27 brass players in the high school group. The distribution of specific instruments within the woodwind and brass families was roughly proportional to standard concert band instrumentation.

The data presented in Table 3.1 were collected from the Participant Data Form completed by each of the 128 participants. The male and female totals reflected the populations of the high school and university band programs selected for the study. There is a relatively even distribution of male and female participants as well as brass players and woodwind players both within and between groups.

Table 3.1

Demographics

	High School		University	
	n	Percent	*n*	Percent
N	64	50	64	50
Male	31	24	36	28
Female	33	26	28	22
Brass	27	21	33	26
Woodwind	37	29	31	24
HS 1	16	13	-	-
HS 2	27	21	-	-
HS 3	13	10	-	-
HS 4	8	6	-	-
U 1	-	-	17	13
U 2	-	-	17	13
U 3	-	-	13	10
U 4	-	-	17	13

Tables 3.2 and 3.3 provide a snapshot of the music instruction experiences for the high school and college participant groups. Not unexpectedly, the tables make apparent a great disparity of music performance experience between the groups. More than twice as many university participants (57) reported experiencing private lessons on their major instrument (and for an average of 4.6 more years) than did high school participants (24). Similar group differences were found in piano lessons.

Variables

Independent Variables

In order to emulate authentic conditions of real music making I involved participants in a listening experience built on four variables—pitch, cent deviation, timbre, and mode of presentation.

Table 3.2

Music Instruction (in Years) – High School Participants

Instruction	HS1	HS 2	HS 3	HS 4
Major Instrument	*n* = 16 *m* = 3.44 *SD* = 3.46 Range = 1-4	*n* = 27 *m* = 4.85 *SD* = 5.60 Range = 1-8	*n* = 13 *m* = 0.31 *SD*= 3.84 Range = 0-4	*n* = 8 *m* = 6.63 *SD*= 3.14 Range = 4-8
Major Instrument Lessons	*n* = 16 *m* = 0.13 *SD* = 1.32 Range = 0-1	*n* = 27 *m* = 0.82 *SD* = 4.91 Range = 0-3	*n* = 13 *m* = 0.54 *SD* = 2.29 Range = 0-2	*n* = 8 *m* = 0.25 *SD* = 1.22 Range = 0-1
Piano Lessons	*n* = 16 *m* = 0 *SD* = 0 Range = 0-0	*n* = 27 *m* = 0.89 *SD* = 9.52 Range = 0-6	*n* = 13 *m* = 0.46 *SD* = 5.76 Range = 0-6	*n* = 8 *m* = 0.13 *SD* = 0.94 Range = 0-1
Vocal Lessons	*n* = 16 *m* = 0.94 *SD* = 8.66 Range = 0-8	*n* = 27 *m* = 0.22 *SD* = 4.97 Range = 0-5	*n* = 13 *m* = 0.08 *SD* = 0.96 Range = 0-1	*n* = 8 *m* = 0.13 *SD* = 0.94 Range = 0-1
Other Lessons	*n* = 16 *m* = 0 *SD* = 0 Range = 0-0	*n* = 27 *m* = 0.04 *SD* = 0.98 Range = 0-1	*n* = 13 *m* = 0.31 *SD* = 3.84 Range = 0-4	*n* = 8 *m* = 0 *SD* = 0 Range = 0-0

Pitch

Since both the Western harmonic system and modern wind instruments are built around the twelve key centers made possible by equal temperament, the present research is based on stimuli constructed in the equal tempered system. B-flat $_4$ (466.165 Hz) and E_4 (329.628 Hz) were selected as the two reference pitches for this study in order to compare participants' discrimination between a common and a rarely-used tuning pitch. B-flat$_4$ was selected due to its popularity as the preferred tuning note in many concert bands. As Morrison (2000) and others

Table 3.3

Music Instruction (in Years) – University Participants

Instruction	U1	U2	U3	U4
Major Instrument	*n* = 17 *m* = 7.2 *SD* = 7 Range = 1-11	*n* = 17 *m* = 9.06 *SD* = 5.19 Range = 7-12	*n* = 13 *m* = 8.9 *SD*= 8.65 Range = 2-12	*n* = 17 *m* = 9.94 *SD*= 9.11 Range = 6-14
Major Instrument Lessons	*n* = 17 *m* = 3.94 *SD* = 10.24 Range = 0-8	*n* = 17 *m* = 6.29 *SD* = 10.66 Range = 0-11	*n* = 13 *m* = 3.92 *SD* = 10.63 Range = 0-10	*n* = 17 *m* = 5.94 *SD* = 15.71 Range = 0-11
Piano Lessons	*n* = 17 *m* = 1.24 *SD* = 9.00 Range = 0-7	*n* = 17 *m* = 2.18 *SD* = 16.98 Range = 0-17	*n* = 13 *m* = 1.38 *SD* = 12.53 Range = 0-13	*n* = 17 *m* = 3.65 *SD* = 17.49 Range = 0-14
Vocal Lessons	*n* = 17 *m* = 0.24 *SD* = 3.01 Range = 0-3	*n* = 17 *m* = 0.59 *SD* = 5.49 Range = 0-4	*n* = 13 *m* = 0.54 *SD* = 4.15 Range = 0-4	*n* = 17 *m* = 0.82 *SD* = 13.60 Range = 0-14
Other Lessons	*n* = 17 *m* = 0 *SD* = 0 Range = 0-0	*n* = 17 *m* = 0.59 *SD* = 8.72 Range = 0-9	*n* = 13 *m* = 0.31 *SD* = 3.84 Range = 0-4	*n* = 17 *m* = 0.29 *SD* = 4.85 Range = 0-5

have pointed out, familiarity with B-flat as a regularly used tuning reference may affect participants' responses to B-flat stimuli. For this reason he used G concert, a pitch not frequently encountered as a tuning note, as a second stimulus in his study. In other research Madsen, Edmonson and Madsen (1969) chose concert F-sharp (369.99 Hz) as the reference tone "in an attempt to avoid eliciting responses confounded by other associations" (p. 1469). I selected E_4 for the same reason: it is not routinely used as a tuning reference in band rehearsals. It has not been used as a reference pitch in the basic or applied literature reviewed, and band compositions and arrangements are rarely scored in E concert.

Both reference pitches were placed in octave four because (a) it is an octave commonly used in band tuning, and (b) pitch perception performance using square wave tones was found to be more accurate near 440 Hz than at higher pitch levels (Spiegel & Watson, 1984).

Cent Deviation

In order to test their ability to judge sharp, in-tune, and flat pitch conditions, participants were asked to indicate lower, same, or higher perceptions of pitch difference in a paired comparison setting. Test tones were constructed at seven levels; one exact duplicate and three different pitches both above and above the B-flat$_4$ (466.165 Hz) and E_4 (329.628 Hz) reference pitch frequencies. Exact frequencies for each test tone are listed in table 3.4 below.

Table 3.4
Cent Deviation Frequencies

Difference	Frequency
B-flat$_4$ +10	468.865
B-flat$_4$ +7.5	468.188
B-flat$_4$ +5	467.513
B-flat$_4$ +0	466.165
B-flat$_4$ -5	464.821
B-flat$_4$ -7.5	464.150
B-flat$_4$ -10	463.480
E_4 +10	331.538
E_4 +7.5	331.059
E_4 +5	330.581
E_4 +0	329.628
E_4 -5	328.677
E_4 -7.5	328.203
E_4 -10	327.729

In pitch discrimination research, Seashore (1919) and Henning and Grosberg (1968) asked participants to indicate pitch differences using only a lower/higher (flat/sharp) designation. Bentley (1966) and Rodman (1981) expanded the requirement using lower/same/higher. Other studies examining pitch discrimination have not specified between sharp and flat in their results

(Cramer & Zeitlin, 1955; Henning & Grosberg, 1968; Spiegel & Watson, 1984; Zeitlin, 1964). Rodman (1981) asked participants to determine whether the second tone of each presented listening pair was lower, the same, or higher than the first tone. Second tones were either not mistuned (0 cent deviation), or mistuned 2, 5, 10, or 15 cents sharp or flat. He reported high school musicians able to identify sharp and flat mistunings at the 5 cent level. Karrick (1998) arbitrarily [his word] chose 6 cents as a threshold for correct responses in his research while Madsen, Edmonson, and Madsen (1969) reported accurate participant discrimination ability at the 10 cent level. Results of these studies informed the decision to limit the span of cent deviations in the present study to 5, 7.5 and 10 cents.

Timbre

Electronically produced tones were used in this study because they are extremely stable compared to recorded tones of human performers, which contain slight pitch variations. These inconsistencies were judged to present an unintended distraction and recorded acoustic tones were therefore discarded in favor of much more stable electronically produced square wave and sawtooth wave tones.

In order to create optimal aural conditions for participants, complex electronic wave forms (square and sawtooth) were chosen as stimuli. More complex electronic wave forms such as square, sawtooth, triangular, and rectangular pulse have been used repeatedly in perception based research (Cassidy, 1989; Cramer & Zeitlin, 1955; Geringer, 1991; Henning & Grosberg, 1968; Rakowski, 1990; Sergeant, 1973; Spiegel & Watson, 1984; Zeitlin, 1964). Research has revealed that more complex wave forms elicit more accurate responses than sinusoidal wave forms when used as test stimuli (Henning & Grosberg, 1968; Sergeant, 1973; Spiegel & Watson, 1984). Also, the square wave is similar to clarinet in timbre, possessing only odd-numbered

harmonics, while the sawtooth wave possesses both even and odd harmonics and resembles the oboe in timbre.

Mode of Presentation

In this study participants heard pitch pairs in two conditions—sequential (reference tone, then test tone separated by one second of silence) and simultaneous (reference tone joined by a superimposed test tone). The decision to present tones sequentially and simultaneously was made in an effort to represent real-world conditions. Real-world pitch discrimination judgments are made by performers in both sequential and simultaneous settings. The act of performing in an ensemble often requires musicians to perform in unison with other performers.

In the simultaneous setting, each reference tone (B-flat$_4$ or E_4) sounds for two seconds and is then joined by a test tone of the same pitch (altered or unaltered), and in one of the two wave forms (square or sawtooth), for two more seconds. The simultaneous and sequential methods of presentation were organized into two separate test sections because I did not wish to add another level of difficulty to an already demanding perception task by inter-mixing the two.

The basic science research sources reviewed for this study, many of which were published in the *Journal of the Acoustical Society of America* (JASA), shared some common design characteristics. Much of this research was conducted using the paired comparison format and in many of these studies both reference and test tones were presented between 60 and 75 dB, and in lengths from 100 to 500 ms with 500 to 750 ms between presentations. Tones of such short duration are valuable in measuring participant perception in some experiments but they do not resemble the types of tone lengths regularly experienced by performing musicians. Also, the frequencies chosen as stimuli in several of the *JASA* studies are produced in multiples of 50 or

100, not frequencies normally associated with the equal tempered tuning system or instrumental performance.

Fastl and Zwicker (1999) presented two settings for detecting sound change. The first type is a tone that changes during its duration and is called a “modulation.” The second type, tone pairs separated by silence, is called “differences.” When “measuring just-noticeable differences, a pause is needed between the sounds to be compared” (p. 181). For pause durations ranging between 0.1s and 2.0s, “the results are independent of the duration of the pause” (p. 181). The one second of silence between tones is in agreement with Harris (1948) whose research indicated that the addition of a brief interval of silence between reference and test tones will effectively eliminate aural distractions such as clicks or “fuzziness” (p. 310).

Dependent Variable

The dependent variable for this study was pitch discrimination, that is, participants’ accuracy in identifying the second pitch in pairs of pitches as being sharp, flat or in-tune to a reference pitch. In order to examine pitch discrimination, I constructed one eight minute paired comparison sequential listening test and one eight minute paired comparison simultaneous listening test that participants completed while wearing headphones in single participant and small group settings. The listening tests account for all sound-based independent variables (pitch, cent deviation, timbre, and listening condition).

Development of Reference and Test Tones

Reference tones used in this study were B-flat$_4$ (466.165 Hz) and E$_4$ (329.628 Hz). Both the reference and the test tones were generated by the ChucK general-purpose real-time audio synthesis and graphics/multimedia programming language (Wang & Cook, 2007). Test tones were generated which varied -10, -7.5, -5, 0, +5, +7.5, and +10 cents from each reference tone.

The frequencies for each cent deviation level were obtained from the Tontechnik Rechner (Sengpiel, 2009) online program and are listed in Table 3.4 above. All frequency values were verified using the formula $b = a \times 2^{n/1200}$ (*a* = reference frequency, *b* = the new frequency, and *n* = cent deviation) (Backus, 1977, p. 349).

The two different timbres for the reference and test tones were created by designating the sawtooth and square wave forms within the ChucK code. The digital analog converter (DAC) then automatically created the specified tones in the specified frequencies. All tones were automatically generated by ChucK at the same amplitude (50% of maximum wave amplitude). Since volume levels have been found to slightly affect pitch perception (Fastl & Zwicker, 1999) they were kept constant across tone pairs in order to control for this effect. Playback volume was set at a level that produced 68 dB at each pair of headphones for each of the sequential tests and 72 dB for each of the simultaneous tests as measured by a decibel meter positioned even with the edge of each headphone's ear pads. When the second tone entered it boosted the volume 4 dB.

Low frequency tones sounding below all of the ChucK generated square and sawtooth tones were eliminated by using the equalization function included in the Audacity digital music editing software (Mazzoni, 2006). The dB levels of all frequencies below 300 Hz for the E_4 tones and 400 Hz for the B-flat$_4$ tones were lowered below -24 dB to render them inaudible and thereby eliminate them as potential distractions.

Development of Pitch Discrimination Test

Once generated, the test tones were paired with the reference tones using the Audacity software; each pair consisting of a reference tone followed by either its altered or unaltered test tone. Spoken number cues were recorded and inserted before each tone pair. The complete listening test consisted of two tasks presented in two sections. For the sequential presentation

section of the test each reference tone and its paired test tone were edited to be two seconds in duration with a break between each pair lasting one second. Much of the research reported in scientific journals dealing with pitch perception presents stimulus tones of much shorter durations (i.e. 10ms to 500ms). However, a few intonation discrimination studies do more closely approximate the present study (Bentley, 1966; Bradshaw & McHenry, 2005; Rodman, 1981). In those studies, with musicians as participants, the reference tones range from one to three seconds in length and are more aligned with actual tuning performance requirements.

For the simultaneous presentation section, each of the reference tones was generated to sound for four seconds with its paired test tone timed to enter after the first two seconds. This resulted in a two second presentation of the reference tone followed by two seconds with both tones sounding together. The same sound engineering process used to prepare the sequential stimuli was used to prepare the simultaneous stimuli.

The sequential task required the participants to listen to the 56 randomly ordered tone pairs. The first tone presented in each pair was always either the B-flat$_4$ (466.165) or E_4 (329.628) reference tone. Each test tone was either the same wave form (square or sawtooth) as its paired reference tone or the opposite, and was either the same frequency, or altered to sound 5, 7.5, or 10 cents above or below its reference. Both reference and test tones were two seconds long and were separated by one second of silence. Each pair of tones was separated by four seconds of silence. This section took approximately 12 minutes to complete.

The simultaneous task required the participants to listen to another 56 pairs of tones of either B-flat$_4$ or E_4. In this setting the reference tones were four seconds in length and each was joined, after sounding alone for two seconds, by a test tone. These pairs were also separated by four seconds of silence. The test tones were again either the same pitch and timbre as the

reference tones, or were altered in the same manner as those for the sequential task. This section took approximately 10 minutes to complete. The answer sheet (Appendix A) was used for both tasks and simply asked participants to circle whether the second tone was lower, the same, or higher than the first tone.

Counterbalancing techniques controlled for order of presentation by listening condition and tone pair. One half of the high school participants and one half of the college participants took the sequential listening test followed by the simultaneous test. The other half of the participants in both groups took the tests in the reverse order. Tone pairs in each section were organized using two different random sequences. One half of both the college and high school groups took the sequential test in random order one, while the other half took it in random order two. Likewise in the simultaneous condition participants were equally divided between random orders three and four. The entire process including data form completion and listening test took approximately thirty minutes.

The two random orders for the sequential test were developed by assigning each possible pitch pairing a number from 1 to 56 and then applying a random sequence which was generated by the online Random Sequence Generator (Haahr, 2009). Two random orders for the simultaneous test were likewise developed. Random orders are listed in Table 3.5 below.

Table 3.5
Table of Random Orders for Test Tones

Reference Tone	Test Tone	Original Order	Random 1 Sequential	Random 2 Sequential	Random 3 Simultaneous	Random 4 Simultaneous
B-flat 0 square	B-flat 0 square	01	12	14	50	49
B-flat 0 square	B-flat +7.5 square	02	47	50	27	34
B-flat 0 square	B-flat +10 square	03	15	21	54	17
B-flat 0 square	B-flat +5 square	04	11	31	24	24
B-flat 0 square	B-flat -5 square	05	29	20	52	11
B-flat 0 square	B-flat -10 square	06	23	48	21	13

(Table 3.5 continued)

B-flat 0 square	B-flat -7.5 square	07	50	49	48	33
E 0 square	E 0 square	08	24	36	1	8
E 0 square	E +7.5 square	09	32	23	33	56
E 0 square	E +10 square	10	49	33	43	12
E 0 square	E +5 square	11	18	32	51	30
E 0 square	E -5 square	12	25	19	9	10
E 0 square	E -10 square	13	08	04	10	44
E 0 square	E -7.5 square	14	35	06	38	53
B-flat 0 sawtooth	B-flat 0 sawtooth	15	40	45	39	39
B-flat 0 sawtooth	B-flat +7.5 sawtooth	16	03	13	40	2
B-flat 0 sawtooth	B-flat +10 sawtooth	17	46	46	16	25
B-flat 0 sawtooth	B-flat +5 sawtooth	18	20	43	22	28
B-flat 0 sawtooth	B-flat -5 sawtooth	19	48	55	15	7
B-flat 0 sawtooth	B-flat -10 sawtooth	20	10	51	32	14
B-flat 0 sawtooth	B-flat -7.5 sawtooth	21	44	12	42	19
E 0 sawtooth	E 0 sawtooth	22	21	41	49	18
E 0 sawtooth	E +7.5 sawtooth	23	13	01	37	43
E 0 sawtooth	E +10 sawtooth	24	19	29	30	6
E 0 sawtooth	E +5 sawtooth	25	26	47	26	48
E 0 sawtooth	E -5 sawtooth	26	41	38	46	42
E 0 sawtooth	E -10 sawtooth	27	01	44	23	22
E 0 sawtooth	E -7.5 sawtooth	28	22	09	31	5
E 0 square	E 0 sawtooth	29	42	28	12	51
E 0 square	E +7.5 sawtooth	30	38	18	44	41
E 0 square	E +10 sawtooth	31	37	53	2	46
E 0 square	E +5 sawtooth	32	30	37	19	37
E 0 square	E -5 sawtooth	33	52	34	36	4
E 0 square	E -10 sawtooth	34	04	42	29	38
E 0 square	E -7.5 sawtooth	35	55	16	13	15
B-flat 0 square	B-flat 0 sawtooth	36	07	27	11	50
B-flat 0 square	B-flat +7.5 sawtooth	37	17	02	3	16
B-flat 0 square	B-flat +10 sawtooth	38	02	03	20	40
B-flat 0 square	B-flat +5 sawtooth	39	28	17	6	54
B-flat 0 square	B-flat -5 sawtooth	40	33	10	55	55
B-flat 0 square	B-flat -10 sawtooth	41	51	08	53	21
B-flat 0 square	B-flat -7.5 sawtooth	42	53	56	17	31
E 0 sawtooth	E 0 square	43	06	40	45	3
E 0 sawtooth	E +7.5 square	44	34	26	14	9
E 0 sawtooth	E +10 square	45	27	05	28	1
E 0 sawtooth	E +5 square	46	45	35	56	32
E 0 sawtooth	E -5 square	47	54	25	4	52
E 0 sawtooth	E -10 square	48	05	24	34	26
E 0 sawtooth	E -7.5 square	49	14	39	41	35
B-flat 0 sawtooth	B-flat 0 square	50	09	11	35	45
B-flat 0 sawtooth	B-flat +7.5 square	51	39	07	5	20
B-flat 0 sawtooth	B-flat +10 square	52	16	54	47	47
B-flat 0 sawtooth	B-flat +5 square	53	43	52	25	36
B-flat 0 sawtooth	B-flat -5 square	54	31	15	8	27
B-flat 0 sawtooth	B-flat -10 square	55	36	30	18	23
B-flat 0 sawtooth	B-flat -7.5 square	56	56	22	7	29

Participants heard each altered test tone four times sequentially and four times simultaneously for a total of eight exposures. For example, the B-flat$_4$ test tone altered to 10 cents sharp was presented in the following pairings;

Sequential presentation orders
1. B-flat$_4$ square wave reference tone paired with B-flat$_4$ +10 cent square wave test tone
2. B-flat$_4$ square wave reference tone paired with B-flat$_4$ +10 cent sawtooth wave test tone
3. B-flat$_4$ sawtooth wave reference tone paired with B-flat$_4$ +10 cent sawtooth wave test tone
4. B-flat$_4$ sawtooth wave reference tone paired with B-flat$_4$ +10 cent square wave test tone

Simultaneous presentation orders
1. B-flat$_4$ square wave reference tone paired with B-flat$_4$ +10 cent square wave test tone
2. B-flat$_4$ square wave reference tone paired with B-flat$_4$ +10 cent sawtooth wave test tone
3. B-flat$_4$ sawtooth wave reference tone paired with B-flat$_4$ +10 cent sawtooth wave test tone
4. B-flat$_4$ sawtooth wave reference tone paired with B-flat$_4$ +10 cent square wave test tone

The four resulting sets of pairs were arranged in eight different orders in an attempt to control for order effect (Table 3.6). Participants were assigned to one of eight groups so that each participant listened to one of two sequential orders and one of two simultaneous orders.

Table 3.6
Test Administration Order Matrix

Group	First Test	Second Test
1	Random Order 1 Sequential	Random Order 3 Simultaneous
2	Random Order 1 Sequential	Random Order 4 Simultaneous
3	Random Order 3 Simultaneous	Random Order 1 Sequential
4	Random Order 3 Simultaneous	Random Order 2 Sequential
5	Random Order 2 Sequential	Random Order 4 Simultaneous
6	Random Order 2 Sequential	Random Order 3 Simultaneous
7	Random Order 4 Simultaneous	Random Order 2 Sequential
8	Random Order 4 Simultaneous	Random Order 1 Sequential

Sony MDR-V150 closed back supra-aural (ear pad) headphones with a dynamic frequency response of 18Hz - 22,000Hz and a sensitivity of 98 dB/mW were chosen instead of speakers to help focus participants' listening attention, control extraneous room noise, eliminate

potential room acoustic effects, and provide consistent volume levels and sound reproduction. Prior studies have used headphones with no reported negative effects due to their use (Cassidy, 1989; Duke, 1985; Karrick, 1998; Madsen, Edmonson & Madsen, 1969; Morrison, 2000).

Reference and test tones were presented via .wav files (24 bit, 48,000 Hz) through the iTunes version 10.1.2.17 music playback program. The iTunes program was run on a Toshiba Satellite A205 laptop computer with an Intel Core2 CPU operating at 1.73 GHz running the Windows Vista Home Premium version 6.0 32-bit operating system with 48 kHz sound output. Output ran through a Boostaroo Model T613-ENC 3-channel Headphone Audio Amplifier/Splitter (20 to 20,000Hz frequency response with a signal to noise ratio -95dBA from clipping) powering the Sony model MDR-V150 stereo headphones connected via Koss Headphone model Y88 Y-Cord Stereophone Splitters.

Pilot Testing

Four graduate students and one university faculty member took both the simultaneous and sequential portions of the test. Their responses led to several changes in audio quality, scripted directions, and test administration. Small audio anomalies were corrected by the researcher. The minute nature of the difference in tuning between reference and test tones and the need for intense concentration was further emphasized in the script. Pilot testers felt that the level of concentration required to adequately address all 56 pairs of tones might be excessive for some high school students. For this reason a one minute timed break was inserted at the half way point in each of the two sections in order to allow participants a moment to rest. Two practice excerpts with instructions were also added at the beginning of the process to allow participants the opportunity to become familiar with the testing process and adjust headphones.

Procedures

Exemption from institutional oversight was requested and granted (see Appendix A.) Permission letters were signed by the high school principal and band director and are included as Appendices B and C. High school participants (*n* = 64) were recruited with the help of the high school band director, and university participants (*n* = 64) with the help of the university marching band and concert ensemble directors. Prior to test administration parental consent forms were collected from all students under the age of eighteen (Appendix D), and adult participant consent form were signed by students over the age of eighteen (Appendix E). All participants were tested in November and December of 2010. All participants completed a data form which requested information concerning their major instrument, age, gender, grade, number of years playing major instrument, and number of years of private instruction on their major instrument (Appendix F).

Participants were then tested via a pre-recorded .wav file. Instructions were pre-recorded and were included in the test recording (Appendix G). Headphones were used to control potential room acoustic and background noise as well as volume levels. All excerpts were played on a laptop computer using a digital music playback program.

All answer sheets (Appendix H) were manually evaluated and 15% were re-evaluated to insure accuracy. Scores were calculated by assigning one point for each correct and 0 points for each incorrect response for each tone pair. The total number of points possible for each participant was 112 (56 points per half). Since two "same timbre" pairs and two "different timbre" pairs were constructed for each tone condition, means were calculated out of two possible points.

Reliability

Twenty-two percent (14) of the participants from each group, high school and college, were retested in January of 2011, and test-retest reliability scores were obtained by dividing the total number of agreements by the sum of agreements plus disagreements. This produced a college reliability score of $R = .64$ and a high school reliability score of $R = .69$, combining for an overall test-retest reliability of $R = .67$. $R = .70$ is generally interpreted as acceptable reliability (Wells & Wollack, 2003). For the present study, reliability therefore approaches acceptable. Although test items were stable between test and retest, the subjective nature of human aural discrimination, the requirement that participants make very fine discriminations, and the time elapsed between test and re-test (roughly 2 months) may have increased inconsistency of response. There is no way to discount guessing in cases where participants were uncertain as to how to respond. It is worth noting that among reliability participants, college correct response scores improved from 50% on the test to 55% on the re-test. High school scores improved slightly from 45% on the test to 46% on the re-test.

CHAPTER 4
RESULTS

The main purpose of this study was to investigate high school and college wind instrumentalists' pitch discrimination when judging pitch pairs separated by 0, 5, 7.5, and 10 cents. A listening test comprised two sections with 56 tone pairs in one section presented sequentially and 56 tone pairs in the other presented simultaneously. Each pair consisted of a reference tone followed by a test tone of the same pitch (B-flat$_4$ or E$_4$), which was either identical to the reference tone, or changed on the basis of cent deviation (0, 5, 7.5, or 10 cents) and/or timbre (square or sawtooth wave). Tone pairs were ordered using four different computer-generated random sequences; two for the sequential section and two for the simultaneous section. The four resulting sets of pairs were further arranged in eight different orders in an attempt to control for order effect (refer to Table 3.6). Participants were assigned to one of eight groups so that each participant listened to one of two sequential orders and one of two simultaneous orders. Participants listened to each tone pair through headphones and circled on an answer sheet whether the test tones were lower, the same, or higher than the paired reference tones.

In the following analyses, means were calculated out of 2. In the listening test, each permutation of pitch, timbre, cent deviation, and presentation mode occurred twice. In other words, participants had two opportunities to respond to any one combination of variables. Two was the smallest cell size, with correct responses being represented by 0, 1, and 2.

In order to determine effect of order of presentation, a One-Way ANOVA was calculated and revealed no significant difference in accuracy among the eight participant groups [F (7,120) = .580, p = .77]. Therefore, for subsequent analysis all participants were pooled together regardless of order.

A Four-Way ANOVA with repeated measures (2 timbres x 2 presentations x 2 pitches x 7 cent deviations) was calculated, and the results (*p* < .05) are presented in Table 4.1.

Table 4.1
Four-Way ANOVA with Repeated Measures for Pitch, Timbre, Presentation, and Cent Deviation

Source	DF	Sum of Squares	Mean Square	F Value	P Value	Partial η^2
Pitch	1	17.88	17.88	44.51	< .0001	.35
Error	127	51.01	.40			
Timbre	1	37.72	37.72	48.14	< .0001	.38
Error	127	99.53	.78			
Presentation	1	12.39	12.39	14.39	.0002	.11
Error	127	109.36	.86			
Cent deviation	6	97.84	16.31	25.93	< .0001	.20
Error	762	479.16	.63			
Pitch x Timbre	1	2.81	2.81	7.88	.0058	.06
Error	127	45.37	.36			
Pitch x Presentation	1	1.29	1.29	3.49	.0641	.03
Error	127	46.82	.37			
Pitch x Cent Deviation	6	234.90	39.15	84.16	< .0001	.66
Error	762	354.46	.47			
Timbre x Presentation	1	4.03	4.03	7.47	.0072	.06
Error	127	68.58	.54			
Timbre x Cent Deviation	6	87.70	14.62	31.26	< .0001	.25
Error	762	356.31	.47			
Presentation x Cent Deviation	6	29.95	4.99	9.52	< .0001	.07
Error	762	399.55	.52			
Pitch x Timbre x Presentation	1	1.40	1.40	3.31	.0713	.03
Error	127	53.57	.42			
Pitch x Timbre x Cent Deviation	6	7.64	1.27	3.23	.0039	.03
Error	762	300.44	.39			
Pitch x Presentation x Cent Deviation	6	38.09	6.35	6.12	< .0001	.13
Error	762	300.05	.39			
Timbre x Presentation x Cent Deviation	6	78.65	13.11	30.64	< .0001	.24
Error	762	326.00	.43			
Pitch x Timbre x Presentation x Cent Deviation	6	8.82	1.47	4.05	.0005	.03
Error	762	276.46	.36			

All four main effects were statistically significant. A significant difference due to the main effect of pitch was found [*F*(1, 127) = 44.51, *p* < .0001], with participants making more accurate intonation responses for the B-flat pitch pairs (*M* = .95, *SD* = .80) than for the E pitch pairs (*M* = .85, *SD* = .74). A significant difference due to the main effect of timbre was also found [*F*(1,

127) = 48.14, *p* < .0001], with participants making more accurate intonation responses for different-timbre tone pairs (*M* = .97, *SD* = .75) than for same-timbre tone pairs (*M* = .83, *SD* = .78). The main effect of presentation [*F*(1, 127) = 14.30, *p* = .0002] revealed another significant difference, with participants making more accurate intonation responses for simultaneously presented tone pairs (*M* = .94, *SD* = .77) than for the sequentially presented tone pairs (*M* = .86, *SD* = .76). Finally, a significant difference due to the main effect of cent deviation was found [*F*(6, 762) = 25.93, *p* < .0001], with test tones mistuned by both 10 cents and 7.5 cents (sharp and flat) being more accurately identified than those mistuned by 5 cents or not mistuned at all (Figure 4.1). Interestingly, the in-tune test tones (*M* = .70, *SD* = .76) were least accurately identified, being correctly labeled an average of only 35% of the time.

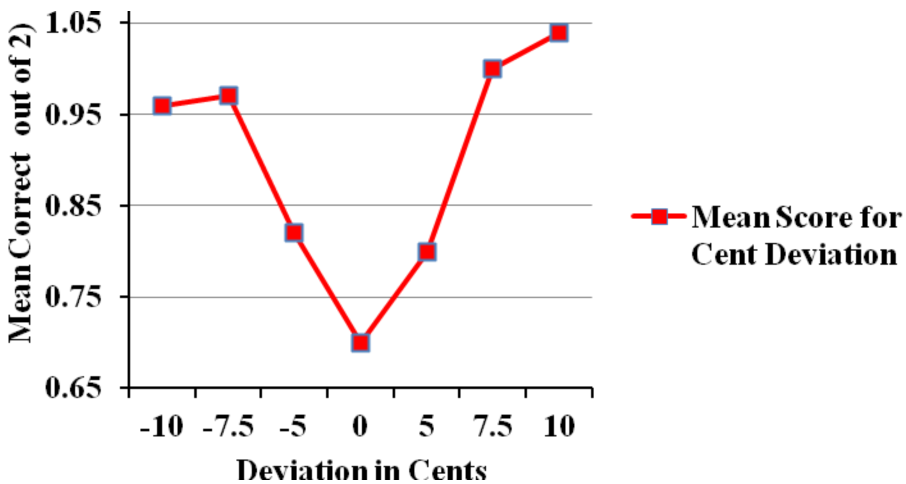

Figure 4.1 Mean Scores for Cent Deviation

Post hoc analysis involving the Scheffé test revealed significant differences in 13 of 21 cent deviation pairs (*p* < .0001). These data, presented in Table 4.2, can be inferred from Figure 4.1. At the +/- 10 cent and +/- 7.5 cent deviation levels, responses are more accurate and more alike, but significantly different than all other cent deviations levels. At the +5, 0, and -5 levels, the only comparison that is significantly different is the 0 to -5 pairing.

Table 4.2
Scheffé Post-Hoc Analysis of the Effect of Cent Deviation (Probabilities)

+10							X
+7.5						X	= .9816
+5					X	**< .0001**	**< .0001**
0				X	= .1558	**< .0001**	**< .0001**
-5			X	**= .0312**	= .9984	**< .0001**	**< .0001**
-7.5		X	**= .0033**	**< .0001**	**= .0002**	= .9884	= .6724
-10	X	> .9999	**= .0091**	**< .0001**	**= .0008**	= .9550	= .5079
	-10	**-7.5**	**-5**	**0**	**+5**	**+7.5**	**+10**

All 2-, 3-, and 4-way interactions were significant ($p < .05$) except pitch x presentation ($p = .06$) and pitch x timbre x presentation ($p = .07$). A few of the lower level interactions provided interesting information. For example, the two-way interaction between pitch and cent deviation is displayed in Figure 4.2. The intonation of E_4 was more accurately perceived at the -10, -7.5, and -5 cent levels while the intonation of B-$flat_4$ was more accurately perceived at the +5, +7.5, and +10 cent levels of mistuning. Both pitches were almost identically perceived when the test tone pitch was unaltered from its paired reference. Accurate response means between B-$flat_4$ and E_4 were separated by more distance at the sharp mistunings than at the flat mistunings.

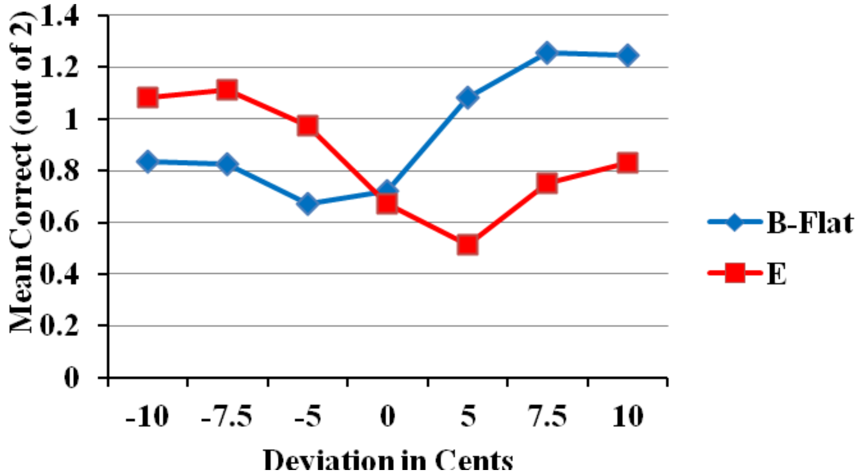

Figure 4.2 Interaction Between Pitch and Cent Deviation

The two-way interaction between timbre and cent deviation is displayed in Figure 4.3. Here participants were more accurate in identifying intonation errors when the pairs were different timbres except in the no pitch change condition where the average different-timbre score for no-change was at least .4 points lower than all other different timbre averages. Mean scores between same and different timbres were closer on the sharp mistunings than on the flat mistunings.

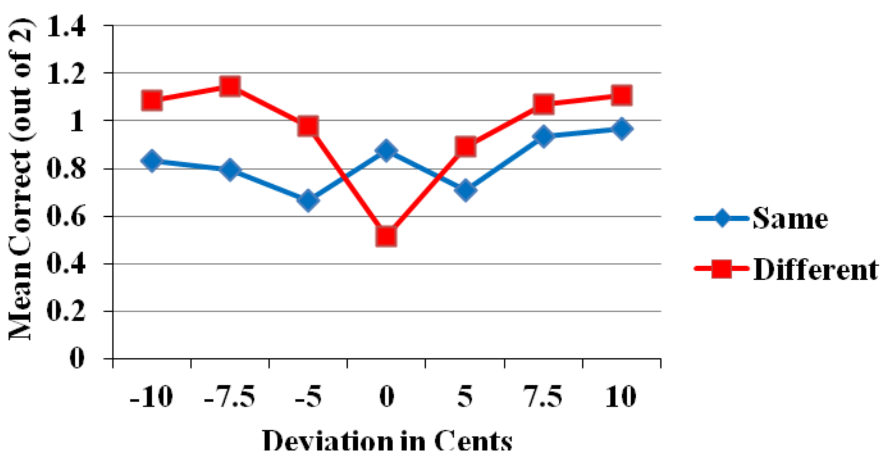

Figure 4.3 Interaction Between Timbre and Cent Deviation

The two-way interaction between reference pitch and timbre is displayed in Figure 4.4. The responses to B-flat$_4$/same timbre (M = .895, SD = .823) and E$_4$/same timbre (M = .756, SD = .729) were both lower than responses to B-flat$_4$/different timbre (M = 1.00, SD = .762) and E$_4$/different timbre (M = .940, SD = .732). However, the difference between correct response means for the same and different E$_4$ reference pitches was greater than that for the B-flat$_4$ reference pitches.

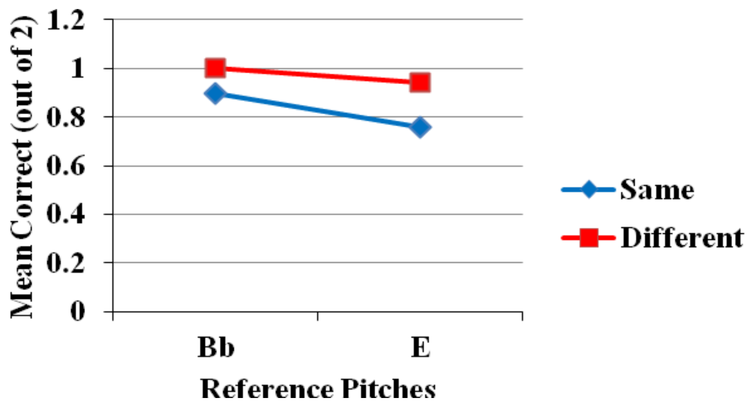

Figure 4.4 Interaction Between B-flat and E Reference Pitches with Same and Different Timbres

All of these interactions are subsumed within a larger four-way interaction presented in Figures 4.5 and 4.6. With the B-flat$_4$ pitch pairs (Figure 4.5), the interaction involving the same and different timbre and sequential and simultaneous presentation is quite pronounced, with the same timbre/sequential accuracy responses being highest at the no-change cent deviation and lowest at most other cent deviations compared to the other three timbre/presentation conditions. Conversely, the different timbre/sequential presentation mean score, at the no-change cent deviation, drops below the other three while scoring above them for all flat mistunings. The same timbre/sequential mean at the no-change level is over one point (50%) above the different timbre/sequential mean. The same-simultaneous and different-simultaneous means present

similarly shaped curves with both presentations showing higher mean scores for the two sharpest mistuned test tones than the same-sequential and different-sequential means.

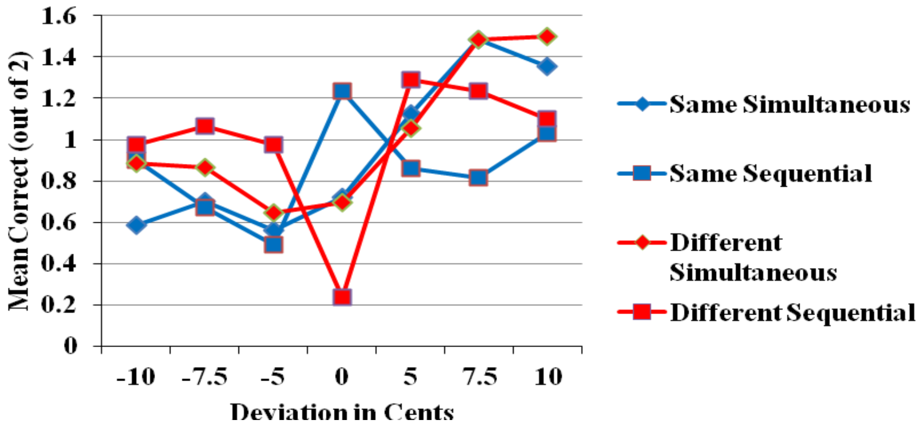

Figure 4.5 Interaction Among B-flat Pitch and Timbre and Presentation and Cent Deviation

These same trends are apparent at the E_4 pitch level (Figure 4.6); however they are less pronounced than at the B-flat$_4$ level. Another interesting difference between the B-flat$_4$ and E_4 graphs is that where B-flat$_4$ mean scores trend better for sharp mistunings, the E_4 mean scores trend better for flat mistunings. Similar to the B-flat$_4$ graph, the same timbre/sequential presentation no-change mean moves up while the E_4 different timbre/ sequential no-change mean once again moves down.

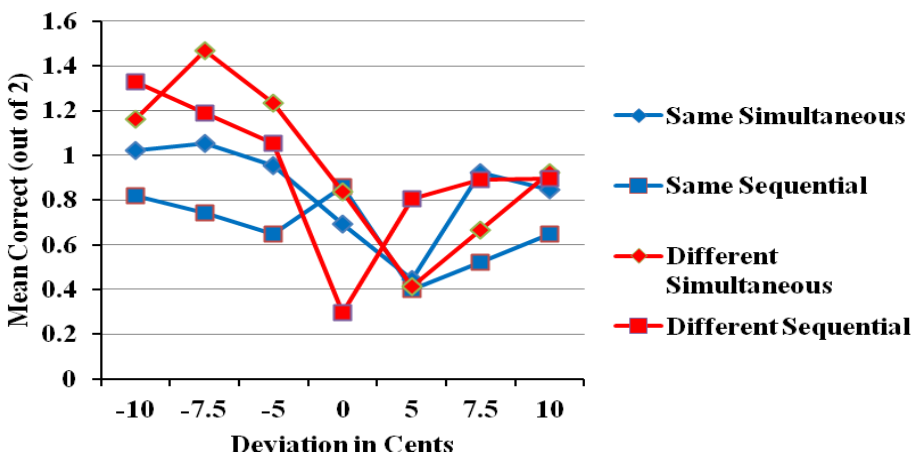

Figure 4.6 Interaction Among E Pitch and Timbre and Presentation and Cent Deviation

The overall shape of the graphs showing the four way interactions in Figure 4.5 and Figure 4.6, not surprisingly, are similar to the shapes outlining the two-way interactions (Figure 4.2, 4.3, and 4.4). The difference at the 0 change cent deviation level displayed in timbre and cent deviation interaction (Table 4.3) is magnified with the addition of presentation in the four-way interactions, yet another indication of the power of context to affect listener response.

In order to further look at the importance of the significant differences, effect size was measured by calculating the partial η^2 statistic. The largest value, partial $\eta^2 = .66$, indicated a strong effect in the pitch by cent deviation interaction, explaining 66% of the variance among the data points. The main effects of timbre and pitch accounted for 38% and 35% respectively, both moderate effects. The remaining comparisons produced partial eta square values of less than .25.

Examination of descriptive statistics revealed that the high school participants consistently performed less accurately than the university participants. The high school group performed with an overall average of 41% correct responses. University participants were more accurate, with an overall average of 49% correct responses. Average percent correct scores for each level of mistuning for each group are presented in Table 4.3.

Table 4.3
Levels of Mistuning: Percent Correct and Ranges

Group	-10	-7.5	-5	0	+5	+7.5	+10
H.S. Percent[a]	42	42	37	30	39	47	47
H.S. Range[b]	2-12	1-15	2-12	0-12	2-15	2-12	3-12
Univ. Percent	54	55	45	40	41	53	57
Univ. Range	3-15	3-14	2-14	0-14	2-14	4-16	3-16

[a] Percent refers to percentage of correct responses
[b] Ranges reflect a minimum possible score of 0 and a maximum possible of 16

On average, high school participants were slightly more accurate when identifying sharp mistunings (44% correct) than flat (41% correct), and correctly identified no-change pairs only

30% of the time. University students were more accurate than high school students on both sides of 0 and almost identical in response to sharp (50% correct) and flat stimuli (51% correct). At the no-change level, they had 40% correct.

Knowledge gained from pilot testing suggested that a listening test comprising 56 tone pairs was both long and demanding enough to potentially cause listener fatigue. Therefore a break was introduced at the half-way point for each presentation setting. Average scores were calculated for pre-break and post-break responses, and the results are presented in Table 4.4. The small decrease in total mean correct response from pre-break to post-break indicates that fatigue was probably not a factor in how listeners responded, although this possibility cannot be ruled out. Interestingly, on the sequential presentation test half, both the university and high school participants' average scores improved by two percentage points. Conversely, on the simultaneous presentation test half, both groups' average scores dropped—by five points for high school and four points for university participants.

Table 4.4
Pre-break versus Post-break Percent Correct

Group	Pre-break	Post-break
High School Sequential	37	39
University Sequential	47	49
High School Simultaneous	46	41
University Simultaneous	52	48
Totals	45.50	44.25

CHAPTER 5
DISCUSSION

This research was motivated by a desire to define a pitch discrimination threshold such that what musicians interpret as being in-tune could be known with some level of certainty. A more definitive answer as to what is perceived as in-tune would be helpful to teachers, conductors, and researchers. Should we consider the definition of in-tune to be when two pitches vary by 0 cents—the literal, non-contextual, non-musical definition— or is there a range of variation around 0 cents that constitutes "in-tune"? Assuming such a range (an assumption that is apparent both anecdotally and in the research literature), what is it?

I designed a listening experience in human perception, not performance, which was ecologically valid in several ways. The variables of pitch (B-flat$_4$ and E$_4$), timbre (square and sawtooth wave forms), mode of presentation (sequential or simultaneous), and cent deviation (plus and minus 10, 7.5, and 5, or 0) were included in order to incorporate into the study some of the elements present in live music making. Results show that my desire to find a clear answer to the threshold question is complicated greatly by aural context. Rather than arriving at a definitive answer, it is apparent that the answer depends on these contextual factors. Although this may sound like a hedge, the detailed nature of this study and its results does provide some clarity with regard to how musicians hear.

Response to Pitch

Results indicate that participants correctly identified pitch discrepancies for B-flat$_4$ pitch pairs (47% correct) significantly better than for E$_4$ pitch pairs (42% correct). This was not surprising because, given its familiarity as the tuning note of choice for bands, musicians may respond to B-flat differently than less "familiar" notes. Many student musicians practice tuning to B-flat frequently, and they often receive feedback as to how accurately they are responding. It

is also the fundamental pitch of most of the brass instruments used in the wind band setting and the tonic pitch of a frequently used key in band music. Morrison's (2000) performance based study found that middle school and high school band students tuned significantly more accurately to a B-flat tuning note than to a series of Gs, which were presented in melodic contexts. Morrison reported low, positive correlations between the B-flat tuning pitch and the four Gs, but a high positive correlation among the four Gs. For student musicians, within-melody context did not matter but other contexts did matter (B-flat vs. G and tuning vs. melodic performance). Morrison's results may explain in part the use of B-flat as the common tuning note for bands, but they raise questions about the effect of B-flat tuning beyond B-flat itself.

It is interesting to note, however, that very little of the research reviewed for the present study included B-flat as a reference pitch. The exceptions were Byo, Schlegel & Clark (2011), Morrison (2000), and Yarbrough, Karrick, & Morrison (1995). Only the research of Byo, Schlegel, and Clark (2011) used B-flat exclusively. In most of the studies examined, researchers opted to use even numbered frequencies and their multiples or avoid B-flat altogether and use other pitches.

Response to Timbre

Research has shown that listeners and performers tend to hear bright timbres as being sharp and dark timbres as flat (Geringer & Worthy, 1999; Wapnick & Freeman, 1980; Worthy, 2000). The content of the wave form itself (i.e. its overtones) was shown to matter in the research of Platt and Racine (1985). They found that subjects tuned better to complex versus sine tones. Ely (1992) concluded that subjects were best at detecting discrepancies when listening to tones with many overtones. Essentially, these results indicate that responses to pitch are affected by timbre.

Likewise in the present study, musicians' responses were affected by timbre. Participants identified pitch discrepancies in different-timbre tone pairs (49% correct) significantly better than they did in same-timbre tone pairs (42% correct). This is supported by Ely (1992) who reported that woodwind players were "significantly better at detecting intonational deviations in duets involving unlike timbral combinations than they were in like timbral combinations" (p. 164). In a study investigating the effect of timbre and octave on the tuning ability of college and high school wind instrumentalists, Byo, Schlegel, and Clark (2011) found that high school musicians tuned similarly to three different timbres (flute, oboe, and clarinet), suggesting that tuning decisions were driven by lower octave displacement of the reference, not timbre. In contrast, Greer (1970) found that brass players were most accurate when tuning to a reference pitch produced by their own instrument. Similarly, Byo (1993), in research on error detection, found that listeners detected flawed performance more effectively in like-timbred settings. Overall, considering this small research base concerning response to timbre, the results are mixed. Notably in the present study, greater accuracy in different timbre pitch pairs was not the case at the no-change level, where same-timbre responses were more accurate. It may be that participants are more accurate detecting larger intonation differences between different timbre pairs, but more accurate at detecting smaller differences in like-timbre pairs.

Response to Mode of Presentation

Simultaneously and sequentially presented tone pairs produced significantly different results. Simultaneously paired reference and test tones were correctly identified 47% of the time, while sequentially presented pairs were correctly identified 43% of the time. By group, university participants made 47% correct choices, while high school participants made 43% correct choices. In speech and hearing research, there has been a wealth of research comparing

responses to sequential (e.g., Darwin, Hukin, & Al-Khatib, 1995) and simultaneous (e.g., McCabe & Denham, 1997) stimuli. Separately, research in music perception and music education has examined sequential presentation of reference and test tones (Henning & Grosberg, 1968; Spiegel & Watson, 1984) and simultaneous presentation pitch matching (e.g., Cassidy, 1989; Hayes, 2009), polyphonic performance (e.g., Geringer, 1991; Greer, 1970), interval tuning (e.g., Brittin, 1993; Karrick, 1998; Rakowski, 1990), and error detection (e.g., Byo, 1993, 1997; Byo & Sheldon, 2000; Schlegel, 2010). The review of literature revealed no research in music in which the effect of sequential and simultaneous presentation was isolated in one study. The results of the present study are not sufficient to suggest that sequential pitch comparison has no place in research or pedagogy.

Although no research was found to support or refute the sequential mode of presentation, it was conjectured that the silence between reference and test tones might have made more apparent pitch discrepancy and/or the direction of pitch discrepancy. Certainly, the sequential mode employed in "down the line" fashion is an often-used tuning procedure in the large ensemble setting. Sundberg and Lindqvist (1973) argued that out-of-tune simultaneously presented pairs should be easily identified due to the presence of beats. Several participants in the present study made unsolicited comments to this effect, that hearing beats served as a cue that the two tones were out of tune. The presence of beats, however, does not help the listener determine the direction of the mistuning. It may also be that not all listeners attend to beats when making pitch decisions. Due to the lack of empirical data, the best course of action at present may be to incorporate both sequential and simultaneous tuning activities into the band rehearsal.

Response to Cent Deviation

Participant responses to cent deviation were found to be significantly different, with the 10 cent and 7.5 cent mistuned pairs in both the sharp and flat direction more accurately identified than 5 cent and 0 cent pairs. This is the most direct answer to the research question, Do participants demonstrate differences in their ability to accurately identify mistunings among cent deviation levels? Yes, they are significantly more accurate detecting pitch discrepancies at the +/- 10 and +/- 7.5 levels than at the +/- 5 and 0 or no-change levels. Notably, their best efforts are 57% correct for university participants and 47% correct for high school participants.

Research has produced results that vary widely regarding the aural acuity of listeners who vary in age and musical experience. For elementary through college aged wind players, Madsen, Edmonson, and Madsen (1969) found an overall threshold of 10 cents. For university musicians, Parker (1983) found accurate identification of mistunings at 20 cents. For high school wind players, Rodman (1983) reported a more discerning discrimination level of 5 cents. For "an unselected group of adults" Seashore (1938, p. 56) reported accurate identification of mistunings at 12 cents. Results from the current study do not identify a specific point, measured in cent deviation from a reference, where participants may be said to accurately detect the smallest pitch change. At the 10 and 7.5 cent levels, the overall correct responses ranged from 42% to 57%. However, these responses were significantly better than responses to the 5 cent and 0 cent levels, where the percent correct range dipped to 30% to 45%.

A less than 50% accuracy rate overall at the 5, 7.5, and 10 cent levels of mistuning appears to be consistent with Madsen, Edmonson, & Madsen's (1969) assertion that aural perception of pitch changes begins at around 10 cents for elementary aged and older participants. However, their results were derived from research conducted in one context. They administered

a listening test with each stimulus presented for 30 seconds in a continuously modulating format. The low correct response rates for the high school participants in the present study (37% at -5, and 39% at +5) are not consistent with Rodman's finding that participants in this age group accurately identified pitch differences at the 5 cent level. Rodman's results derive from yet another context. As in the present study he asked participants to judge the second tone in tuned and mistuned pairs as lower, the same, or higher than its paired stimulus, but he used the 12 pitches from the C chromatic scale spread over a span of five octaves. Both studies examined participants' abilities to discriminate slight differences between tones in pairs at different levels of mistuning. But where the present study presented the different levels of mistuning in random order, Rodman began his test, like Bentley (1966) and Seashore (1938), with larger mistunings (15 cents in this case) and then systematically reduced the levels of mistuning (15-10-5-2) as the test progressed. Both of these differences (octave and random order) may have contributed to the different results reported by the two studies.

Results from the present study support the assertion that context plays a very important role in pitch perception. When comparing the results of the present study to the literature reviewed, it is apparent that contexts are different in every case. Parker (1983) presented participants with one second long tones in pairs, with the second tone mistuned from 0 to 100 cents in 10 cent increments. Geringer (1983) and Geringer and Witt (1985) combined both performance and perception in their research. Miles (1972) conducted intonation training sessions that combined both instrumental performance and the use of an electronic Intonation Trainer. These differences in context make cross-study comparisons difficult and may contribute to the different levels of perception reported—from Rodman's reported 5 cent discrimination level for high school musicians to Parker's reported 20 cent level for university student

musicians. Much of the variation in responses within and between studies may very well have been affected by variation in musical context—and the present study only scraped the surface of musical context.

Data in the present study indicate that musicians within a college or high school band are not a homogeneous group relative to aural acuity (around 40% in both groups hear well at +/- 5 cents; 60% do not). For example, in order to decide on a threshold of 20 cents, some of Parker's participants would have scored better than 20 cents correct and some would have scored worse. The idea of discovering one number as the point of demarcation between what musicians can and cannot hear no longer seems germane. Logically, if it were possible to find a pitch perception threshold, it would not be one threshold. Data from the present study suggest that there may be multiple thresholds—depending on aural and musical contexts.

In the present study participants had to listen and make quick decisions. In the sequential setting each tone lasted for two seconds with a one second break in between. In the simultaneous setting the reference tone sounded alone for two seconds before being joined for two more seconds by the test tone. In both settings participants had four seconds between the pairs in which to decide whether the test tones were lower, the same as, or higher than their paired references. The length of time participants have to make decisions about pitch may be an important factor affecting accuracy of perception and performance. Are split-second decisions advantageous to considered decisions? In real music making, many tuning and intonation decisions have to be made quickly. The length of time participants have been given to decide, when comparing pitches in perception tasks, has varied in the literature from a few milliseconds to a few seconds. Research into the *adaptive unconscious* (Gladwell, 2005) suggests that humans are capable of making accurate judgments in many different situations in very short

amounts of time without consciously considering possible outcomes. In many music performance settings musicians also have very little time to perceive, decide, and make adjustments regarding their pitch levels.

Numerous studies (Geringer, 1978; Geringer & Witt, 1985; Madsen, 1966; Madsen, Edmonson, & Madsen, 1969; Morrison, 2000; Yarbrough, Morrison, & Karrick, 1997) reported that participants erred more in the direction of sharpness, or put another way, they identified flat better than sharp. This was also true in the present perception-based study when participants responded to E_4. However, it was not true with their responses to B-$flat_4$. Participants responded in a manner consistent with previous research when the stimulus was the unfamiliar E, but were inconsistent when the stimulus was the familiar B-flat.

Also in the present study, as with Madsen, Edmonson and Madsen (1969), the in-tune test tones were the least accurately identified. Combined high school and university averages revealed more "same" responses missed (35% correct) than either "higher" (47% correct) or "lower" (46% correct). Poor response to the 0 cent deviation pairs might be explained by the fact that, given the nature of the task, participants were primed to hear differences even when none actually existed. The fact that participants had one-half the number of chances in “same” stimuli than they did in “different” (higher and lower) stimuli may shed some light on this seemingly low rate of accurate selection. When considering a test protocol with three possible responses for each question, chance alone (guessing) would suggest an approximate correct selection rate of 33%. During informal post-test discussions several subjects made unsolicited comments regarding a lack of confidence, and that they guessed when unsure of the correct response. Since only the ten lowest scoring participants (9%) scored 36 or fewer correct responses out of 112 (32% correct), and the remaining 118 subjects scored higher than 33% correct, it can be argued

that 91% of participants used cognitive strategies rather than simply guessing. Guessing, however, cannot be ruled out.

Interactions

More important than the significant main effects were the significant interactions that, considered in total, indicate that pitch discrimination may be as much about other factors as it is about pitch. The main effect showed that pitch (B-flat or E) matters because listeners responded differently, that is, better to B-flat and worse to E. The pitch by cent deviation interaction, which produced the largest effect size of all (partial $\eta^2 = .66$), provided evidence that what matters goes even deeper. Participants' within-pitch responses to B-flat varied depending on the direction of mistuning. They responded more accurately to B-flat when it was sharp than when it was flat. Conversely, they responded to E more accurately when it was flat than when it was sharp. It is conjecture, but perhaps this shows musicians responding as expected to an "unfamiliar" pitch and not as expected to a "familiar" pitch. Importantly, there was yet another different aural reaction to B-flat and E at the 0 cent deviation level; responses to each were nearly the same.

In the timbre by cent deviation interaction the effect size (partial $\eta^2 = .25$) indicates this interaction was responsible for 25% of the variance among data points. The different timbre pitch pairs were correctly identified at a higher rate than were the same timbre pairs, except at the 0 cent deviation level where there was a starkly different response pattern. Pedagogically, it is logical to think it necessary to establish the sound of in-tune in one's ear as a precursor to judging out-of-tune. In other words, "I know what in-tune sounds like. This chord doesn't sound in-tune, so it must be out-of-tune." One might expect responses to in-tune stimuli to be more accurate than to out-of-tune stimuli. Not so in this interaction which shows musicians' responses to different timbre in-tune stimuli to be the least accurate of all cent deviation

responses (only 26% correct). The closest out-of tune response (same timbre, -5 cents) was seven percentage points better (33% correct). Responses to same timbre in-tune were also better (44% correct), but not as accurate as responses to +7.5 (47% correct) and +10 (49% correct). Perhaps this indicates that one place to begin to better equip student musicians aurally is to make a concerted effort to help them detect the sound of in-tune. To set them up to succeed, these data indicate that this process might best start with same-timbre stimuli.

In the two-way interaction between pitch and cent deviation at the 0 cent deviation level participant responses were virtually identical. When timbre was introduced as a factor at the 0 cent level a spike occurred for same timbre and a dip occurred for different. When presentation mode was introduced with B-flat pitch, a larger spike for same and a larger dip for different occurred. This effect was mirrored with E, but not to the same degree. Essentially, as the context changed at the no-change level, the results changed. These differences raise an important question, Is the aural challenge different under conditions of in-tune versus conditions of out-of-tune? In pedagogy, perhaps it is worth considering starting with out-of tune and moving toward in-tune. Maybe the ear needs the advantage provided by comparison—of the sound of out-of-tune and the change in contrast as it moves toward in-tune.

The main effect of timbre showed that in pitch comparisons involving same or different timbres, timbre makes a difference. In this instance the participants responded better to different timbre pairs. The timbre and pitch interaction also suggests that deeper, more complex relationships exist between factors. For both timbre conditions, participants were more accurate identifying mistuned B-flat pairs than E pairs. Also, participants' responses to same/different timbre were more alike on B-flat pairs than they were on E pairs. As discussed above, participants' familiarity with B-flat may explain the better scores.

The four-way interactions among B-flat pitch, timbre, presentation mode, and cent deviation, and among E pitch, timbre, presentation mode, and cent deviation reflect their component two-way interactions. The relatively small effect size (partial $\eta^2 = .03$) would suggest that these interactions are not very important. However, they do reveal some interesting comparisons. The addition of mode of presentation (sequential or simultaneous) in this comparison reveals the same spike and dip at 0 cent deviation for both E and B-flat that appears in Figure 4.3 where the two pitches are combined. In the four-way interaction, the magnitude of the spike for same timbre/sequential mode and the dip for different timbre/sequential mode is far greater than that shown in the two-way interaction and for B-flat than for E. As with the interaction between pitch and cent deviation, the flat cent deviations were generally more accurate on E, and the sharp cent deviations generally more correct on B-flat.

It may be beneficial for researchers, teachers, and conductors to think about aural response as a reaction to a multiple component sound condition rather than pitch, timbre, and presentation mode as separate items. An overview of the contents of conducting, score study, and rehearsal texts (Battisti & Garofalo, 1990; Green, 1961; Kohut, 1973; McBeth, 1972; Neidig, 1964; Rudolph, 1950; Weerts, 1976) revealed a lack of attention given to both tuning and musical context (e.g. instrumentation, scoring, register, dynamics, etc.).

Goolsby's (1997) study of the use of verbal instruction by band directors ($N = 30$) in a series of videotaped rehearsals revealed a similar lack of focus on intonation. He found that expert, novice, and student teachers addressed a number of different areas, with all three groups spending the most time on rhythm. Intonation was the ninth most frequently addressed variable by expert teachers ($M = 10.6$ times across 20 rehearsals) behind rhythm (27.5), articulation

(21.4), expression (16.7), listening (13), tone (11.8), style (11.3), airstream (11), and dynamics (11).

A somewhat startling contrast is presented in Colprit's (2000) finding that string teachers stopped most often (11.5% of the time) to address intonation problems. In the string world where pitch can be anywhere the performer puts finger to string, there is much more attention paid to intonation. In the band world where a "roughly correct" pitch often occurs with correct fingering, intonation receives much less attention. If the expectation is high-quality performance in band, perhaps band conductors should heed the intonation practices of string teachers/conductors (at least in the time spent category).

Results from the current study advise against simplistic approaches to tuning/intonation and reveal a need for teachers/conductors to approach tuning and intonation in varied ways. Perhaps many band directors approach tuning and intonation with little concern for the context that comprises the aural experience. Building "sound" awareness on a comprehensive level, involving all of the facets of sound, and using the music played and the flawed performance of student musicians to expose intonation challenges, demands a teacher/conductor who is aware of the complexities involved in decision-making about sound.

For example, band teacher/conductors might work to develop lesson plans that address the many components that comprise musical context. They might choose to slow down musical passages with a focus on listening to specific target pitches. In such instances it might be advantageous to direct attention to one specific player or one specific section. They might identify a model pitch and have the performers sustain (simultaneous condition) while bending their ears toward that model. In the band rehearsal setting the texture or dynamics of a certain passage may make it impossible for all members of the ensemble to hear the model. In such

cases a sequential approach, working with one section or one player at a time, might be a better choice. As Morrison's (2000) research revealed, musicians may tune notes in a melodic context differently than they tune a single note. Teacher/conductors may decide to begin with single notes and chords, and then gradually move towards the intonation of specific notes within melodic contexts. Other rehearsal considerations such as doublings with instruments of different timbres, matching pitch with instruments in different octaves, and the intonation tendencies of different notes on different instruments are all conditions that require careful consideration and planning.

Age and Experience

Not surprisingly, university musicians were found to perform with more accurate pitch discrimination in every area measured. The combined average score for correct responses (sharp, flat, and in-tune) for the university students was 49% while the high school participants averaged 41%. These results are consistent with previous research (e.g., Geringer & Worthy, 1999; Madsen, Edmonson, & Madsen, 1969; Platt & Racine, 1985), which revealed a tendency for older, more experienced participants to discriminate pitch differences more accurately than younger, less experienced participants.

University participants responded correctly 50% of the time for sharp stimuli and 51% for flat stimuli, compared to high school participants who only responded 44% correct for sharp and 41% correct for flat. Both groups performed least accurately to the in-tune pairs, with university participants averaging 40% correct and high school students averaging only 30% correct. The difference in the sharp and flat scores of the university participants is almost identical, and the difference between high school sharp and flat responses is slight.

In the present study the high school group consistently performed below the level of the university participants at the seven levels of tuning and mistuning tested (Table 4.6). This finding is in agreement with results from prior studies in that it reports different levels of pitch discrimination for different age groups. Of the top 25 performers, only two were high school students. Of the bottom 25 performers, 16 were high school students.

Fatigue

Participants in the present study's pilot testing process reported that the length of the listening test, coupled with the degree of concentration required, was tiring. Therefore a one-minute break was inserted at the mid points of both the simultaneous and sequential halves of the test in an attempt to lessen the potential effect of listener fatigue. In order to examine the possible effect of fatigue, average correct responses were calculated for the test halves. A slight drop in combined average scores between the first half (45.5% correct) and second half (44.25% correct) was too small to suggest that fatigue was a factor in participants' performance (Table 4.4). Essentially both groups of participants performed slightly better after the break in the sequential listening portion (42% pre-break vs. 44% post-break), but not as well after the break in the simultaneous portion (49% pre-break vs. 45% post-break). It is possible that the simultaneous task was more demanding and therefore resulted in the lower post-break average score. Or, perhaps participants learned as they progressed through the earlier portion of the sequential listening task resulting in improved performance later in that test.

Conclusions

Conclusions are organized according to research question.

1. Do participants demonstrate differences in their ability to accurately identify mistunings among cent deviation levels (plus and minus 10, plus and minus 7.5, plus or minus 5, and 0)? If so, what is the nature of these differences?

Yes. The present study found significant main effects and significant interactions, indicating that there are differences in participants' ability to accurately identify mistunings. The highest correct response rates occurred at the 7.5 and 10 cent mistuning levels; the lowest at the no-change level. University students averaged higher scores than did high school students at every level of mistuning and were correct 49% of the time. High school students were correct 41% of the time.

When considering the group data provided by this large number of student musicians in conjunction with Parker's (1983) research, it is accurate to say that accuracy rates of 50% and up occur between the 10 and 20 cent deviation levels. Also, just over 40% of university musicians and just under 40% of high school musicians were accurate discriminating at the 5 cent deviation level. Given these figures, it is clear that neither a high school band nor a university band is a homogeneous group as regards pitch discrimination; within the same band, there is a sizable group who fail to accurately detect 10 cent deviations and a sizable group who accurately detect 5 cent deviations. Though it may be tempting to focus on percent correct responses that struggle to get to 50%, it is important to realize that these data provide a look at how student musicians heard in a limited-time-to-decide perception task across 128 test items that varied in relevant contextual factors.

In keeping with previous research results, a logical hypothesis would have been that participants would be more accurate detecting flat mistunings than sharp. In the current study this only occurred with one of two reference pitches. It may be that participants are better at detecting sharp mistunings with some pitches and flat mistunings with others. But it is important to remember that this was a perception task (there was no performance element), and it was conducted with specific variables of pitch, timbre, listening condition, and degrees of mistuning.

2. Does timbre (same and different) affect participants' accurate identification of mistuned tones?

Yes. Participants were significantly more accurate identifying mistuned different-timbre tone pairs (49% correct) than they were identifying same-timbre tone pairs (42% correct). At the no-change level, participants' greater accuracy under same-timbre conditions may be indicative of more accurate perception of larger intonation differences in different timbre conditions and more accurate perception of smaller intonation differences in like-timbre conditions.

One important and unexpected finding of this research revealed participants to be least accurate when judging pitches in the 0 cent deviation condition. It has seemed logical to assume that hearing in tune is the same as hearing out of tune, except that in tune is the 0 cent deviation form of out of tune. This may not be the case. It may be that hearing in tune and out of tune are two different processes and the participants in this study were better at the out of tune process (no matter how much cent deviation) than they were at the in tune process.

3. Does pitch (B-flat$_4$ and E$_4$) affect participants' accurate identification of mistuned tones?

Yes. Participants' responses for all levels of mistuning (both sharp and flat) were significantly more accurate for the B-flat pitch pairs (48% correct) than for the E pitch pairs (43% correct)‒listeners responded differently to two different pitches. Perhaps more than

anything else, these results suggest that pedagogy in intonation and tuning in the band setting must extend beyond a narrow, beginning of rehearsal, one-note tuning process. Morrison's (2000) lack of correlation between mass tuning (static B-flats) and melodic playing (Gs within a melody), combined with the results from the current study, suggest a need for band conductors to approach tuning and intonation in more than one way. Simply put, band mass tuning procedures could benefit from the use of more than one pitch and one approach.

4. Does listening condition (simultaneous and sequential) affect participants' accurate identification of mistuned tones?

Yes. Participants were more accurate identifying mistunings in simultaneously presented tone pairs (47% correct) than in sequentially presented tone pairs (43% correct). The presence of beats in the simultaneously presented pitch pairs may have contributed to the higher level of accuracy in the simultaneous condition. The timed nature of the test may have also affected performance. Four seconds between tone pairs only gave participants limited time in which to make decisions in both the sequential and simultaneous settings.

5. Do age and experience affect participants' accurate identification of mistuned tones?

Perhaps. The descriptive statistics derived from the current study trend in the direction of numerous previous studies (e.g., Duke, 1985; Madsen, 1974; Madsen, Edmonson, & Madsen, 1969), and suggest that the age and experience levels of participants makes a difference with the older, more experienced university participants outperforming the high school group in every task measured. Factors that were not variables in the present study such as age and years of private instruction may have also affected participants' perception accuracy.

Future Research

The timed nature of the tasks in the present study required participants to make immediate choices and thus removed contemplation from the equation. In pitch matching studies the participants are usually afforded much more time to match reference tones (e.g., Byo, Schlegel, & Clark, 2011; Miles, 1973) and thus more time to decide whether they are sharp, in tune, or flat. The act of matching a tone while performing affords participants the opportunity to compare and adjust their responses over a longer span of time than that afforded in this study and many of the perception based listening studies reviewed. In fact Miles (1973) found his beginning band subjects learned to perform beat-free intervals over time during practice sessions in a simultaneous setting.

Designing listening tests that incorporate different lengths of time between tone pairs may better inform our understanding of the decision making process regarding intonation. Studies specifically designed to compare results from different length reference and test tones might also be valuable in this regard. Research of this type could incorporate recorded examples of wind instruments as stimuli since the sequential presentation of tones does not require the perfect tone reproduction necessary for simultaneous settings.

A comparison of the effects of simultaneous and sequential settings on pitch perception is another area that could benefit from continued research. Much of the intonation research to-date has been conducted in a sequential format. The simultaneous setting has primarily been used in interval production or melody and pitch matching tasks. Research designed to isolate the effect of the sequential and simultaneous presentation modes on different levels of mistuning might inform conductors' decisions as to which mode would be most effective for use during the tuning process in rehearsals. Currently there is little data available on this topic.

Results from intonation perception tasks incorporating simultaneous presentation of stimuli may reveal effects caused by participant fatigue. Listening to stimulus and test tones presented concurrently may require greater concentration and therefore be more tiring. Research into the effect of different test lengths may also provide valuable information in this regard.

The presence of beats may be a signal to listeners that an intonation problem exists, but beats do not inform listeners of direction of mistuning. One of the possible disadvantages of sequential presentation of stimuli is that beats are not present. It is also important, however, to consider that all listeners may not be responding to beats. Some performers may possess a keener sense of pitch and be capable of perceiving out of tune pitch pairs without needing to attend to beats.

Because the square wave stimuli in the present study only possessed odd numbered harmonics, there were fewer harmonics present and therefore fewer opportunities for harmonics to interact when test tones were slightly mistuned. Research limited to one wave form possessing both even and odd harmonics (e.g. sawtooth), in both tuned and mistuned conditions, presented in sequential and simultaneous settings using only one reference pitch may provide more specific and more ecologically valid information in this area.

Another line of research might incorporate a pretest-posttest format built around a training regimen designed to improve participants' perception and/or performance abilities through practice over a period of weeks or months. Training sessions involving repeated listening sessions designed to gradually move participants from significantly mistuned to very slightly mistuned pairs of pitches over a period of weeks or months might prove beneficial. I found that my ability to accurately identify mistunings at all three levels tested improved throughout the process of building, pilot testing, and administering the listening tests. I took the

test on several occasions and was able to improve my performance over a span of several months due to repeated exposure to the tuned and mistuned pitch pairs. I experienced the most improvement at identifying test tones which had been mistuned at the five cent level.

REFERENCES

Backus, J. (1977). The acoustical foundations of music (2nd ed.). New York: W. W. Norton & Company.

Ballard, D. (2006). Relationships between college level wind instrumentalists' achievement in intonation perception and performance (Doctoral dissertation, Indiana University). *ProQuest Information and Learning Company,* 3298145*.*

Battisti, F. (1993). List listening targets for students. In J. Casey (Ed.), *Teaching techniques and insights for instrumental music educators*. Chicago, IL: GIA Publications, Inc.

Battisti, F. (1998). Frank Battisti. In J. E. Williamson & K. L. Neidig (Eds.), *Rehearsing the band* (pp. 1-10). Cloudcroft, NM: DBA Neidig Services.

Battisti, F. & Garofalo, R. (1990). *Guide to score study for the wind band conductor*. Ft. Lauderdale, FL: Meredith Music Publications.

Bentley, A. (1966). *Musical ability in children and its measurement*. New York: October House, Inc.

Bernstein, J. & Oxenham, A. (2003). Pitch discrimination of diotic and dichotic tone complexes. *Journal of the Acoustical Society of America, 113*(6)*,* 3323-3334.

Bradshaw, E. & M. McHenry. (2005). Pitch discrimination and pitch matching abilities of adults who sing inaccurately. *Journal of Voice*, *19*(3), 431-39.

Brittin, R. V. (1993). Effects of upper and lower-register accompaniment on intonation. *Journal of Band Research, 29*(1)*,* 43-50.

Burns, E. M. & Ward, W. D. (1978). Categorical perception - phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, *63*(2)*,* 456-468.

Byo, J. L. (1993). The influence of textural and timbral factors on the ability of music majors to detect performance errors. *Journal of Research in Music Education*, *41*(2), 156-167.

Byo, J. L. (1997). The effects of texture and number of parts on the ability of music majors to detect performance errors. *Journal of Research in Music Education, 45*(1), 51-66.

Byo, J. L., Schlegel, A. L., & Clark, N. A. (2011). Effects of stimulus octave and timbre on the tuning accuracy of secondary school instrumentalists. *Journal of Research in Music Education, 58*(4) 316–328.

Byo, J. L. & Sheldon, D. A. (2000). The effect of singing while listening on undergraduate music majors' ability to detect pitch and rhythm errors. *Journal of Band Research, 36*(1), 26-46.

Cassidy, J. W. (1989). The effect of instrument type, stimulus timbre, and stimulus octave placement on tuning accuracy. *Missouri Journal of Research in Music Education, 26*, 7-23.

Cohen, A. (1961). Further effects of intensity upon the pitch of pure tones. *Journal of the Acoustical Society of America, 33*, 1363-1376.

Colprit, E. J. (2000). Observation and analysis of suzuki string teaching. *Journal of Research in Music Education, 48*(3), 206-221.

Cramer, R. & Zeitlin, L. (1955). Frequency discrimination of pure and complex tones. Report; Psychology Department of the Army Medical Research Laboratory, Project No. 6-95-20-001. Fort Knox, Kentucky.

Darwin, C. J., Hukin, R. & Al-Khatib, B. (1995). Grouping in pitch perception: Evidence for sequential constraints. *Journal of the Acoustical Society of America, 98*(2), 880-885.

Delezenne, M. (1826). Mémoire sur les valeurs numériques des notes de la gamme. *Recueil des travaux de la Société des Sciences de Lille,* 1-71.

Duke, R. A. (1985). Wind instrumentalists' intonational performance of selected musical intervals. *Journal of Research in Music Education, 33*(2), 101-111.

Duke, R. A., Geringer, J. M., & Madsen, C. K. (1988). Effect of tempo on pitch perception. *Journal of Research in Music Education, 36*(2), 108-125.

Ely, M. C. (1992). Effects of timbre on college woodwind players' intonational performance and perception. *Journal of Research in Music Education*, *40*(2), 158-167.

Fastl, H. & Zwicker, E. (1999). *Psychoacoustics: Facts and models* (2nd ed.). New York: Springer-Verlag. http://www.springerlink.com.libezp.lib.lsu.edu/content/u60r2w/

Fechner, G.T. (1860). *Elemente der psychophysik* [Elements of psychophysics]. Leipzig: Breitkopf und Härtel.

Fyk, J. (1985). Vocal pitch-matching ability in children as a function of sound duration. *Bulletin of the Council for Research in Music Education, 85*, 76-89.

Fyk, J. (1987). Duration of tones required for satisfactory precision of pitch matching. *Bulletin of the Council for Research in Music Education*, *91*, 38-44.

Fyk, J. (1995). *Melodic intonation, psychoacoustics, and the violin*. Zielona Góra: Organon.

Geringer, J. M. (1976). Tuning preferences in recorded orchestral music. *Journal of Research in Music Education, 24*(4), 169-176.

Geringer, J. M. (1978). Intonational performance and perception of ascending scales. *Journal of Research in Music Education, 26*(1), 32-40.

Geringer, J. M. (1983). The relationship of pitch-matching and pitch-discrimination abilities of preschool and fourth-grade students. *Journal of Research in Music Education, 31*(2), 93-99.

Geringer, J. M. (1991). Temporal discrimination of modulated intensity in music excerpts and tones. *Journal of Research in Music Education*, *39*(2), 113-120.

Geringer, J. M. (1995). Continuous loudness judgments of dynamics in recorded music excerpts. *Journal of Research in Music Education*, *43*(1), 22-35.

Geringer, J. M., & Witt, A. C. (1985). An investigation of tuning performance and perception of string instrumentalists. *Bulletin of the Council for Research in Music Education, 85,* 90-101.

Geringer, J. M., & Worthy, M. D. (1999). Effects of tone-quality changes on intonation and tone-quality ratings of high school and college instrumentalists. *Journal of Research in Music Education, 47*(2), 135-149.

Gladwell, M. (2005). *Blink*. New York: Little, Brown and Company.

Goolsby, T. W. (1997). Verbal instruction in instrumental rehearsals: a comparison of three career levels and preservice teachers. *Journal of Research in Music Education, 45*(1), 21-40.

Green, E. (1961). *The modern conductor*. Englewood Cliffs, NJ: Prentice-Hall.

Greer, R. D. (1970). The effect of timbre on brass-wind intonation, In E. Gordon (Ed.), *Experimental research in the psychology of music* (pp. 65-94). Iowa City, IA: University of Iowa Press.

Haahr, M. (2009). Random sequence generator. Retrieved January 26, 2010 from www.random.org/sequences.

Hall, D. E., & Hess, J. T. (1984). Perception of musical interval tuning. *Music Perception*, *2,* 166-195.

Harris, D. J. (1948). Discrimination of pitch: suggestions toward method and procedure. *The American Journal of Psychology, 61*(3), 309-32.

Hartmann, W. M., McAdams, S., & Smith, B. K. (1990). Hearing a mistuned harmonic in an otherwise periodic complex tone. *Journal of the Acoustical Society of America*, *88,* 1712–1724.

Hayes, T. E. (2009). Achievement of middle school wind instrumentalists in performance and perception of intonation. (Doctoral dissertation, Indiana University). *ProQuest Information and Learning Company,* 3358917.

Henning, F. & Grosberg, S. (1968). Effects of harmonic components on frequency discrimination. *Journal of the Acoustical Society of America, 44*, 1386-1389.

Karrick, B. (1998). An examination of the intonation tendencies of wind instrumentalists based on their performance of selected harmonic musical intervals. *Journal of Research in Music Education, 46*(1), 112-127.

Kennedy, M. & Kennedy, J. (2007). *The concise oxford dictionary of music*. New York: Oxford University Press.

Kohut, D. L. (1973). *Instrumental music pedagogy*. Englewood Cliffs, NJ: Prentice-Hall.

Lehman, A. C., Sloboda, J. A., & Woody, R. H. (2007). *Psychology for musicians: Understanding and acquiring the skills*. Oxford: Oxford University Press.

Lindley, M. (2001a). Equal Temperament. In S. Sadie & J. Tyrrell (Eds.), *The new grove dictionary of music and musicians, 8* (pp. 275-276). New York: Grove's Dictionaries, Inc.

Lindley, M. (2001b). Just Intonation. In S. Sadie & J. Tyrrell (Eds.), *The new grove dictionary of music and musicians, 13* (pp. 290-295). New York: Grove's Dictionaries, Inc.

Lindley, M. (2001c). Tuning. In S. Sadie & J. Tyrrell (Eds.), *The new grove dictionary of music and musicians, 25* (pp. 884-885). New York: Grove's Dictionaries, Inc.

Loosen, F. (1993). Intonation of solo violin performance with reference to equally tempered, Pythagorean, and just intonation. *Journal of the Acoustical Society of America, 93*(1), 525–539.

Madsen, C. K. (1966). The effect of scale direction on pitch acuity in solo vocal performance. *Journal of Research in Music Education, 14*(4), 266-275.

Madsen, C. K. (1974). Sharpness and flatness in scalar vocal performance. *Science de l'Art: Scientific Aesthetics, 9*, 91–97.

Madsen, C. K., Edmonson, F. A., & Madsen, C. K. (1969). Modulated frequency discrimination in relationship to age and musical training. *Journal of the Acoustical Society of America, 46*(6), 1468-1472.

Madsen, C. K., & Geringer, J. M. (1976). Preferences for trumpet tone quality versus intonation. *Bulletin of the Council for Research in Music Education, 46*, 13-22.

Madsen, C. K. & Geringer, J. M. (1981). Discrimination between tone quality and intonation in unaccompanied flute/oboe duets. *Journal of Research in Music Education, 29*, 305-313.

Madsen, C. K. & Madsen, C. H. (1997). *Experimental research in music* (3rd ed.). Raleigh, NC: Contemporary Publishing Company.

Mason, J. A. (1960). Comparison of solo and ensemble performances with reference to Pythagorean, just, and equi-tempered intonations. *Journal of Research in Music Education, 8*/1, 31-38.

Mazzoni, D. (2006). Audacity (Version 1.2.5): Retrieved October 15, 2009, from http://audacity.sourceforge.net/

McBeth, F. (1972). *Effective performance of band music*. San Antonio, TX: Southern Music Company.

McCabe, S. & Denham, M. (1997). A model of auditory streaming. *Journal of the Acoustical Society of America, 101*, 1611–1621.

McMurray, A. (1998). Allan McMurray. In J. E. Williamson & K. L. Neidig (Eds.), *Rehearsing the band* (pp. 1-10). Cloudcroft, NM: DBA Neidig Services.

Miles, E. M. (1972). Beat elimination as a means of teaching intonation to beginning wind instrumentalists. *Journal of Research in Music Education*, *20*(4), 496-500.

Morrison, S. (2000). Effect of melodic context, tuning behaviors, and experience on the intonation accuracy of wind players. *Journal of Research in Music Education*, *48*(1), 39-51.

Morrison, S. & Fyk, J. (2002). Intonation. In R. Parncutt & G. McPherson (Eds.), *The science and psychology of music performance creative strategies for teaching and learning* (pp. 183-197). London: Oxford University Press.

Neidig, K. (1964). *The band director's guide*. Englewood Cliffs, NJ: Prentice-Hall.

Nichols, R. (1947). Listening: questions and problems. *Quarterly Journal of Speech 33*(1), 83-86.

Olson, H. F. (1967). *Music, physics and engineering*. New York: Dover Publications.

Parker, 0. (1983). Quantitative differences in frequency perceptions by violinists, pianists, and trombonists. *Bulletin of the Council for Research in Music Education, 76*, 49-58.

Pickler, A. (1966). History of experiments on the musical interval sense. *Journal of Music Theory*, *10*(1), 54-95.

Platt, J. R., & Racine, R. J. (1985). Effect of frequency, timbre, experience, and feedback on musical tuning skills. *Perception and Psychophysics, 38,* 543–553.

Plomp, R. (1964). The ear as a frequency analyzer. *Journal of the Acoustical Society of America, 36*(9), 1628–1636.

Plomp, R., & Mimpen, A. M. (1968). The ear as a frequency analyzer, II. *Journal of the Acoustical Society of America, 43,* 764–767.

Price, H. E. A proposed glossary for use in affective response literature in music. *Journal of Research in Music Education, 34*(3), 151-159.

Rakowski, A. (1990). Intonation variants of musical intervals in isolation and in musical contexts. *Psychology of Music*, *18*, 60-72.

Rodman, M. E. (1981). *Range distribution, timbre, and cent calibration as factors in pitch discrimination capacities*. Unpublished doctoral dissertation retrieved January 15, 2010, from Texas Tech University: http://etd.lib.ttu.edu/theses/available/etd-03252009-31295001544831/

Rudolph, M. (1950). *The grammar of conducting*. New York: Schirmer Books.

Schellenberg, E. G. (2002). Asymmetries in the discrimination of musical intervals. *Music Perception*, *19*(2), 223–248.

Schlegel, A. (2010). The effect of directed attention score study procedures on music majors' error detection in three-part instrumental music. (Doctoral dissertation, Louisiana State University). *Louisiana State University Electronic Theses and Dissertations*.

Schubert, E., Wolfe, J. & Tarnopolsky, A. (2004). *Spectral centroid and timbre in complex, multiple instrumental textures*. Paper presented at the eighth International Conference on Music Perception and Cognition, Evanston, IL.

Seashore, C. (1919). *Manual of instructions and interpretations for measures of musical talent*. New York: Columbia Graphophone Company.

Seashore, C. (1938). *Psychology of music*. New York: McGraw-Hill.

Sengpiel, E. (2009). *Tontechnik rechner*. Mikrofonaufnahmetechnik und tonstudiotechnik. [Computer software] http://www.sengpielaudio.com/calculator-centsratio.htm

Sergeant, D. (1973). Measurement of pitch discrimination. *Journal of Research in Music Education, 21*, 3-19.

Shower, E. & Biddulph, R. (1931). Differential pitch sensitivity of the ear. *Journal of the Acoustical Society of America, 3,* 275-287.

Siegel, J. A. & Siegel, W. (1977). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception and Psychophysics, 21*(5), 399-407.

Sogin, D. W. (1989). An analysis of string instrumentalists' performed intonational adjustments within an ascending and descending pitch set. *Journal of Research in Music Education, 37*(2), 104-111.

Spiegel, M. F., & Watson, C. S. (1984). Performance on frequency-discrimination tasks by musicians and non-musicians. *Journal of the Acoustical Society of America, 76*, 1690–1695.

Stevens, S., Volkmann, J, & Newman, E. (1937). A scale for the measurement of the psychological magnitude of pitch. *Journal of the Acoustical Society of America, 8*, 185-190.

Sundberg, J. E. & Lindqvist, J. (1973). Musical octaves and pitch. *Journal of the Acoustical Society of America 54,* 922-928.

Swaffield, W. R. (1974). Effect of melodic parameters on ability to make fine-tuning responses in context. *Journal of Research in Music Education, 22*(4), 305-312.

Thompson, W. F., & Schellenberg, E. G. (2002). Cognitive constraints on music learning. In R. Colwell & C. Richardson (Eds.), *The new handbook of research on music teaching and learning* (pp. 461–486). New York: Oxford University Press.

Truax, B. (Ed.). (1999). *Handbook for acoustic ecology* (2nd ed.). Accessed from internet at [http://www.sfu.ca/sonic-studio/handbook/Critical_Band.html]

Vorce, F. (1964). The effect of simultaneous stimulus on vocal pitch accuracy (Doctoral dissertation, The Florida State University). *Dissertation Abstracts International, 25,* 5327A.

Vos, J. (1988). Subjective acceptability of various regular twelve-tone tuning systems in two-part musical fragments. *Journal of the Acoustical Society of America, 83,* 2383–2392.

Wagner, M. (2009). *Introductory musical acoustics* (4th ed.). Raleigh, NC: Contemporary Publishing Company of Raleigh.

Wang, G. & Cook, P. (2007). Wang, G. & Cook, P. (2007). ChucK [Computer software and manual]. Retrieved October 15, 2009, from http://chuck.cs.princeton.edu

Wapnick, J., & Freeman, P. (1980). Effects of dark-bright timbral variation on the perception of flatness and sharpness. *Journal of Research in Music Education*, *28*(3), 176-184.

Weber, E.H. (1834). De pulsu, resorptione, auditu et tactu: Annotationes anatomicae et physiologicae (On stimulation, response, hearing and touch: Notes on anatomy and physiology). Leipzig: Koehler.

Weerts, R. (1976). *Handbook of rehearsal techniques for the high school band*. West Nyack, NY: Parker Publishing Company.

Wells, C. & Wollack, J. (2003). *An instructors' guide to understanding test reliability*. Madison, WI: Testing & Evaluation Services, University of Wisconsin.

Wever, E. & Bray, C. (1930). The nature of acoustic response: the relation between sound frequency and frequency of impulses in the auditory nerve. *Journal of Experimental Psychology, 13*(5), 373-387.

Wever, E. & Bray, C. (1936). The nature of acoustic response: the relation between sound intensity and the magnitude of responses in the cochlea. *Journal of Experimental Psychology, 19*(2), 129-143.

Worthy, M. D. (2000). Effects of tone-quality conditions on perception and performance of pitch among selected wind instrumentalists. *Journal of Research in Music Education, 48a*(3), 222-236.

Yarbrough, C., Karrick, B., & Morrison, S. J. (1995). Effect of knowledge of directional mistunings on the tuning accuracy of beginning and intermediate wind players. *Journal of Research in Music Education, 43*(3), 232-241.

Yarbrough, C., Morrison, S. J., & Karrick, B. (1997). The effect of experience, private instruction, and knowledge of directional mistunings on the tuning performance and perception of high school wind players. *Bulletin of the Council for Research in Music Education, 134*, 31–42.

Young, W. (1973). The Bentley measures of musical abilities: a congruent validity report. *Journal of Research in Music Education, 21*(1), 74-79.

Zanette, D. H. (2008). Zipf's law and the creation of musical context, *Musicae Scientiae, 10*, 3-18.

Zeitlin, L. (1964). Frequency discrimination of pure and complex tones. *Journal of the Acoustical Society of America 36*, 207-209.

APPENDIX A
IRB REQUEST FOR EXEMPTION

Application for Exemption from Institutional Oversight

LSU
Institutional Review Board
Dr. Robert Mathews, Chair
131 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.6792
irb@lsu.edu
lsu.edu/irb

Unless qualified as meeting the specific criteria for exemption from Institutional Review Board (IRB) oversight, ALL LSU research/ projects using living humans as subjects, or samples, or data obtained from humans, directly or indirectly, with or without their consent, must be approved or exempted in advance by the LSU IRB. This Form helps the PI determine if a project may be exempted, and is used to request an exemption.

-- Applicant,Please fill out the application in its entirety and include the completed application as well as parts A-E, listed below, when submitting to the IRB. Once the application is completed, please submit two copies of the completed application to the IRB Office or to a member of the Human Subjects Screening Committee. Members of this committee can be found at http://www.lsu.edu/screeningmembers.shtml

-- A Complete Application Includes All of the Following:
- **(A)** Two copies of this completed form and two copies of part B thru E.
- **(B)** A brief project description (adequate to evaluate risks to subjects and to explain your responses to Parts 1&2)
- **(C)** Copies of all instruments to be used.
  *If this proposal is part of a grant proposal, include a copy of the proposal and all recruitment material.
- **(D)** The consent form that you will use in the study (see part 3 for more information.)
- **(E)** Certificate of Completion of Human Subjects Protection Training for all personnel involved in the project, including students who are involved with testing or handling data, unless already on file with the IRB. Training link: (http://phrp.nihtaining.com/users/login.php.)
- **(F)** IRB Security of Data Agreement: (http://www.lsu.edu/irb/IRB%20Security%20of%20Data.pdf)

1) Principal Investigator: Norman Alan Clark **Rank:** Graduate Student

Dept: Music Education **Ph:** 225-270-7722 **E-mail:** clarkn@tigers.lsu.edu

2) Co Investigator(s): please include department, rank, phone and e-mail for each

James L. Byo, Carl Prince Matthies Memorial Professor of Music Education
Department of Music Education, School of Music
225-578-2593, jbyo@lsu.edu

IRB# E5220 LSU Proposal #______
(X) Complete Application
(X) Human Subjects Training

3) Project Title: Direction of Mistuning, Range of Cent Deviation, and Timbre as Factors in Musicians' Pitch Discrimination in Simultaneous and Sequential Listening Conditions

Study Exempted By:
Dr. Robert C. Mathews, Chairman
Institutional Review Board
Louisiana State University
203 B-1 David Boyd Hall
225-578-8692 | www.lsu.edu/irb
Exemption Expires: 10-13-2013

4) Proposal? (yes or no) no **If Yes, LSU Proposal Number** ______

Also, if YES, either ◯ **This application completely matches the scope of work in the grant**
OR ◯ **More IRB Applications will be filed later**

5) Subject pool (e.g. Psychology students) High school and university band students

*Circle any **"vulnerable populations" to be used**: (children <18; the mentally impaired, pregnant women, the ages, other). Projects with incarcerated persons cannot be exempted.

6) PI Signature N. A. Clark **Date** 30 Sep 2010 (no per signatures)

**** I certify my responses are accurate and complete.** If the project scope or design is later changes, I will resubmit for review. I will obtain written approval from the Authorized Representative of all non-LSU institutions in which the study is conducted. I also understand that it is my responsibility to maintain copies of all consent forms at LSU for three years after completion of the study. If I leave LSU before that time the consent forms should be preserved in the Departmental Office.

Screening Committee Action: Exempted ✓ Not Exempted _____ Category/Paragraph 1

Reviewer Mathews **Signature** Robert C Mathews **Date** 10/14/10

APPENDIX B
SCHOOL ADMINISTRATOR CONSENT LETTER

Traci McCorkle
Principal
Lauren Avery
Associate Principal

East Ascension High School
612 East Worthey Road
Gonzales, LA 70737
(225) 391-6100

Carli Francois
Randall Loving
Walter Traveler
Assistant Principals

November 17, 2010

To Whom It May Concern;

Mr. N. Alan Clark is hereby approved to conduct his Ph.D. research study, *Direction of Mistuning, Range of Cent Deviation, and Timbre as Factors in Musicians' Pitch Discrimination in Simultaneous and Sequential Listening Conditions* with the band program here at East Ascension High School. I look forward to reviewing the results of his study.

Sincerely,

Traci McCorkle

Mrs. Traci McCorkle
Principal
East Ascension High School

APPENDIX C
TEACHER CONSENT FORM

TEACHER CONSENT FORM

1. Title: Direction of Mistuning, Range of Cent Deviation, and Timbre as Factors in Musicians' Pitch Discrimination in Simultaneous and Sequential Listening Conditions

2. Performance Sites: Louisiana State University, Parkview Baptist High School, and East Ascension High School

3. Contacts: The following investigators are available for questions about this project:
N. Alan Clark, principal investigator, Louisiana State University
clarkn@tigers.lsu.edu
225-270-7722

Dr. James L. Byo, Carl Prince Matthies Professor of Music Education, faculty advisor
jbyo@lsu.edu
225-578-2593

4. Purpose of Study: This inquiry is designed to answer the following primary questions: At what point do instrumentalists begin to discern minute changes in pitch? Is there a difference between participants' ability to detect sharp and flat pitches? Does the stimulus wave form affect perception of pitch change? Do participants discern pitch changes better in a sequential format, or a simultaneous format, or does it matter? Secondarily, the study will attempt to answer the following: Does age affect pitch perception accuracy? Does experience affect pitch perception accuracy? Does instrument family affect pitch perception accuracy?

5. Participants: Participants will be students enrolled in the band programs at Parkview Baptist High School, Baton Rouge, Louisiana; East Ascension High School, Gonzales, Louisiana; and the Louisiana State University Band program.

6. Number of Participants: 120 (Sixty from the two high schools and sixty from LSU)

7. Study Procedures: Participants will complete a data inform in which they will indicate their age, years of private study, and gender. Participants under the age of 18 will obtain parent/guardian approval prior to testing. Participants 18 or older will give informed consent before the test begins. Following the procurement of this information, participants listen to a series of pre-recorded tones and indicate whether each pair of tones is the same, or how it is different by circling "lower, same, higher" on an answer sheet.

8. Benefits: There will likely be no immediate benefit to participants. Potentially, study results will provide evidence regarding subjects' ability to detect intonation differences.

9. Risks/Discomforts: There is no known risk involved in this project. Participants are musicians and are skilled in listening, performing, and tuning; though they will have different skill levels in each of these musical domains. The process poses no physical or mental discomfort. Recorded audio volume levels will be equalized and set at a comfortable level for all participants.

10. Right to Refuse: Participation in this study is voluntary. At any time, the subject may withdraw from the study without penalty or loss of any benefit to which the subject may otherwise be entitled.

11. Privacy: The study is confidential. Codes will link data to identity. Records will be maintained in secure office storage by the principal investigator only. Results of the study may be published but no names or identifying information will appear in any publication. Data will be kept confidential unless release is legally compelled.

12. Financial Information: Participants will not receive financial compensation for participation and will not incur financial cost.

Thank you for your time and willingness to participate in this study. Please return this form to the principal investigator as soon as possible.

Signatures:
The study has been discussed with me and all my questions have been answered. I may direct additional questions regarding study specifics to the investigators. If I have questions about subjects' rights or other concerns, I can contact Robert C. Matthews, Chairman, LSU Institutional Review Board, 225-578-8692, irb@lsu.edu, www.lsu.edu/irb. I agree to participate in the study described above and acknowledge the researchers' obligation to provide me with a copy of this consent form if signed by me.

Teacher Name (print): PATRICIA ROUSSEL

Teacher Signature: Pat Roussel

Date: 11/17/10

APPENDIX D
PARENTAL CONSENT FORM

1. Title: Direction of Mistuning, Magnitude of Cent Deviation, and Timbre as Factors in Musicians' Pitch Discrimination in Simultaneous and Sequential Listening Conditions

2. Performance Sites: Louisiana State University, Parkview Baptist High School, and East Ascension High School

3. Contacts: The following investigators are available for questions about this project:

N. Alan Clark, principal investigator, Louisiana State University
clarkn@tigers.lsu.edu
225-270-7722

Dr. James L. Byo, Carl Prince Matthies Professor of Music Education, faculty advisor
jbyo@lsu.edu
225-578-2593

4. Purpose of Study: This inquiry is designed to answer the following primary questions: At what point do instrumentalists begin to discern minute changes in pitch? Is there a difference between subjects' ability to detect sharp and flat pitches? Does the stimulus wave form affect perception of pitch change? Do subjects discern pitch changes better in a sequential format, or a simultaneous format, or does it matter? Secondarily, the study will attempt to answer the following: Does age affect pitch perception accuracy? Does experience affect pitch perception accuracy? Does instrument family affect pitch perception accuracy?

5. Subjects: Subjects will be students enrolled in the band programs at Parkview Baptist High School, Baton Rouge, Louisiana; East Ascension High School, Gonzales, Louisiana; and the Louisiana State University Band program.

6. Number of Subjects: 120 (Sixty from the two high schools and sixty from LSU)

7. Study Procedures: Subjects will complete a data inform in which they will indicate their age, years of private study, and gender. Subjects under the age of 18 will obtain parent/guardian approval prior to testing. Subjects 18 or older will give informed consent before the test begins. Following the procurement of this information, subjects listen to a series of pre-recorded tones and indicate whether each pair of tones is the same, or how it is different by circling "lower, same, higher" on an answer sheet.

8. Benefits: There will likely be no immediate benefit to subjects. Potentially, study results will provide evidence regarding subjects' ability to detect intonation differences.

9. Risks/Discomforts: There is no known risk involved in this project. Subjects are musicians and are skilled in listening, performing, and tuning; though they will have different skill levels in

each of these musical domains. The process poses no physical or mental discomfort. Recorded audio volume levels will be equalized and set at a comfortable level for all subjects.

10. Right to Refuse: Participation in this study is voluntary. At any time, the subject may withdraw from the study without penalty or loss of any benefit to which the subject may otherwise be entitled.

11. Privacy: The study is confidential. Codes will link data to identity. Records will be maintained in secure office storage by the principal investigator only. Results of the study may be published but no names or identifying information will appear in any publication. Data will be kept confidential unless release is legally compelled.

12. Financial Information: Subjects will not receive financial compensation for participation and will not incur financial cost.

I understand the scope and intent of this study and all my questions have been answered. I may direct additional questions regarding study specifics to the investigators. If I have questions about subjects' rights or other concerns, I can contact Robert C. Matthews, Chairman, LSU Institutional Review Board, 225-578-8692, irb@lsu.edu, www.lsu.edu/irb. I agree to allow my child to participate in the study described above and acknowledge the researchers' obligation to provide me with a copy of this consent form if signed by me.

Participant Age: ______

Participant Name: ____

Parent/Guardian Name (please print): ____

Parent/Guardian Signature: __________________________ Date: ___________

APPENDIX E
ADULT PARTICIPANT CONSENT FORM

1. Title: Direction of Mistuning, Magnitude of Cent Deviation, and Timbre as Factors in Musicians' Pitch Discrimination in Simultaneous and Sequential Listening Conditions

2. Performance Sites: Louisiana State University, Parkview Baptist High School, and East Ascension High School

3. Contacts: The following investigators are available for questions about this project:
N. Alan Clark, principal investigator, Louisiana State University
clarkn@tigers.lsu.edu
225-270-7722

Dr. James L. Byo, Carl Prince Matthies Professor of Music Education, faculty advisor
jbyo@lsu.edu
225-578-2593

4. Purpose of Study: This inquiry is designed to answer the following primary questions: At what point do instrumentalists begin to discern minute changes in pitch? Is there a difference between subjects' ability to detect sharp and flat pitches? Does the stimulus wave form affect perception of pitch change? Do subjects discern pitch changes better in a sequential format, or a simultaneous format, or does it matter? Secondarily, the study will attempt to answer the following: Does age affect pitch perception accuracy? Does experience affect pitch perception accuracy? Does instrument family affect pitch perception accuracy?

5. Subjects: Subjects will be students enrolled in the band programs at Parkview Baptist High School, Baton Rouge, Louisiana; East Ascension High School, Gonzales, Louisiana; and the Louisiana State University Band program.

6. Number of Subjects: 120 (Sixty from the two high schools and sixty from LSU)

7. Study Procedures: Subjects will complete a data inform in which they will indicate their age, years of private study, and gender. Subjects under the age of 18 will obtain parent/guardian approval prior to testing. Subjects 18 or older will give informed consent before the test begins. Following the procurement of this information, subjects listen to a series of pre-recorded tones and indicate whether each pair of tones is the same, or how it is different by circling "lower, same, higher" on an answer sheet.

8. Benefits: There will likely be no immediate benefit to subjects. Potentially, study results will provide evidence regarding subjects' ability to detect intonation differences.

9. Risks/Discomforts: There is no known risk involved in this project. Subjects are musicians and are skilled in listening, performing, and tuning; though they will have different skill levels in

each of these musical domains. The process poses no physical or mental discomfort. Recorded audio volume levels will be equalized and set at a comfortable level for all subjects.

10. Right to Refuse: Participation in this study is voluntary. At any time, the subject may withdraw from the study without penalty or loss of any benefit to which the subject may otherwise be entitled.

11. Privacy: The study is confidential. Codes will link data to identity. Records will be maintained in secure office storage by the principal investigator only. Results of the study may be published but no names or identifying information will appear in any publication. Data will be kept confidential unless release is legally compelled.

12. Financial Information: Subjects will not receive financial compensation for participation and will not incur financial cost.

The study has been discussed with me and all my questions have been answered. I may direct additional questions regarding study specifics to the investigators. If I have questions about subjects' rights or other concerns, I can contact Robert C. Matthews, Chairman, LSU Institutional Review Board, 225-578-8692, irb@lsu.edu, www.lsu.edu/irb. I agree to participate in the study described above and acknowledge the researchers' obligation to provide me with a copy of this consent form if signed by me.

Participant Age: ______

Participant Name (please print): ____

Participant Signature: _______________________________________ Date: ________________

APPENDIX F
PARTICIPANT DATA FORM

Participant Number: ______________ Test Number: ______________

Age ________________

Gender M F

Grade/College Year ____

Major Instrument: ____

Years playing major instrument ________________

Years of private instruction on major instrument __________________

Minor Instrument: ____

Years playing minor instrument ________________

Years of private instruction on minor instrument __________________

Years of private piano instruction __________________

Years of private vocal instruction __________________

Other music instruction ______________________________ Number of years _____________

APPENDIX G
SCRIPT FOR PITCH DISCRIMINATION TESTING SESSION

[SECTION ONE]

“Section one.”

“Thank you for agreeing to participate in this project. In this first of two sections I am going to ask you to listen to a series of tones which are presented in pairs. Sometimes both tones will be of the same quality (or timbre) and sometimes the second tone will be different. Also, sometimes both tones will be exactly the same pitch, and sometimes the second tone will be different. The differences are often quite small, so please listen very carefully to the first tone of each pair and compare it to the altered second tone. If the second tone sounds lower (or flat), circle “lower” on your answer sheet for that pair. If it sounds higher (or sharp) circle “higher”, and if it sounds the same mark “same” on your answer sheet. Again, many of the pitch differences are very small, so listen carefully. There are 56 pairs of tones in this section. We will take a timed, one minute break after number 28. Please be careful to follow the numbering system as it is laid out on the answer sheet. Do you have any questions?”

[*Operator will pause recording*]

[*Resume playback*]

“Here are two practice examples. Please mark your answer sheet for practice examples one and two. Let the test monitor know if your headphone volume needs to be adjusted after marking these examples.”

[*Practice examples*]

“We will now begin section one.”

[SECTION TWO]

“Section two.”

“In this section there are also 56 pairs of tones. Here the second tone will join the first tone after the first tone has sounded for two seconds. The two tones will then sound together for two more seconds. Please circle on your answer sheet whether the second tone is lower, the same, or higher than the first tone. As in section one, many of the pitch differences are quite small, so listen very carefully. There are 56 pairs of tones in this section and we will take a one minute

break after number 28. Please be careful to follow the numbering system as it is laid out on the answer sheet. Do you have any questions?"

[*Operator will pause recording*]

[*Resume playback*]

"Here are two practice examples. Please listen carefully to the examples and mark your answer sheet for practice examples one and two."

[*Practice examples*]

"We will now begin section two."

APPENDIX H
PITCH MATCHING TEST ANSWER SHEET

Participant Number: ______________ Test Number: ______________

Section one practice examples;

[Practice example 1] Lower Same Higher

[Practice Example 2] Lower Same Higher

SECTION ONE

1.	Lower	Same	Higher	16.	Lower	Same	Higher
2.	Lower	Same	Higher	17.	Lower	Same	Higher
3.	Lower	Same	Higher	18.	Lower	Same	Higher
4.	Lower	Same	Higher	19.	Lower	Same	Higher
5.	Lower	Same	Higher	20.	Lower	Same	Higher
6.	Lower	Same	Higher	21.	Lower	Same	Higher
7.	Lower	Same	Higher	22.	Lower	Same	Higher
8.	Lower	Same	Higher	23.	Lower	Same	Higher
9.	Lower	Same	Higher	24.	Lower	Same	Higher
10.	Lower	Same	Higher	25.	Lower	Same	Higher
11.	Lower	Same	Higher	26.	Lower	Same	Higher
12.	Lower	Same	Higher	27.	Lower	Same	Higher
13.	Lower	Same	Higher	28.	Lower	Same	Higher
14.	Lower	Same	Higher	------------------BREAK------------------			
15.	Lower	Same	Higher	29.	Lower	Same	Higher

30.	Lower	Same	Higher	44.	Lower	Same	Higher
31.	Lower	Same	Higher	45.	Lower	Same	Higher
32.	Lower	Same	Higher	46.	Lower	Same	Higher
33.	Lower	Same	Higher	47.	Lower	Same	Higher
34.	Lower	Same	Higher	48.	Lower	Same	Higher
35.	Lower	Same	Higher	49.	Lower	Same	Higher
36.	Lower	Same	Higher	50.	Lower	Same	Higher
37.	Lower	Same	Higher	51.	Lower	Same	Higher
38.	Lower	Same	Higher	52.	Lower	Same	Higher
39.	Lower	Same	Higher	53.	Lower	Same	Higher
40.	Lower	Same	Higher	54.	Lower	Same	Higher
41.	Lower	Same	Higher	55.	Lower	Same	Higher
42.	Lower	Same	Higher	56.	Lower	Same	Higher
43.	Lower	Same	Higher				

END OF SECTION ONE

Section two practice examples;

[Practice example 3]	Lower	Same	Higher
[Practice Example 4]	Lower	Same	Higher

SECTION TWO

1.	Lower	Same	Higher	20.	Lower	Same	Higher
2.	Lower	Same	Higher	21.	Lower	Same	Higher
3.	Lower	Same	Higher	22.	Lower	Same	Higher
4.	Lower	Same	Higher	23.	Lower	Same	Higher
5.	Lower	Same	Higher	24.	Lower	Same	Higher
6.	Lower	Same	Higher	25.	Lower	Same	Higher
7.	Lower	Same	Higher	26.	Lower	Same	Higher
8.	Lower	Same	Higher	27.	Lower	Same	Higher
9.	Lower	Same	Higher	28.	Lower	Same	Higher
10.	Lower	Same	Higher	------------------BREAK------------------			
11.	Lower	Same	Higher	29.	Lower	Same	Higher
12.	Lower	Same	Higher	30.	Lower	Same	Higher
13.	Lower	Same	Higher	31.	Lower	Same	Higher
14.	Lower	Same	Higher	32.	Lower	Same	Higher
15.	Lower	Same	Higher	33.	Lower	Same	Higher
16.	Lower	Same	Higher	34.	Lower	Same	Higher
17.	Lower	Same	Higher	35.	Lower	Same	Higher
18.	Lower	Same	Higher	36.	Lower	Same	Higher
19.	Lower	Same	Higher	37.	Lower	Same	Higher

38. Lower Same Higher

39. Lower Same Higher

40. Lower Same Higher

41. Lower Same Higher

42. Lower Same Higher

43. Lower Same Higher

44. Lower Same Higher

45. Lower Same Higher

46. Lower Same Higher

47. Lower Same Higher

48. Lower Same Higher

49. Lower Same Higher

50. Lower Same Higher

51. Lower Same Higher

52. Lower Same Higher

53. Lower Same Higher

54. Lower Same Higher

55. Lower Same Higher

56. Lower Same Higher

END OF SECTION TWO

VITA

N. Alan Clark has taught music and conducted at all levels from middle school through professional military bands. While teaching at Kathleen High School in Lakeland, Florida he also served as an adjunct music faculty member at Florida Southern College, and as an instructor with the Suncoast Sound Drum and Bugle Corps. In 1987 he entered the United States Air Force (USAF) and served as saxophone section leader and Drum Major with the USAF Band of the West in San Antonio, Texas. In 1990 he was commissioned and appointed Deputy Commander of the Band of the USAF in Europe. In 1993 Lieutenant Clark was appointed Deputy Commander of the Air Force Band of Flight in Dayton, Ohio and in 1996 he assumed command of The Band of the USAF Reserve. Major Clark accepted the appointment as Air Force Reserve Officer Training Corps Commandant of Cadets at Louisiana State University in April 2004, and he retired from the Air Force in 2007 after twenty years of service. Major Clark is a member of the National Association for Music Education, the National Band Association, the College Band Directors National Association, Pi Kappa Lambda, Kappa Kappa Psi, and Phi Mu Alpha. He holds both the Bachelor of Music Education and the Master of Fine Arts in Saxophone Performance degrees from the University of Florida as well as the Master of Science in International Relations from Troy University.