University of Alberta

Discrimination of musical intervals by humans and chickadees: Cue salience modulated by timbre

by

Lee Michael Vilinsky

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

Psychology

Fall 2013
Edmonton, Alberta

Abstract

Musical consonance/dissonance, roughly defined by its characteristic stability/instability, has been shown to be a relatively salient feature of sound. The extent to which the salience of this property varies as a function of timbre, a property that distinguishes two sounds of the same pitch and loudness, is currently unknown. A Go/No Go operant task was employed to test how humans (Experiment 1) and black-capped chickadees (Experiment 2) discriminate synthetic and piano musical intervals of varying consonance/dissonance. Humans that discriminated synthetic intervals had proportionally higher error rates for intervals where the upper notes were close in pitch whereas humans that discriminated piano stimuli had more errors to stimuli related by consonance. Chickadees showed a similar trend for synthetic intervals but not for piano intervals. Taken together, these findings suggest that timbre modulates the salience of consonance/dissonance for auditory discrimination tasks but that the relative salience varies across species.

Acknowledgements

I would like to give my biggest and sincerest thanks to my supervisor Dr. Christopher Sturdy for all of his guidance and support throughout my graduate school career. The countless hours spent designing, interpreting, and discussing the projects constituting this thesis is met with great appreciation. Without his knowledge and advice, the journey to this point would not have been possible. I would also like to thank my graduate committee, Dr. Peter Hurd and Dr. Dallas Treit, for their guidance, time, and commitment as well.

I would like to extend special thanks to Dr. John Lind, Brian Dupuis, and all of my lab mates (Marisa Hoeschele, Allison Hahn, Lauren Guillette) for their suggestions and advice while working on these projects. Surely, the thoughtful discussions involved with the design and analyses of these experiments in addition to some of the data collection helped this thesis come to fruition.

Finally, I thank my parents, Lisa and Dean Vilinsky, for always being there for me and believing in me. And a very special thanks to Susan Larcombe for the continuous encouragement throughout the duration of this thesis and the immeasurable love and support that is unendingly provided.

Table of Contents

List of Tables

List of Figures

(A), subsequently make the S+ box (B) turn green. Feedback following a response was displayed in an adjacent white box (C).

Figure 2.4 (Page 31) Results for groups for human Experiment 1a. Each line on each graph represents a particular S- stimulus. Proportion of subjects responding is plotted on the y-axis and trial number is on the x-axis. A) Phase 1 and 2 for participants reinforced for P8. B) Phase 1 and 2 for participants reinforced for P5. C) Phase 1 and 2 for participants reinforced for M7. D) Phase 1 and 2 for participants reinforced for TT.

Figure 2.5 (Page 36) Results for groups for human Experiment 1b. Each line on each graph represents a particular S- stimulus. Proportion of subjects responding is plotted on the y-axis and trial number is on the x-axis. A) Phase 1 and 2 for participants reinforced for P8. B) Phase 1 and 2 for participants reinforced for P5. C) Phase 1 and 2 for participants reinforced for M7. D) Phase 1 and 2 for participants reinforced for TT.

Figure 3.1 (Page 56) Average Discrimination Ratio (DR) for each S- stimulus across the eight Vincentized blocks for chickadees in the synthetic condition (see text for details). Panel (A) shows phase 1 and phase 2 chickadee discrimination for the P5 S+ group and panel (B) shows phase 1 and phase 2 discrimination for the TT S+ group. Error bars show 95% confidence intervals.

Figure 3.2 (Page 60) Average Discrimination Ratio (DR) for each S- stimulus across the eight Vincentized blocks for chickadees in the piano condition (see text for details). Panel (A) shows phase 1 and phase 2 chickadee discrimination for the

Chapter 1

Introduction

As with any modality of perception, audition employs both bottom-up and top-down processing. For example, a sound that is heard in the environment can be immediately classified as urgent, pleasing, familiar, or novel based on various acoustic properties of the signal. To address how organisms process auditory information, it is thus paramount to consider “the nature of the stimulus, the input without which no processing can take place” in behavioral research (Garner, 1970, p.1). As Garner suggests in his 1970 thesis on information processing, focusing on stimulus complexity allows a researcher to draw conclusions about aspects of the stimulus that govern behavior from a bottom-up perspective. Furthermore, paying close attention to the stimulus at hand allows one to determine those aspects of the stimulus that are separable (perceived independently) and integral (perceived dependently) – a critical distinction when planning an experiment that uses stimuli of multiple dimensions (Garner, 1970).

On the other hand, understanding behavior from a top-down perspective can prove equally important in that it breaks down a system to gain insight into its subsystems. Top-down perception involves schema that reflect knowledge and expectation and consequently has the ability to modify the initial organization derived from bottom-up processing. For example, when an auditory stimulus is familiar (e.g., a mother's voice) and is partially occluded, the listener will understand the occluded portions' meaning more quickly and accurately relative to an unfamiliar stimulus that is also partially occluded (e.g., a stranger's voice). Naturally, an experimental task that necessitates learning probes faculties involved with top-down processing. One argument against the top-down perspective, however, is that it is a "black box" explanation of behavior – that is, it is sometimes difficult to attribute the behavioral outcome of a specific variable. A major proponent of this argument was the late J.J. Gibson, supporting an ecological or "direct realist" theory of information processing (Gibson, 1966). According to Gibson's theory, the proximal stimulus is very rich with information and, as a result, leads to a direct (i.e., not a reconstructive or representative) perception. This view rejects the usage of learning, memory, and any sort of cognitive processing in understanding the world through perception.

As a result of this ongoing debate, both approaches should not be studied separately, but rather together. They both aim to understand the cause of a single output, behavior, as determined by the effects of various stimuli, though their experimental foci may not have much overlap. This synergistic approach to understanding auditory perception and cognition is the foundation of much behavioral research since the 1970’s and is a basic premise of the research discussed here as applied to music and its adaptive significance.

1.1 The Evolution of Music

Within the realm of auditory information processing lies a subdivision of research focused on the perception of music and its underlying constituents. Using music as a model system to understand information processing encapsulates both lower- and higher-order cognitive functioning and has revealed itself to be instrumental for informing scientists of broader issues in cognitive psychology over the last thirty years (Tirovolas & Levitin, 2011; Levitin, 2010). Music, akin to language, is regarded as a universal to humankind, and as such the study of music has yielded many interesting results behaviorally, developmentally, neurologically, and clinically. For example, in a landmark study investigating innate preferences for musical stimuli, Zentner and Kagan (1998) suggest that human infants are biologically predisposed to treat some sounds as more pleasing than others based on their musical structure (Zentner & Kagan, 1998). This finding of a preference for specific musical stimuli has not only been replicated (Schellenberg et al., 2005), but has been shown in other species as well (Chiandetti & Vallortigara, 2011), which suggests that its perception is rooted in acoustic properties. Harmony, the relation among simultaneous tones, and its purpose in musical structure is simply one aspect of music that researchers have addressed in years past but has proved useful in spanning across research fields in memory, attention, perception, categorization, emotion, and experience (Levitin, 2010).

Understanding music through an evolutionary lens, however, is a more recent development. It has been proposed that music and language are intrinsically linked and therefore are not processed independently in the human brain (Patel, 2008). The underlying mechanisms that have fostered the evolution of music, and its debated co-evolution alongside language, are therefore of much interest because they illuminate the degree to which music serves a communicative function (see Hauser & McDermott 2003 for a review). Given this, many have suggested that music might be a biological adaptation to function in a group, facilitating cohesion, courtship, and leisure (see McDermott & Hauser, 2005 for a review). Still, Steven Pinker perhaps most famously addressed music as "auditory cheesecake…a superfluous confection" that has come about as a byproduct of auditory functions that have evolved for other reasons (Pinker, 1997). Though both sides of this debate present evidence that supports their respective position, it is worth noting that they are not mutually exclusive. Some aspects of music may very well be peripheral to the auditory system as a whole and others may have evolved specifically as music-adaptive functions. The question then becomes which specific aspects of music are unique and/or innate. As it turns out, one major theme generated by this question is the creation of cross-species behavioral experiments: that are rooted in the field of comparative psychology.

One major objective of comparative psychology is to determine cognitive processing differences between humans and non-human animals. Comparative experiments have examined music perception in non-human animals and have demonstrated categorization and preference behavior based on musical constructs (e.g., chords, consonance/dissonance) in a multitude of species (e.g., Hoeschele et al., 2012; Tierney, Russo, & Patel, 2011; Watanabe, Uozumi, & Tanaka, 2005). Though experiments such as these address issues that are seemingly simple to humans, they (and others like them) have laid the foundation for a fertile avenue of research that aims to disentangle the evolutionary origins of music. The most important contribution studies like these have made to the field of comparative music psychology is the realization that the perception of music is not uniquely human and thus deserves to be recognized as serving some degree of specific adaptive function.

The goal of the research presented in this thesis is to elucidate species-specific differences in the perception of musical consonance/dissonance. The emergence of this property, defined by Western harmony in human music, has been debated for generations (Helmholtz 1885; Roederer 1975; Tenney 1988; Hartmann 1998). The question still remains: why is harmony defined as it is? Comparing how humans and nonhumans attend to this property may lead to telling conclusions regarding this question. Before investigating this property through behavior, however, it is paramount to understand the foundation of music as determined by the physical properties of sound.

1.2 Musical Properties of Sound

Pitch is the subjective perception of the fundamental frequency (F_0) of a sound source and allows the ordering of sounds on a frequency-related scale (Plack et al., 2005). Pitch is often viewed as extending along two dimensions in a helix formation: pitch height and pitch chroma (see Figure 1.1; Shepard, 1982). Pitch height is a linear attribute that has a positive relationship with frequency, though this relationship is complex. Pitch chroma refers to a sound quality shared by tones that have the same octave interval, or a doubling of frequency (Bachem, 1950; Shepard, 1982). The three-dimensional space occupying the helix model of pitch allows circular organization, but it is important to note that distance between notes is not accurately represented by this model. Moreover, perception of pitch is related to frequency by a non-linear process (Licklider, 1951); it is different than, and sometimes more than, simply a reflection of energy at specific frequencies in the spectrum of the physical signal. For example, pitch is perceived to increase more rapidly than frequency for tones below 1000 Hz but less rapidly for tones above 1000 Hz (Stevens, Volkmann, & Newman, 1937).

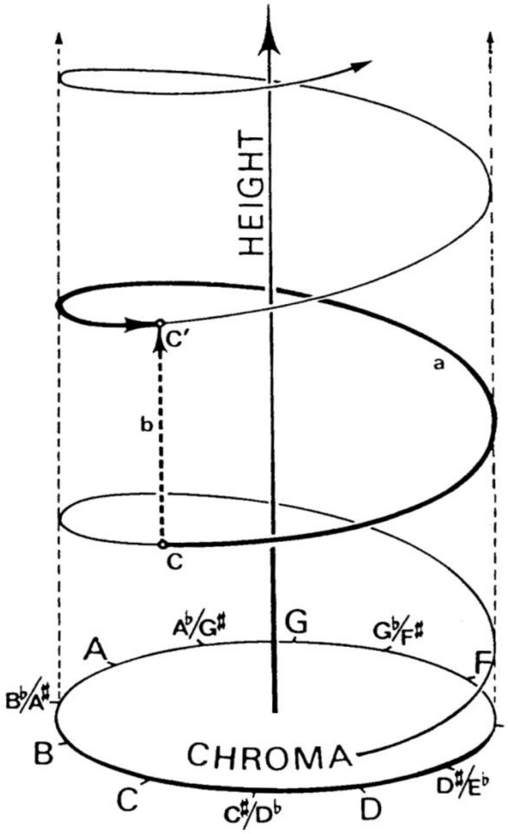


Figure 1.1 Helix model of pitch. Pitch height is represented by an increase or decrease in vertical space. Musical notes that have the same chroma are on the same vertical axis in pitch space. (From "Approximation to Uniform Gradients of Generalization by Monotone Transformations of Scale" by Roger N. Shepard. From Stimulus Generalization, D. I. Mostofsky (Ed.), Copyright (c) 1965 by the Board of Trustees of the Leland Stanford Jr. University, renewed 1993.

Though pitch height has been shown to be an extremely salient feature of sound in humans and non-human animals, few sounds in the environment consist of strictly the fundamental frequency. Most sounds are complex and contain higher frequencies (harmonics) above the fundamental, providing additional information about the sound’s source to the organism perceiving a sound (Grey, 1978; Singh, 1986). The complexity of a given sound varies in regard to the spectral distribution of power across the upper harmonics and the temporal envelope (attack, steady-state, and decay). These spectral and temporal cues comprise what is known as timbre (Luce, 1963; Grey, 1976).

Formally, timbre is the psychological sensation by which a listener can discriminate two sounds with the same pitch and volume. It is represented by the power distribution across integer multiples (2×, 4×, 10×, etc.) of the fundamental frequency. The vibration frequencies of harmonics from natural sounds are rarely precise multiples of the fundamental, however. The ‘inharmonicity’ of integer multiples that are not whole numbers is another factor that gives naturally-produced sounds an additional characteristic of warmth (Handel, 1989). By simply adding harmonics to a pure tone, however, the fundamental frequency is still perceived, but its timbre is changed. Moreover, manipulating the distribution of power across these harmonics alters timbre as well. A listener will perceive a trumpet, piano, or violin playing ‘middle C’ (262 Hz) as the same pitch but will be able to identify the sound source because of differences in timbre. Previous research has shown that perception of timbre and pitch interacts in that they are not perceived independently (Melara & Marks, 1990; Krumhansl & Iverson, 1992). Ultimately, it is the combination of these sinusoidal vibrations across the frequency spectrum that determines the timbre of a sound.

Sinusoid vibrations arise in an immense array of mechanical systems ranging from strings on a piano to the vocal folds of vocalizing species. What all of the instances that produce these systematic vibrations have in common is that the restoring force is equal to the displacement – that is, air molecules actively equalize their distribution in space following a disruption in force (see Handel, 1989). Moreover, any oscillation or vibrating pattern that repeats indefinitely can be built up from (or is analyzable into) the sum of its parts. This is known as Fourier's theorem and dictates that, as applied strictly to psychology and psychophysics, the whole of a given stimulus is equal to the sum of its parts (Lejeune-Dirichlet, 1829). A Fourier analysis is useful for auditory research in that it decomposes periodic signals, such as sinusoids, into the sum of a (possibly infinite) set of simple oscillating functions. This way of viewing auditory signals has proved paramount in contributing to our understanding of how sound functions in the environment.

Finally, a relatively salient feature of sound that is often overlooked is musical consonance/dissonance. Though many definitions and adjectives have been applied to the phenomenon, consonance and dissonance is generally understood to be a sound's stability and instability, respectively (see Cazden, 1980). Consonance/dissonance perception is only possible with complex sounds that encompass two or more musical notes, for it is the relative frequency differences among the notes that make it either consonant or dissonant. For example, a note C4 (C in the fourth octave, 262Hz) paired with G4 (392Hz) creates a musical interval of a perfect fifth, defined by its frequency ratio of 3:2. A sound that contains a simple frequency ratio is classified as consonant because there are more occurrences at which the two frequencies are in phase with each other (Plomp & Levelt, 1965). That is, there is less fluctuation in amplitude over time. The same root note C4 (262Hz) paired with F#4 (370Hz) creates a musical interval of a tri-tone; Defined by its frequency ratio of 32:45, it is significantly more complex than a perfect fifth. As such, it is classified as a dissonant interval. This “frequency-ratio” theory, also known as the Natural Law Theory (Cazden, 1959), is generally accepted, but other research suggests that factors such as culture (Lundin, 1947), musicianship (Pitt, 1994), and even individual differences (McDermott, Lehr, & Oxenham, 2010) contribute to this phenomenon as well.

Other models of consonance/dissonance perception do exist though most do not account for dissonant sounds that exceed a critical bandwidth (Boomsliter & Creel, 1961; Greenwood, 1961). The critical bandwidth is a frequency processing unit for the auditory system that functions as a filter. Auditory roughness (dissonance) results from the inability of this processing unit to resolve inputs whose frequency difference is smaller than the critical bandwidth, a confined frequency range (Kameoca & Kuriyagawa, 1969, Plomp & Levelt, 1965). If two frequencies occur beyond their respective critical bands, then the bands do not overlap and thus roughness does not occur. For tone pairs with higher frequencies (e.g., tones greater than 500 Hz), the width of the critical band is a constant frequency ratio between two and three semitones. Tone pairs under 500 Hz have a decreasing critical bandwidth as frequency also decreases (Plomp & Levelt, 1965). Thus, the relative consonance of simultaneous pure tones is a function of absolute frequency distance independent of the simplicity or complexity of frequency ratios.

An important illustration is the differences in consonance/dissonance for musical intervals that possess timbre as opposed to pure tones. The Pythagorean ratios discussed above refer to the fundamental frequency of two notes. For two complex tones – tones that contain more frequencies than just the fundamental – higher harmonics create more overlapping frequencies of the two tones. The resulting harmonic overlap will either be more or less in phase with each other and is correlated with the ratio of the fundamental frequencies (Plomp & Levelt, 1965). A phenomenon known as beating arises when the overlapping frequencies are near misses of each other (i.e., within a critical frequency bandwidth). When the harmonic frequencies overlap exactly, beating is absent. The presence, absence, and the degree of beating determines what Helmholtz (1954) referred to as auditory roughness and is the physical foundation of consonance/dissonance.

For example, consider two complex tones whose frequencies correspond to a ratio of 2:1 (octave). Half of the harmonics of the lower tone are present in the harmonic series of the upper tone, while all of the harmonics of the higher tone are present in the series of the lower tone. For tones that stand in a ratio of 3:2, one third of the harmonics of the lower tone are present in the series of the higher tone, while half of the harmonics of the higher tone are present in the series of the lower tone. Thus, sensations of beating arising from harmonics that are close but not identical in pitch are less likely between tones related by simple frequency ratios (more common harmonics) than between tones related by more complex ratios (fewer common harmonics). The presence of harmonics in these intervals adds many more layers of frequency ratios to the overall stimulus. An interval with power distributed across many harmonics as opposed to an interval consisting of strictly fundamental frequencies will therefore contain more overlapping critical bands which, in turn, increase auditory roughness.

Taken together, these properties of sound vary on distinct and sometimes related dimensions. Moreover, the complexities of single sounds are exacerbated by the complexities of sound sequences or context. Single sounds studied in isolation makes their intricacies easier to control for and consequently allows researchers to conclude that certain sounds are perceived in certain ways. One thing in common with all of these properties, however, is that their objective physicality bears a complex relationship to their subjective percept; pitch is the complex perception of frequency and timbre is the complex perception of harmonic spectra and temporal envelope. This distinction between objective and subjective relationships is critical when considering cross-species examinations of stimulus perception because a stimulus defined by objective, physical properties may be perceived differently among the species being compared.

1.3 Current Studies

The current set of studies explores the processing of musical stimuli in two different species: humans (*Homo sapiens*) and songbirds – specifically, black-capped chickadees (*Poecile atricapillus*). Music research has focused most extensively on humans as it was previously thought that it was a uniquely human percept and activity. Songbirds, however, present a unique opportunity for a cross-species comparison of auditory perception and cognition for a few critical reasons. Songbirds, as well as humans, exhibit vocal learning, the ability to modify vocal signals as a direct result of experience associated with a tutor, such as a parent (Thorpe, 1954; Marler, 1970; see Nottebohm, 1972 for a review). This ability to imitate and form templates of specific vocalizations is highly evolved, present in only a handful of all vocalizing species, and is used in a way to better facilitate communication among conspecifics (see review by Doupe & Kuhl, 1999). Another reason is that pitch perception has been studied extensively in black-capped chickadees. Much of this research suggests that this species has exceptional pitch sensitivity as compared to humans, including both absolute and, to a lesser extent, relative pitch (e.g., Weisman & Ratcliffe, 1989; Weisman, Njegovan, & Ito, 1994; Weisman et al., 2010). As such, songbirds are chosen by many researchers interested in disentangling the evolutionary origins of music because of their reliance on acoustic input and communication.

As mentioned previously, the focus of the current experiments employs a synergy of bottom-up and top-down processing. The main questions being addressed are: (1) whether musical consonance/dissonance is a uniquely human percept and (2) whether the saliency of this dimension varies with other aspects of sound. Although a majority of recent studies discern valuable information into scientific understanding of the first question, most do not address issues related to the second. Moreover, such studies do not address the relative saliencies of these properties when pitted against each other in an operant learning paradigm. Much of the literature on human auditory perception is focused on pitch while consonance/dissonance has only had a recent surge of scientific inquiry (e.g., Trainor et al., 2002; Chiandetti & Vallortigara, 2011; Krohn et al., 2007). It is known that consonance/dissonance is a salient property of sound, but it is unknown how this saliency fluctuates relative to other properties of sound that are either present or absent.

One challenge to evaluating the possible importance of consonance/dissonance relative to pitch in a discrimination task is that different musical intervals are produced by manipulation of pitch. The research question thus is carefully formulated to distinguish between learning disparities that are due to differences in pitch and those that can be attributed to the emergent property of musical consonance/dissonance. As such, the current study employs musical intervals to control for relative pitch, the ability to perceive the relation between two or more notes, and consonance/dissonance equally. Given what is known about consonance/dissonance and its relation to timbre, it is expected that human subjects exposed to pure tones will rely more heavily on the pitch of the upper notes of the intervals to discriminate the sounds. Furthermore, human subjects exposed to the same intervals with timbre present will rely more on the relative consonance/dissonance among the intervals to discriminate them because the increased number of interactions among frequencies will make the property sound more salient. If these predictions are correct, then it will suggest that timbre modulates whether consonance/dissonance or pitch is a more salient cue for auditory discrimination. Though it has been shown that songbirds can discriminate consonant and dissonant chords and use timbre as a central component for auditory processing (e.g. Braaten & Hulse, 1991), it is unknown how the salience of consonance/dissonance modulates with timbre. There are therefore no *a priori* predictions for chickadees in regard to response differences to consonant and dissonant stimuli with different spectral structures.

Chapter 2

Discrimination of Synthetic and Piano Musical Intervals in Humans (*Homo sapiens*)

2.1 Introduction

The evolutionary origin of music has puzzled scientists since at least the time of Darwin (1871) though it has only recently become a popular research topic aimed at disentangling its adaptive significance (Tirovolas & Levitin, 2011; McDermott & Hauser, 2005).The origin of music may not be simple to uncover, however, as the music faculty integrates many domains and serves a variety of functions that all may vary in evolutionary history (Fitch, 2006). An important first step, then, to elucidate music's evolutionary history is to determine which specific features of music are unique to humans, illuminating attributes that may have been targets for natural selection.

Much of the debate that surrounds the evolution of music involves experiential factors that exist within the human species. As such, comparative research between humans and nonhuman animals has become a primary method to probe for the uniqueness of certain properties of music as it allows researchers to control for experience (Chiandetti & Vallortigara, 2011; Watanabe, Uozumi, & Tanaka, 2005). This approach has made great strides in advancing our understanding of music perception and production across a variety of species, most notably those that rely on perceiving and producing vocalizations to communicate (e.g. Hoeschele et al., 2012). Specifically, songbirds have proved to be an effective model for cross-species comparisons as they are, similarly to humans and only a handful of other animals (such as dolphins, whales, and bats), vocal learners (Thorpe, 1954; Marler, 1970; Jarvis, 2004). When considering the ability of these nonhuman species to understand specific properties of Western music, however, it is paramount for the researcher to employ variations along the dimensions of sound that are used for testing.

Much of the early literature addressing music perception in humans has focused on the perception of pitch, the perception of the fundamental frequency (F_0) of a sound. This may largely be due to the finding that pitch is central to human communication and has great significance in Western music (Krumhansl, 1991). In Western music, however, the absolute values of pitches are generally less important than the relationships among them. The former is referred to as absolute pitch and the latter as relative pitch, though the ability to use both cues to determine auditory significance is employed by many vocalizing species. But this does not imply that both cues are treated equally in the species that use these abilities. Humans use relative pitch through the transposition of melodies in order to sing the melody at a more natural vocal range (Patel, 2003). Humans are much less adept at absolute pitch perception, this ability occurring in less than 1 in 10,000 individuals (Bachem, 1955). Humans do show strong relative pitch processing from about the age of 6 (Takeuchi & Hulse, 1993) and it has been shown to be important for learning certain languages (Deutsch et al., 2006). Finally, there is evidence that humans can further develop both abilities through rehearsal (Pitt, 1994; Brady, 1970;).

Two musical notes presented simultaneously will elicit sensations of consonance (pleasantness) or dissonance (unpleasantness). Many attempts have been made to explain this phenomenon through culture (Butler & Datson, 1968), biological preparedness (Trainor, Tsang, & Cheung, 2002), and individual differences (McDermott, Lehr, & Oxenham, 2010). Ancient ideology (e.g., Pythagoras ca. 600 BC; see Plomp & Levelt, 1965 for a historical review) had hypothesized that the ratio between the fundamental frequencies of two notes determined whether an interval was consonant or dissonant – simple ratios (e.g., 200 Hz and 300 Hz; 2:3) yielded consonance and relatively complex ratios yielded dissonance (e.g., 200 Hz and 340 Hz; 10:17). Though this explanation holds true for most Western music, it is more recently thought not to be the determinant of consonance/dissonance but a correlate with acoustic properties that are important to the auditory system (McDermott & Hauser, 2008).

Consonance/dissonance perception has more recently been compared in nonhuman species with inconclusive results. Java sparrows (*Lonchura oryzivora*; Watanabe, Uozumi, & Tanaka, 2005) and European starlings (*Sturnus vulgaris;*

Hulse, Bernard, & Braaten, 1995) can discriminate consonant and dissonant chords and generalize this discrimination to a new set of chords, showing that they can use relative pitch to solve a task. In listening preferences tasks, however, chicks have shown an innate preference for consonance (Chiandetti & Vallortigara, 2011) but some species of nonhuman primates have failed to show any preference even when humans preferred consonant stimuli in a similar task (McDermott & Hauser, 2004). Though the case for consonance/dissonance perception in non-human species is somewhat divergent, studies like these suggest that consonance/dissonance may be an emergent property that is not unique to the human species.

The most prevalent explanation of consonance/dissonance is attributed to Helmholtz (1954) and his assertion that sounds are distinguished by a phenomenon known as beating. Beating is a result of fluctuations in amplitude over time and occurs when the frequencies of two tones are within a critical frequency bandwidth of each other (Plomp & Levelt, 1965). The simple ratios that dictate the fundamental frequencies of consonant musical intervals cause many of the upper components (harmonics) of a complex interval to overlap at exactly the same frequency, leading to relatively less beating than a dissonant interval characterized by a complex ratio and harmonics that do not overlap. Thus, it is plausible that the saliency of consonance/dissonance will be diminished in a musical interval composed of two pure tones (i.e. the fundamental frequencies) compared with a musical interval composed of two complex tones with upper harmonics. Indeed, experiments comparing the relative consonance/dissonance of musical intervals composed of two components or more than two components have shown exactly this (Kameoka & Kuriyagawa, 1969a; Kameoka & Kuriyagawa, 1969b). The question remains whether timbre modulates consonance/dissonance perception in a nonhuman species.

Although a majority of recent studies discern valuable information into scientific understanding of this phenomenon, most do not address the relative saliency of other properties of sound when considering musical consonance/dissonance. It is for this reason that I address consonance/dissonance perception in humans (*Homo* sapiens) as a function of different timbres in Experiments 1a and 1b. The purpose of these two experiments was to establish how humans learn to discriminate synthetic (Experiment 1a) and piano (Experiment 1b) musical intervals. The overall goal was to systematically determine how timbre influences the perception of musical intervals that vary in consonance/dissonance. My expectations for Experiments 1a and 1b were: (1) humans should be able to discriminate all intervals but that the level of expertise and the nature of category formation during the task may vary with musical experience (see Burns & Ward, 1977; Zatorre & Halpern, 1979) and (2) the addition of harmonic frequencies for piano intervals will influence the pattern of responding compared to synthetic intervals. Given the increased beating sensations that arise from intervals with harmonics as opposed to without, it is expected that consonance will be a more salient cue for discriminating piano intervals.

2.2 General Method

Stimuli

Two sets of stimuli were created and corresponded with either musical intervals composed of synthetic pure tones or musical intervals composed of synthetic piano tones with upper harmonics. The first set of stimuli consisted of sinusoids that were 1000 ms in duration and were generated using Audacity version 1.2.6 (Pittsburg, PA). The frequency (note) range of these pure tones was from 261.63 Hz (C4) to 659.26 Hz (E5). Individual notes were then synthesized to create temporally simultaneous musical intervals. Intervals of minor second (m2), tri-tone (TT), perfect fifth (P5), major seventh (M7), and octave (P8) were composed on root notes C4 and E4 (see Figure 2.1). All stimuli were corrected for their respective DC offsets and ramped in amplitude from -96dB to 0dB at onset (0-10ms) and from 0dB to -96dB before offset (990-1000ms).

The second set of stimuli was composed using the default MIDI piano instrument in Anvil Studio (Version 2.26, Willow Software, Shoreline, WA, USA). All five intervals (m2, TT, P5, M7, and P8) on root notes C4 and E4 were composed using this software and underwent the same normalization procedures as the stimuli in the first set. All stimuli from both sets were exported as 16-bit .wav files and were sampled at 44.1 KHz. The piano setting in Anvil Studio incorporates an automatic band-pass filter in that the highest harmonic for any interval was no greater than 10,000 Hz (see Figure 2.2). Figure 2.2 illustrates the spectral differences between these two sets of stimuli.

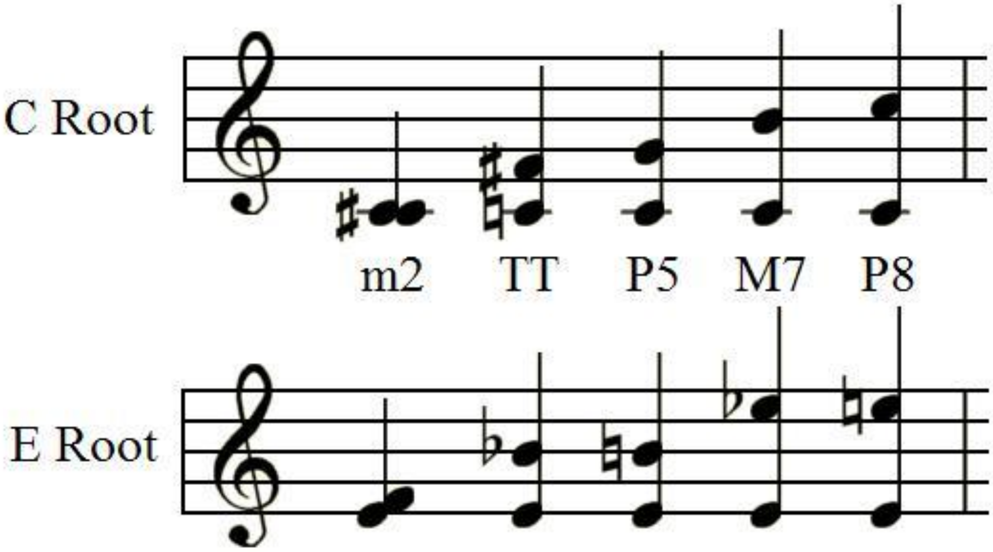


Figure 2.1 Musical notation of intervals used in the current study. The top staff shows the five intervals composed on root note C4 and the bottom staff shows the same intervals composed on root note E4. In order of ascending pitch of the upper note: minor second (m2), tri-tone (TT), perfect fifth (P5), major seventh (M7), and octave (P8).

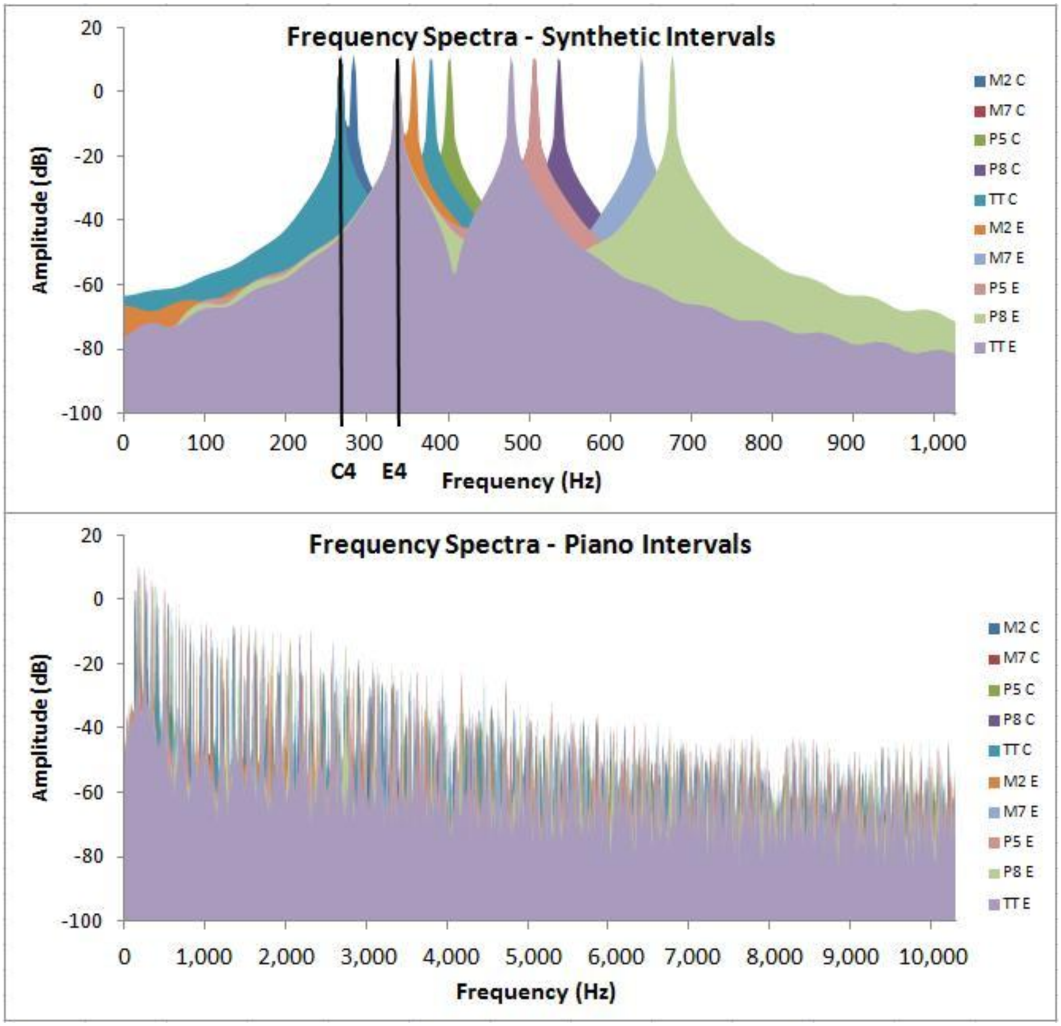


Figure 2.2 Frequency spectrum analysis of pure tone (top) and piano (bottom) stimuli used in all three experiments. The two black lines in the upper panel represent C4 and E4 of the Western chromatic scale, the root notes from which the stimuli used in the current study are composed. Notably, the pure tone musical intervals consist of only F_0 while the piano intervals possess many integer multiples of the F_0.

Apparatus

For each session, up to four participants were run in a single room and were stationed at their own desk, chair, and desktop computer. Each computer was custom-built but had identical components (Intel Core i7 930 CPU, Asus P6T

SE motherboard, Creative Soundblaster Audigy SE sound card) and was outfitted with Sennheiser HD 280 headphones (response 8 Hz - 25 kHz, Sennheiser Canada, Montreal). Custom software programmed in Visual Basic (used by Weisman et al., 2010) presented stimuli and recorded subject responses. All stimuli were presented through the headphones at 75dB(C) +/- 0.1dB(C) and were calibrated using a Brüel & Kjær Type-2239 Sound Level Meter (Brüel & Kjær Canada Ltd., Point Claire, Quebec, Canada).

Design

Participants were tested using a Go/No-Go operant paradigm to discriminate one of the musical intervals from the four other intervals (see Table 2.1). P8 (octave), M7 (major seventh), P5 (perfect fifth) and TT (tri-tone) were all treated as rewarded (S+) stimuli between subjects while the other four stimuli within a condition were unrewarded (S-). The one S+ stimulus was presented on 50% of the trials while the four S- stimuli were presented an equal number of times comprising the other 50% of the trials. For all conditions, a second phase of discrimination was implemented after completion of the first phase with the same intervals and reward contingencies except the intervals differed in root note. Root note order was counterbalanced across subjects within a condition to control for effects of absolute pitch discriminations.

Design of Experiments

Experiment	Species	Timbre	S+ Stimulus (*N*)			
1a	Human	Synthetic	P5 (*32*)	TT (*37*)	P8 (*37*)	M7 (*35*)
1b		Piano	P5 (*31*)	TT (*30*)	P8 (*31*)	M7 (*31*)
2a	Chickadee	Synthetic	P5 (*6*)	TT (*6*)		
2b		Piano	P5 (*6*)	TT (*6*)		

Table 2.2 Design of all experiments. All five intervals (P5, P8, TT, M7, m2) were presented in all conditions. The column labeled “S+ Stimulus” refers to the interval treated as S+ while the four other stimuli were S- (see text for details). Numbers in parentheses next to the S+ groups represent the total sample size for that condition.

Procedure

Upon entering the experimental testing room, participants were randomly assigned to one of the four computer stations and were provided a written description of the task but not the specific goals of the research. Each participant was asked to complete a brief questionnaire about their musical and linguistic background (see Appendix A for a copy of the questionnaire). After all participants in a session had completed the survey, subjects were screened for absolute pitch using a standardized test (Baharloo et al., 1998; see below for details). Following the screening test, an experimenter verbally explained the Go/No-Go task used during the experiment.

Participants were told to use the mouse cursor to click a button on the screen containing the words “PLAY TONE” in order to begin a trial. Clicking the “PLAY TONE” button subsequently and immediately presented a stimulus through the headphones. After the entire duration of the 1000-ms stimulus, the participant was able to click a button labeled “S+” during a 3-second window (see Figure 2.3). If they chose to respond, they received feedback in a box adjacent to the S+ button: either “correct” or “incorrect”, depending on which stimulus was played. If they did not respond, they would receive no feedback. The task was to respond to stimuli that were reinforced (S+) and not to respond to stimuli that lacked reinforcement (S-). If a response was made during an S- trial, the next trial was delayed by 5 s. Participants were told during the experimenter's elaboration of the task that the stimuli presented were placed into separate categories but they were not given any further instruction on how to categorize them.

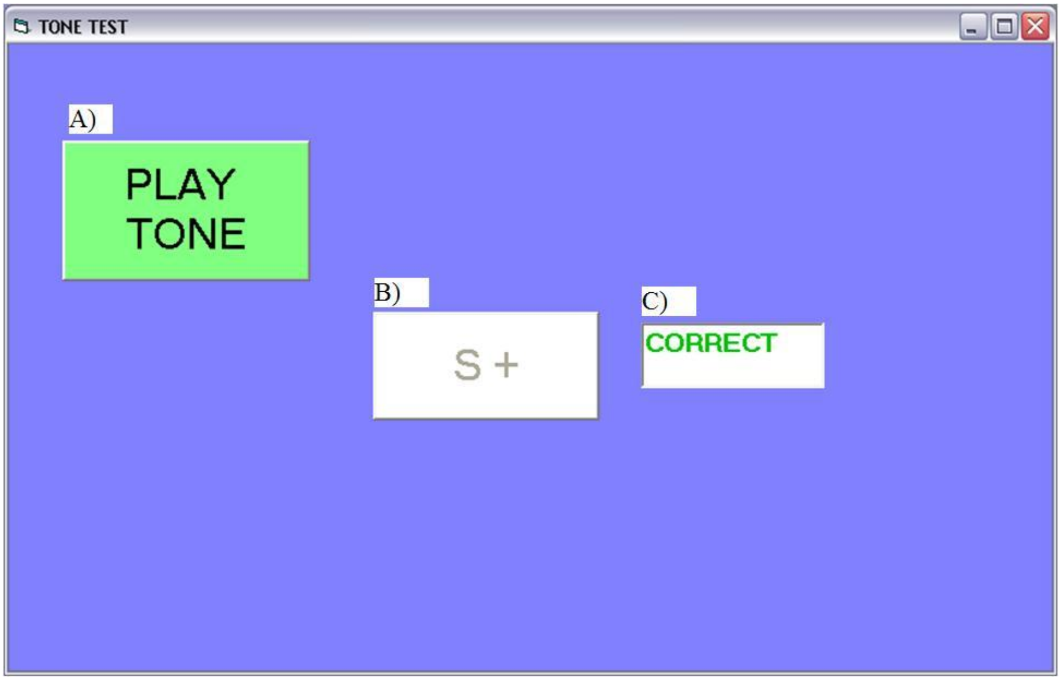


Figure 2.3 The experimental interface participants used for Experiments 1a and 1b. Participants used a mouse cursor to click "PLAY TONE" (A), subsequently make the S+ box (B) turn green. Feedback following a response was displayed in an adjacent white box (C).

Both Experiments 1a and 1b consisted of two phases of 96 trials each. Stimulus presentation was randomized without replacement within each phase of each experiment; that is, each stimulus was randomly presented before a second, randomly-selected iteration of the same stimuli was used.

Response Measures

A discrimination ratio (DR) was used as a primary indicator of performance. The DR is calculated by dividing the percent response for the S+

stimulus by the sum of the percent response to all stimuli (i.e., S+ and S- responses). The DR for each S- stimulus was also calculated by dividing the percent responses for the S+ stimulus by the sum of the percent response for the S+ stimulus plus the S- stimulus of interest. A DR of 0.5 indicates response patterns equivalent to chance; higher than 0.5 indicates more responding to the S+ and lower than 0.5 indicates more responding to the S-. A DR of 1.0 indicates perfect discrimination (i.e., zero responses to the S-). DRs were calculated overall as a basis for performance criterion and also for each S- stimulus to determine differences in performance among the stimuli.

2.3 Experiment 1a: Discrimination of Synthetic Musical Intervals

The goal of Experiment 1a was to use discrimination performance for each stimulus to evaluate quantitative differences in perceiving musical categories of consonance and dissonance for synthetic musical intervals composed of pure tones. Therefore, only the synthetic intervals were used in this study. Reinforcement contingencies were manipulated between subjects to determine, if any, biases in responding resulting from category formation through reinforcement.

Method

Subjects

One-hundred forty-six undergraduate students at the University of Alberta took part in the experiment for course credit. All subjects reported that they had normal hearing. Subjects were screened for absolute pitch (AP, see below); three tested positive for at least Level-1 AP (see Baharloo et al., 1998). These subjects were excluded from final analyses because their abilities are not representative of the general population. The remainder of the subjects was spread relatively equally across conditions (see Table 2.1).

Statistical Analyses and Exclusion Criteria

AP Test. This test has been shown to provide an effective way to measure AP by having human subjects make an instantaneous judgment of the note name and octave number of a randomly selected pure tone within the range of C2 (65 Hz) and B6 (1975 Hz). Subjects are awarded 0.5 points for correctly identifying the name of the note and 0.5 points for indentifying its octave number; 0.25 points is awarded for selecting a note name that is +/- one semi-tone from the note presented. A summation of points awarded after 40 trials determined AP ability: less than 20 points is no AP, between 20 and 25 points is Level 1 AP, and greater than 25 points is Level 2 AP.

Musical Experience. A musical experience index was calculated for each participant based on the number of years of training they had privately (e.g., lessons), in a group (e.g., school band), and how long they had been self taught. The index was calculated by summing the number of years of private lessons with one half the number of years in a group or self-taught in order to weight private instruction more heavily. This procedure was a modified version of one that has been used by Russo et al. (2005).

Performance and Exclusion Criteria. Participants that did not obtain a DR > .80 by the end of the second phase were excluded from final analyses ($N = 2$). One of these two participants seemed to rush through the experiment and responded S+ on nearly all trials (Phase 1 = 90 out of 96 trials, Phase 2 = 95 out of 96 trials). The other participant responded S+ on very few trials during phase 1 (4 out of 96 trials) and expressed confusion during phase 2 (36 out of 96 trials).

Differences in performance for each stimulus across phases were analyzed using 4 (S- stimulus) × 2 (Phase) ANOVAs for each reinforcement group. Post-hoc pairwise comparisons (Bonferroni-corrected) were then used to evaluate performance differences among the S- stimuli. We also determined if there were any effects of stimulus generalization from the first phase of discrimination to the second phase by using single sample *t*-tests comparing the proportion of subjects responding to the first presentation of each S- stimulus (collapsed across stimuli) between phases against chance level (0.5). An alpha level of .05 was used to judge significance of all ANOVAs and an alpha level of .00625 (4 × 2 / .05) was used to judge significance for all pairwise comparisons. All statistical tests were conducted in PASW (version 18, IBM Corporation, Armonk, New York).

Results

Most participants learned the discrimination extremely quickly, regardless of reinforcement condition (see Figure 2.4). Figure 2.4 shows the proportion of participants responding to each S- (non-reinforced) stimulus for the first twelve presentations of both S- and S+ stimuli for participants across all reinforcement groups.. Critically, almost all subjects learned to discriminate the S+ from all of the S- stimuli quite well by trial 4 with exception of one stimulus, which varied with reinforcement group. An explanation of these differences is presented below.

Musical Experience

In order to assess any effect of musical experience on performance, correlations between participants' musicianship index and overall DR's for both phase 1 and phase 2 were conducted. Though most subjects did not have an index greater than 5, the correlations for blocks 1 and 2 were not significant (r = -0.12, *p* = .489; r = .076, *p* = .685, respectively). These correlations were collapsed across all four reinforcement groups.

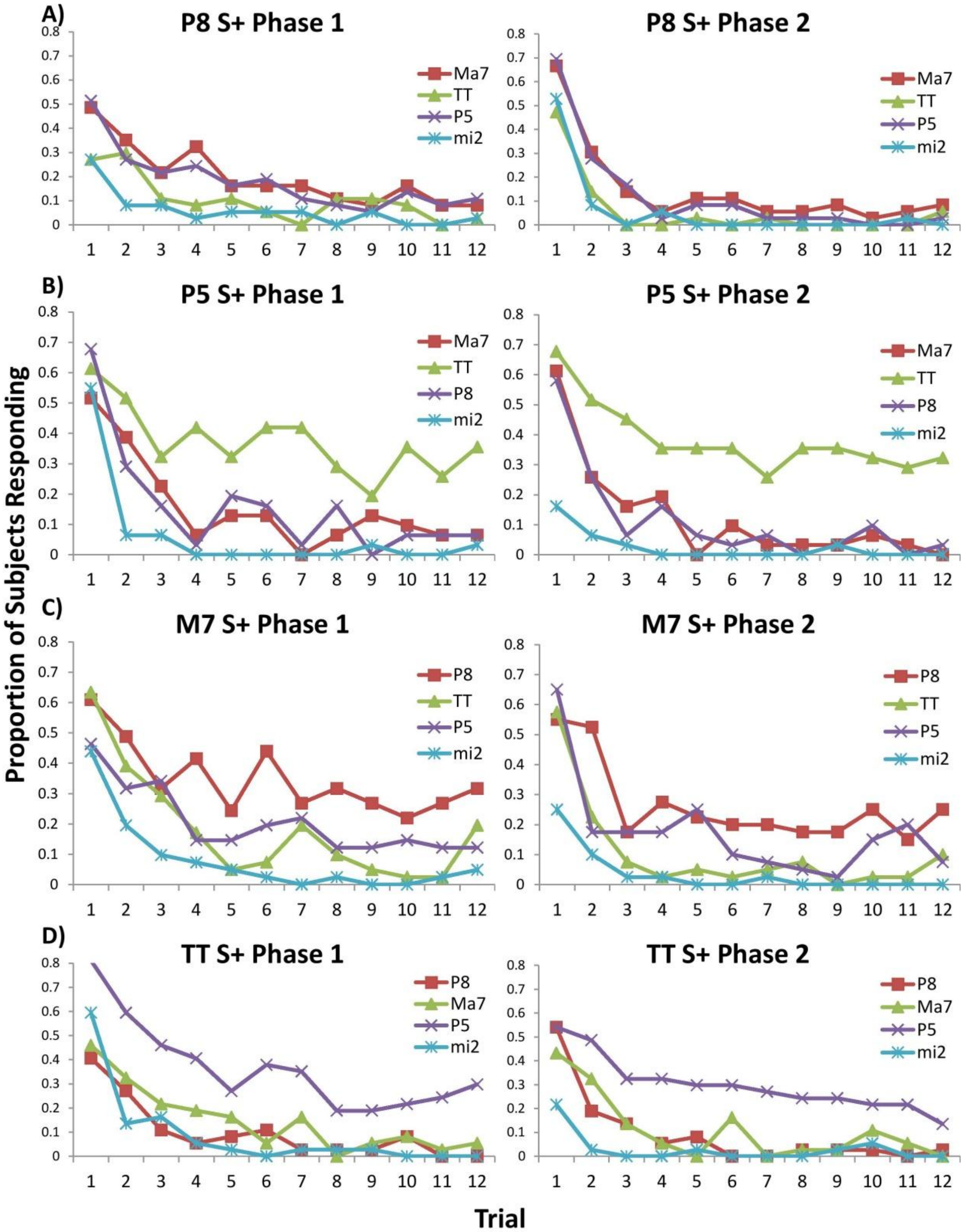


Figure 2.4 Results for groups for human Experiment 1a. Each line on each graph represents a particular S- stimulus. Proportion of subjects responding is plotted on the y-axis and trial number is on the x-axis. A) Phase 1 and 2 for participants reinforced for P8. B) Phase 1 and 2 for participants reinforced for P5. C) Phase 1

and 2 for participants reinforced for M7. D) Phase 1 and 2 for participants reinforced for TT.

Overall Performance

P5 Reinforced. An effect of phase was observed (*F*(1, 93) = 7.115, $p < .012$, partial $\eta^2 = .187$) such that overall performance was significantly greater in phase 2 than in phase 1. There was a main effect of stimulus for participants reinforced for P5 (*F*(3, 93) = 41.709, $p < .001$, partial $\eta^2 = .574$), suggesting that there were performance differences among the non-reinforced stimuli. Pairwise comparisons for phase 1 show that TT (*M* = .74) was discriminated with the least accuracy compared to m2 (*M* = .94, $p < .001$), P8 (*M* = .86, $p < .001$), and M7 (*M* = .86, $p < .001$). Notably, the m2 appeared to be easiest to discriminate. Pairwise comparisons in phase 2 show a similar trend, though the average DRs were higher than in phase 1 overall. This trend is supported by the single sample *t*-tests comparing the proportion of subjects responding during the first trials of each phase: phase 1 showed no significant difference from chance level (*t*(32) = 1.15, $p = .252$) whereas phase 2 discrimination was significantly greater than chance (*t*(32) = 3.41, $p = .031$).

TT Reinforced. Participants in the TT reinforced group also showed a significant effect of phase (*F*(1, 108) = 14.086, $p = .001$, partial $\eta^2 = .281$) and stimulus (*F*(3, 108) = 52.580, $p < .001$, partial $\eta^2 = .594$). Pairwise comparisons for phase 1 showed that performance for P5 (*M* = .73) was significantly poorer compared with m2 (*M* = .92, $p < .001$), P8 (*M* = .91, $p < .001$), and M7 (*M* = .87, $p < .001$). The single sample *t*-tests comparing the proportion of subjects responding during the first trials of each phase again showed no significant difference from chance level ($t(36) = 1.01$, $p = .342$) whereas phase 2 discrimination was significantly greater than chance ($t(36) = 4.41$, $p = .012$).

P8 Reinforced. Participants reinforced for P8 did not show a significant effect of phase ($F(1, 108) = 3.594$, $p = .066$, partial $\eta^2 = .091$) but showed a significant effect of stimulus ($F(3, 108) = 18.270$, $p < .001$, partial $\eta^2 = .337$). Pairwise comparisons for phase 1 showed that performance for the m2 ($M = .95$, $p < .001$) and TT ($M = .91$, $p < .01$) were significantly greater than M7 ($M = .85$, $p < .01$), and P5 ($M = .86$, $p < .01$), with a similar pattern of results in phase 2. However, in both phases, performance between M7 and P5 were not different from each other. The single sample *t*-tests comparing the proportion of subjects responding during the first trials of each phase again showed no significant difference from chance ($t(36) = 1.41$, $p = .271$) whereas phase 2 discrimination was significantly greater than chance ($t(36) = 3.92$, $p = .015$).

M7 Reinforced. Participants reinforced for P8 showed a significant effect of phase ($F(1, 120) = 24.906$, $p < .001$, partial $\eta^2 = .384$) and stimulus ($F(3, 120) = 45.919$, $p < .001$, partial $\eta^2 = .534$). Pairwise comparisons for phase 1 showed that performance for m2 ($M = .92$, $p < .001$), TT ($M = .85$, $p = .02$), and P5 ($M = .84$, $p = .02$) were significantly greater than P8 ($M = .74$, $p < .01$). The single sample *t*-test for phase 1 showed no significant difference from chance ($t(40) = 1.50$, $p = .281$) whereas phase 2 discrimination was significantly greater than chance ($t(40) = 2.92$, $p = .036$).

2.4 Experiment 1b: Discrimination of Piano Musical Intervals

The main goal of Experiment 1b was to determine if subjects would show a similar pattern of responding to the same intervals but with a different spectral structure (i.e., timbre). Therefore, the only difference in Experiment 1b was that the piano stimulus set was employed rather than the pure tone stimulus set. The presence of higher harmonics within the stimuli used for Experiment 1b should facilitate consonance/dissonance to be a more salient cue for auditory discrimination. It was thus expected that non-reinforced stimuli more closely related by consonance or dissonance to reinforced stimuli would have a higher proportion of errors relative to other non-reinforced stimuli.

Method

Subjects

One-hundred twenty-three undergraduate students at the University of Alberta took part in the experiment for course credit. As in Experiment 1a, all participants reported that they had normal hearing and were subject to an AP test; no participant possessed any significant level of AP.

Apparatus, Procedure, Stimuli, and Analyses

The same apparatus, design, and procedure was employed for Experiment 1b as in Experiment 1a. The only difference was the stimuli used for this experiment which were the musical intervals composed of piano tones rather than synthetic pure tones. The same statistical analyses from Experiment 1a were employed in Experiment 1b. Again, all analyses used an alpha level of .05 to judge statistical significance and were performed using PASW v.18.

Results

As in Experiment 1a, most participants learned the discrimination extremely quickly, regardless of stimulus or reinforcement type. Most participants had a DR greater than .80 by the fourth trial (see Figure 2.5). One participant did not meet criterion of DR = .80 and two participants had a DR < .60 for a particular stimulus.

Musical Experience

Again, the correlations between the musicianship index and DRs for blocks 1 and 2 were not significant (r = -0.07, $p = .784$; r = .100, $p = .582$, respectively).

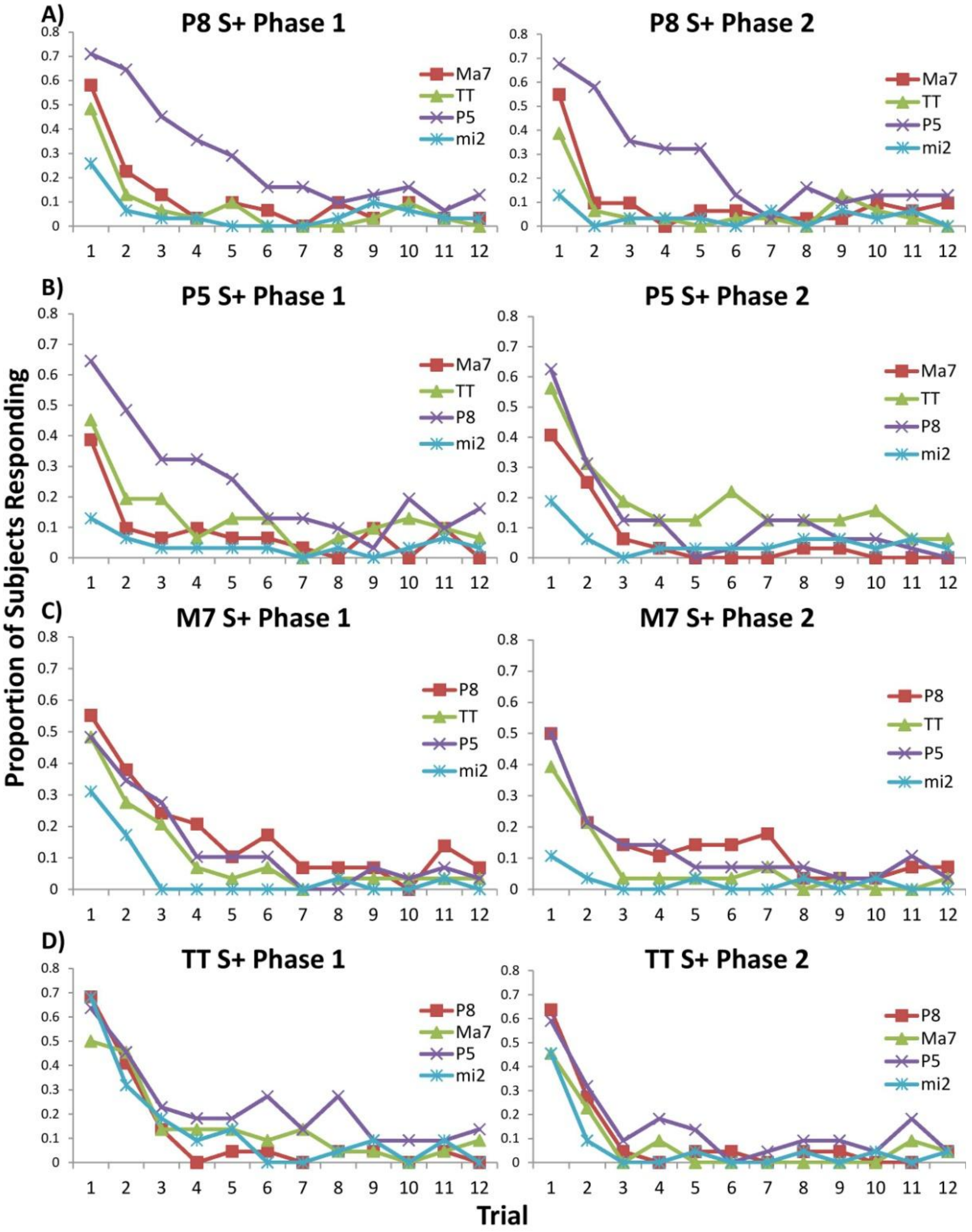


Figure 2.5 Results for groups for human Experiment 1a. Each line on each graph represents a particular S- stimulus. Proportion of subjects responding is plotted on the y-axis and trial number is on the x-axis. A) Phase 1 and 2 for participants reinforced for P8. B) Phase 1 and 2 for participants reinforced for P5. C) Phase 1 and 2 for participants reinforced for M7. D) Phase 1 and 2 for participants reinforced for TT.

Overall Performance

Figure 2.5 shows data for participants across all reinforcement groups. Almost all subjects learned to discriminate the S+ from all of the S- stimuli quite well by trial 4, though the discrimination performance across stimuli tended to differ between reinforcement conditions. For P5 and P8 S+ conditions, the proportion of errors across time is higher for one particular S- stimulus relative to all other S- stimuli. The TT and M7 S+ conditions do not show this divergent effect. An explanation of these results is presented below.

P5 Reinforced. No effect of phase was observed ($F(1, 90) = 2.115$, $p = .089$, partial $\eta^2 = .117$) but there was a main effect of stimulus for participants reinforced for P5 ($F(3, 90) = 20.389$, $p < .001$, partial $\eta^2 = .405$) for both phases, suggesting that there were differences in performance among the stimuli over time. Pairwise comparisons for phase 1 show that P8 ($M = .80$) was discriminated with the least accuracy compared to m2 ($M = .95$, $p < .001$), TT ($M = .88$, $p < .001$), and M7 ($M = .91$, $p < .001$). Pairwise comparisons in phase 2 show a similar trend, though the average DRs were higher than in phase 1 overall. This trend is supported by the single sample *t*-tests comparing the proportion of subjects responding during the first trials of each phase: phase 1 showed no significant difference from chance ($t(32) = 1.15$, $p = .252$) whereas phase 2 discrimination was significantly greater than chance ($t(32) = 3.41$, $p = .031$).

TT Reinforced. Participants in the TT reinforced group showed a significant effect of phase ($F(1, 87) = 8.245$, $p < .01$, partial $\eta^2 = .297$) and stimulus ($F(3, 87)$

= 17.472, $p < .001$, partial $\eta^2 = .376$). Pairwise comparisons for phase 1 showed that performance for P5 ($M = .82$) was significantly poorer compared with m2 ($M = .88$, $p < .001$), P8 ($M = .90$, $p < .001$), and M7 ($M = .87$, $p < .001$). Notably, performance overall was poorer than results from the P5 reinforcement condition. Single sample *t*-tests comparing the proportion of subjects responding during the first trials of each phase again showed no significant difference from chance level ($t(36) = 1.01$, $p = .342$) whereas phase 2 discrimination was significantly greater than chance ($t(36) = 4.41$, $p = .012$).

P8 Reinforced. Participants reinforced for P8 also showed a significant effect of phase ($F(1, 90) = 10.283$, $p < .01$, partial $\eta^2 = .401$) and stimulus ($F(3, 90) = 53.169$, $p < .001$, partial $\eta^2 = .639$). Pairwise comparisons for phase 1 showed that performance for m2 ($M = .96$, $p < .001$), TT ($M = .93$, $p < .01$), and M7 ($M = .90$, $p < .01$) were significantly greater than P5 ($M = .79$, $p < .01$), with a similar pattern of results in phase 2. Compared with participants in the P8 reinforcement condition in Experiment 1, performance for the P5 was markedly poorer and performance for P8 was substantially better. The single sample *t*-tests for the first trials of each phase again showed no significant difference from chance ($t(36) = 1.41$, $p = .271$) in phase 1 whereas phase 2 discrimination was significantly greater than chance ($t(36) = 3.92$, $p = .015$).

M7 Reinforced. Finally, participants reinforced for M7 did not show a significant effect of phase ($F(1, 93) = 4.115$, $p = .077$, partial $\eta^2 = .152$) but showed a significant main effect of stimulus ($F(3, 90) = 17.663$, $p < .001$, partial $\eta^2 = .371$). Pairwise comparisons for phase 1 showed that performance for m2 (*M*

= .95, $p < .001$) was best, and the descending order of difficulty for other stimuli were TT ($M = .89$), P5 ($M = .88$), and P8 ($M = .84$). Discrimination performance for m2 was significantly greater than the other non-reinforced stimuli, but there were no significant differences among the remaining three. The single sample *t*-test for phase 1 showed no significant difference from chance ($t(40) = 1.50$, $p = .281$) whereas phase 2 discrimination was significantly greater than chance ($t(40) = 2.92$, $p = .036$).

2.5 Discussion

My experiments showed that humans are able to discriminate musical intervals of varying consonance with the same root note. Participants were able to discriminate the intervals minor second (m2), tri-tone (TT), perfect fifth (P5), major seventh (M7) and octave (P8) in a Go/No-go task with remarkable accuracy in a small number of trials. This overall level of expertise was prevalent regardless of the reinforcement contingency and the spectral composition of the stimuli. This is not surprising as previous work, that has used a similar task with triadic chords, showed analogous rates of learning among human subjects (Hoeschele et al., 2012). However, comparing participants' expertise with regard to individual stimuli suggests more complex patterns of learning which did indeed differ between reinforcement contingency and timbre.

Experiment 1a

Experiment 1a sought to elucidate discrimination of synthetic musical intervals composed of pure tones. Because pure tones do not possess harmonic frequencies that elicit dissonance due to the presence of beating, it was expected that consonance/dissonance would not be a salient cue for discrimination purposes. Thus, S- stimuli that are related by consonance (e.g., P5 and P8) or dissonance (e.g., TT and M7) to the S+ stimulus should not have been mis-categorized and therefore should have generated few errors. Generally, participants across all reinforcement conditions followed this trend.

Discrimination was significantly above chance by around the third presentation for S- stimuli related by consonance or dissonance. However, discrimination was also quite high by the third presentation for most S- stimuli in general with one exception. For P5 S+, M7 S+, and TT S+ groups, performance for one S- stimulus appeared significantly worse relative to all other stimuli. Participants in the P5 S+ group made the most erroneous responses to the TT S- stimulus. The reverse was true for participants in the TT S+ group: they made the most errors for the P5 S- stimulus. Those in the M7 S+ group had a significantly higher proportion of errors for the P8 S- stimulus. The P8 S+ group showed the poorest discrimination performance for the M7 S- stimulus, though this difference was less pronounced than the other three groups. Differential responses based on a given participants’ reinforcement condition is telling in two ways: (1) it shows no biases towards the absolute frequency relationships between the two notes of a given interval and (2) the S- stimuli that hold the highest proportion of errors across reinforcement groups are all related based on the relative pitch of the upper note compared with the S+ stimulus. In other words, S- stimuli with an upper note that was close in pitch to the upper note of the S+ stimulus were most difficult to discriminate in the synthetic stimuli condition.

This finding is somewhat perplexing as humans are able to discriminate the pitch between two pure tones within the frequency range of our stimulus set to a degree of 1 Hz (Handel, 1989). Although most participants learned to discriminate all S- stimuli equally across many trials, the errors at the onset of learning for both phases show that discrimination of synthetic musical intervals is more difficult when the upper note of the comparison interval is close in pitch. Previous research has shown that human pitch discrimination is quite accurate in both musicians and non-musicians to a magnitude of less than one semitone (Tervaniemi et al., 2005). The smallest frequency difference between the upper notes of the synthetic stimuli used was roughly 12 Hz (F#4 in TT on the C4 root note and G4 in P5 on the C4 root note). However, each stimulus used in the current study consisted of a root note and an upper note, so it is possible that the root note served to interfere with the perception of distinct categories based on pitch alone. We did not find any correlations related to performance and musical training, though previous work indicates distinct categorical perception of musical intervals that deviate by a semitone or less in trained musicians (Burns & Ward, 1978). Burns and Ward also tested untrained musicians who showed that the discrimination threshold was roughly one semitone. Thus, the perceived difference between intervals with the same root note, but with upper notes that deviate by one semitone, should not be salient for untrained listeners as the difference in pitch is close to or at threshold. The absence of significant correlations between performance and musicianship may be due to the relatively few participants with substantial musical experience in the current study.

The higher proportion of errors for S- stimuli whose upper note is closely related in pitch to that of the S+ stimulus suggests that participants across all conditions tended to focus on the pitch of the upper note of each S- stimulus to solve the task. Conversely, the relatively rarity of errors for S- stimuli related by consonance/dissonance to the S+ stimulus suggest that participants did not attend to the relative harmonicity of the intervals. Given this, it is logical to conclude that consonance/dissonance did not act as a salient cue for discriminating these synthetic intervals.

Experiment 1b

Experiment 1b used stimuli with the same manipulations as Experiment 1a except the Experiment 1b stimuli possessed higher harmonic frequencies. The addition of these harmonics using piano tones that comprised the same intervals in Experiment 1a should have increased the relative salience of consonance and dissonance. This was examined by comparing relative rates of error across the various non-reinforced intervals in the Go/No-go task.

Again, participants were able to discriminate all intervals with great accuracy and in relatively few trials. As shown by Figure 2.5, a trend similar to that in Experiment 1a emerged in that one particular stimulus had a significantly higher proportion of errors relative to other S- stimuli, but this was only true for the P5 S+ and P8 S+ groups. In the M7 S+ and TT S+ groups, discrimination among all S- stimuli seemed to be fairly consistent. However, the specific S- stimuli that were the most difficult to discriminate in the P5 S+ and P8 S+ groups were different than those for the same reinforcement groups in Experiment 1a.

Participants in the P5 S+ and P8 S+ groups had higher error rates when comparing stimuli closely related in consonance/dissonance. For example, participants in the P5 S+ group had the highest proportion of errors for the P8 S- stimulus relative to all other S- stimuli; participants in the P8 S+ group had the highest proportion of errors for the P5 S- stimulus. It is interesting to note, however, that the M7 S+ and TT S+ groups did not show this trend. For both groups, the highest proportion of errors across all trials was attributed to the S- stimulus that had an upper note closest in pitch to the S+ stimulus. Thus, it is only the participants reinforced with consonant intervals that seemed to use the relative consonance/dissonance as a cue for auditory discrimination more so than the pitch of the upper notes.

These findings suggest that the relative consonance/dissonance is a salient cue for auditory discrimination but only when participants are reinforced for consonant intervals and only when complex, harmonically rich stimuli are discriminated. Much research has suggested an innate preference for consonance in humans (e.g., Trainor, Tsang, & Cheung, 2002; Butler & Daston, 1968), though this is, to our knowledge, the first example in which reward contingencies for consonant and dissonant stimuli in a learning paradigm affect performance differentially. One plausible explanation for these differences in strategies to solve the task may be related to exposure through enculturation. For example, it may be that because much Western music utilizes consonant harmonies involving perfect fifths and octaves, listeners perceive intervals related by consonance as more similar than intervals related by dissonance. This may be true regardless of musicianship, as we did not find significant correlations between performance and musical experience for participants in Experiment 1b.

Taken together, these two experiments suggest that timbre modulates whether the relative consonance or the pitch of the upper note of musical intervals is a more salient cue for auditory discrimination. Listeners exposed to synthetic intervals attended to the pitch of the upper notes of the stimuli they were discriminating. Listeners exposed to piano intervals attended to the relative consonance among the stimuli as a cue for discrimination, but only when the reinforced interval was consonant. These findings are in line with Helmholtz's (1885) widely accepted explanation of consonance in that it is distinguished from dissonance by the absence of beating. However, these experiments are the first to show the relative saliency among pitch relations with and without harmonics in a learning paradigm. The data presented here suggest that this effect is not determined by musical experience. However, as consonance and dissonance are both extremely common in Western music and media, the answer may lie within culture and development. Further studies may discern spectral and temporal features of timbre that contribute to this saliency switch and the degree to which experience with Western music affects consonance/dissonance perception.

Chapter 3

Discrimination of Synthetic and Piano Musical Intervals in Black-Capped Chickadees (*Poecile atricapillus*)

3.1 Introduction

Using an animal model to investigate music perception has led to telling conclusions about the potential adaptive significance of music. This is because animals are thought to lack music as it is defined by human cultures. As such, music-related traits found in animals through tasks that gauge musical sound preferences, discriminations, and generalizations are likely to represent a general-purpose mechanism that may have evolved long ago. Though most studies comparing human and non-human music perception have emphasized their differences rather than potential homologies, some studies are suggesting that the rules that govern Western music may have a more biological basis.

Pitch, addressed previously as being one of the well-studied phenomena in humans, is also well-studied in songbirds (Weisman et al., 1990; Doupe & Kuhl, 1999). In many cases, songbirds rely on absolute pitch more readily than relative pitch to solve certain tasks (Cynx, Hulse, & Polyzois, 1986) and even outperform many other species (including humans) on tasks that require the use of solely absolute pitch (Weisman et al., 1998; Weisman, Njegovan, Williams, Cohen, & Sturdy, 2004). As in humans, songbirds can further develop both relative and absolute pitch abilities through rehearsal (Weisman & Ratcliffe, 1989). Little is known about how passerine songbirds use these abilities to process musical intervals, the combination of two pitches. Furthermore, less is known about how chickadees process the relative consonance/dissonance of these intervals, a resulting property of the two pitches.

Experiments 2a and 2b employ black-capped chickadees (*Poecile* atricapillus) in an operant discrimination task. Black-capped chickadees were chosen for a few critical reasons: (1) They have an evolutionarily complex auditory system in common with other songbirds that is strikingly similar to the mammalian auditory system (Durand, Tepper, & Cheng, 1992; Hughes, Nowicki, & Lohr, 1998; Wooley & Cassady, 2003; Mooney, 2009), (2) there is substantial literature illustrating their pitch discrimination abilities in similar tasks (e.g., Weisman & Ratcliffe, 1989; Guillette et al., 2010; Hoeschele, Guillette, & Sturdy, 2012), and (3) their survival depends heavily on perceiving biologically relevant acoustic stimuli in their natural habitat. The third point allows us to compare their responses with humans to non-biologically relevant stimuli – in this case, synthetic musical intervals. Taken together, studying black-capped chickadees offers a unique opportunity for a comparative approach toward understanding if sounds defined by Western music are perceived the same way as in humans.

In Experiments 2a and 2b, I tested chickadees using the same stimuli in an operant discrimination paradigm that is directly comparable to that employed in the first human experiments and has been shown to be an effective method for determining differences between stimulus responses (see Hoeschele et al., 2012). Previous research has shown that songbirds make reliable discriminations faster and more accurately when note pairs are linked by a common frequency ratio (Weisman, Njegovan, & Ito, 1994) and that they are able to determine differences in consonant and dissonant triads (Watanabe, Uozumi, & Tanaka, 2005). It is unknown whether these performance differences modulate with timbre, though the property is an extremely important cue for discrimantion (Cynx, Williams, & Nottebohm, 1990). As such, we expected in Experiments 2a and 2b that chickadees would discriminate stimuli with harmonics faster than those without and that the pattern of responses across the stimuli would be different than humans.

3.2 General Method

Animals

Twenty-five black-capped chickadees (*Poecile atricapillus*, 13 female and 12 male – gender identified by DNA analysis; Griffiths et al., 1998) at least one year of age (determined by the shape and color of their outer tail retrices; Pyle, 1997) were used in both experiments conducted between January and June, 2013.

These birds were captured from sites in and around the city of Edmonton, Alberta, Canada (53° 34’ N, 113° 25’ W) and the Barrier Lake Field Station in Kananaskis County, Alberta, Canada (51° 02’ N, 115° 03’ W) between January, 2009 and March, 2012. Prior to the experiments, all birds were housed individually in Jupiter Parakeet cages (30 × 30 × 40 cm; Rolf C. Hagen, Inc., Montreal, Canada) in colony rooms on a natural light cycle and had *ad libitum* access to lab food (Mazuri Small Bird Maintenance Diet; Mazuri, St. Louis, MO), water (supplemented with vitamins three times a week, Hagen, Rolf C. Hagen, Inc., Montreal, Canada), grit, and cuttle bone. Birds were continuously given a mixture of greens and eggs twice a week and one superworm (*Zophobas morio*) three times a week as supplements. During the experiments, however, access to lab food was provided subsequent to an appropriate operant response (access to water, grit, and cuttle bone remained *ad libitum*). Additionally, one superworm in the morning and afternoon was given to each bird every day – the mixture of greens and eggs was not provided during the experiments. All birds had previous experience with auditory discriminations using similar operant paradigms though they were naïve to the stimulus sets used in the current experiment.

Apparatus

During both experiments, each bird was housed in its own modified colony room cage with three perches and dispensers for water and grit. Each cage was placed inside a ventilated sound-attenuating chamber illuminated by a 9-W full-spectrum fluorescent light bulb. One wall of the cage had an opening (11 × 16 cm) that gave access to a motor-driven feeder (Njegovan, Hilhorst, Ferguson, &

Weisman, 1994). Infrared cells placed in the feeder and on the perch closest to the feeder tracked the position of the bird – the latter of which enabled the bird to initiate a trial. Next to the feeder was either a Fostex FE108Σ or Fostex FE108EΣ full-range speaker (effective frequency response of 200-16,000Hz, though both models are flat in the frequency range of interest; Fostex Corp., Japan) pointed into the cage that presented the stimuli between 70 and 75dB(A) from the perch that required a bird to initiate a trial (calibrated using the same Brüel & Kjær Type-2239 sound level meter from the human experiments). Trials in the chamber were scheduled automatically by a computer and single-board unit (Palya & Walter, 2001) which also recorded responses. The computer read and presented the stimulus files from one of its optical CD drives and was powered by either a Cambridge A300 Integrated Amplifier (Cambridge Audio, London, England) or a NAD 310 Integrated Amplifier (NAD Electronics, London, England) and connected to the speaker inside of the chamber (see Sturdy & Weisman, 2006 for details).

Stimuli

The same stimulus sets used in the human Experiments 1a and 1b were used in the chickadee Experiments 2a and 2b, respectively. The only difference is that 1000 ms of silence was added to the end of each stimulus using SIGNAL version 5.10.24 (Engineering Design, Berkeley CA). This change made it possible to use the stimuli on CDs as part of our standard protocol without altering the native attack and steady state portion of the stimuli.

Design

The design of both chickadee experiments was similar to that of the human experiments. Experiment 2a tested chickadees using synthetic stimuli and Experiment 2b tested chickadees using piano stimuli. The one difference, however, was that all of the reinforcement conditions in the human experiments were not present in the chickadee experiments. For Experiments 2a and 2b, only the P5 and TT were treated as S+ stimuli. These conditions were chosen to ensure that at least one consonant stimulus (P5) and one dissonant stimulus (TT) shared reinforcement roles between individuals while moderating the total number of chickadees used for both experiments.

Procedure

The Go/No-Go task used in the human experiments was employed in the current experiment using a directly comparable paradigm. Birds were rewarded with food for correctly responding to an S+ stimulus and were punished with a timeout and darkness in their respective chambers for 30 s following an incorrect response to an S- stimulus. This basic reward/punishment procedure was applied, most critically, to the discrimination phases (see below).

Non-differential Training. After each bird underwent our standard shape training procedures, ensuring appropriate use of the request perch and feeder, non-differential training began. The goal of this phase was to ensure that all birds heard and attended to all stimuli with which they would later discriminate equally, therefore diminishing possible response biases.

A bird initiated a trial by landing on the request perch, subsequently breaking the infrared beam, and remaining there for an average of 1 s (ranging between 900 and 1,100 ms). After which, one of the 10 stimuli that a bird would hear in its subsequent discrimination phases was presented randomly and without replacement. The bird would be reinforced with access to food for 1 s if it remained on the request perch for the duration of the stimulus and then moved to the feeder within 1 s after stimulus offset. Reinforcement was followed with an inter-trial interval (ITI) of 30 s with the chamber light turned on, during which no new trials could occur. Interrupted trials in which the bird failed to remain on the request perch for the duration of the stimulus resulted in an ITI of 30 s with the chamber lights turned off. This was to ensure that birds listened to the entire stimulus before making a response. If the bird did not move to the feeder during the 1-s reinforcement window after stimulus offset, a new trial could initiate after moving away from the request perch and then back onto it or by remaining on the request perch for an ITI of 60 s. This long ITI was to promote engagement in the task and discourage sitting on the request perch to initiate trial after trial. This also indirectly increased the probability that a response was made on a given trial. Once all 10 stimuli were played in a random order, they were played again in a novel random order. This sequence repeated until the bird finished the non-differential training phase.

Data from this phase was analyzed in blocks of 500 trials. In order to progress to the first discrimination phase, each bird had to meet two criteria: (1) respond by moving to the feeder on a high proportion of trials ($\geq$ 60%) for six blocks and (2) have less than or equal to a 3% difference in responding between all future S+ and S- stimuli for four blocks. The first criterion was to foster adequate responding and the second criterion was to eliminate stimulus response biases among future reward contingencies. Once these criteria were met, a bird progressed to the first discrimination phase.

Discrimination. Similar to the human experiments, there were two phases of discrimination and each were identical except for the stimuli presented in each phase. The root note with which the interval was composed was again counterbalanced within all timbre and reinforcement conditions. The second phase of discrimination contained the stimuli composed from the root note that the bird did not discriminate during the first phase (C4 or E4). Therefore, as in the human experiments, each phase of discrimination required a bird to discriminate one S+ stimulus from the four S- stimuli.

The same procedures for non-differential training were used in discrimination with a few crucial differences. One difference was that rather than being reinforced for a response to every stimulus, birds were reinforced only for responding by moving to the feeder after the presentation of one particular stimulus depending on which treatment group they belonged to. A second difference was that the lights went out in the chamber for 30 s following an incorrect Go response to any of the S- stimuli. Finally, identical to the human experiments, the S+ stimulus was presented 50% of the time and all four S- stimuli comprised the other 50% of stimulus presentations (12.5% each).

Data from each discrimination phase were analyzed in 480-trial blocks. Progressing from the first phase to the second phase of discrimination and from the second phase to completing the experiment required meeting one of two criteria: (1) an overall DR of 0.8 or greater for six blocks with the last two blocks being consecutive or (2) reaching a total of 36,000 trials (75 blocks).

3.3 Experiment 2a: Discrimination of Synthetic Musical Intervals

The purpose of Experiment 2a was to determine differences in the rates and level of expertise at which chickadees discriminate synthetic musical intervals that vary in consonance/dissonance. The same synthetic stimuli used for human Experiment 1b were used in this study. Furthermore, the same Go/No-Go task was employed using an apparatus adapted for chickadee use in order to be directly comparable to results obtained from the human experiments.

Method

Animals

Twelve black-capped chickadees (*Poecile atricapillus*, six female and six male) were used in this experiment. The six female and six male chickadees were dispersed evenly across reinforcement conditions P5 S+ and TT S+ for non-differential training and discrimination phases.

Statistical Analyses

Because chickadees had great variability in the number of trials to reach either criterion, a normalizing procedure (i.e., Vincentizing) was applied to each bird’s acquisition curve in order to compare chickadee responses both across reinforcement conditions and with the human data from Experiment 1a. Vincentizing was calculated by taking the total number of bins required to meet criterion and dividing this value by eight vincentized blocks (i.e., the fewest number of blocks an individual bird took to reach criterion). The quotient of this number became the averaging quantile determining the total number of blocks to average per Vincentized block. The remainder of this number became the number of blocks added to the average of the middle two Vincentized blocks (i.e. blocks 4 and 5; see Kling & Riggs, 1971). For example, a chickadee that took 26 bins to reach criterion would have three bins averaged together in each Vincentized block except blocks 4 and 5, which would have four bins. This is because dividing 26 by 8 into whole numbers (in this example, 3 remainder 2) would make the quotient (3) the grouping size for each Vincentized block. The remainder (2) is dispersed evenly between blocks 4 and 5 (i.e, both the 4th and 5th Vincentized block would have 3 + 1 = 4 actual blocks averaged together).

To determine any potential differences in responding due to the sex of the birds or the absolute frequency values of the stimuli, we first conducted a 2 (Sex) × 2 (Root note) ANOVA (collapsed across reinforcement condition) using the number of blocks to reach criterion as a measure of overall performance. For each reinforcement condition, we conducted a three-way ANOVA comparing the DR for Phase × Vincentized blocks × S- Stimulus to determine differences in the rates of learning and the degree of expertise between the non-reinforced stimuli. Bonferroni-corrected pairwise comparisons were then performed to determine differences in performance among the S- stimuli. Finally, an analysis of savings from phase 1 to phase 2 of discrimination was performed by comparing the linear regression coefficients of each individual’s Vincentized data for each S- stimulus using a 4 (S- Stimulus) × 2 (Phase) ANOVA. All analyses were conducted using

PASW v.18 and an alpha level of 0.05 was used to determine statistical significance in all tests.

Results

The ANOVA using the number of blocks to reach criterion as a measure of overall performance revealed that there were no significant differences of sex (*F*(1,19) = .141, *p* = .711) or root note (*F*(1,19) = .156, *p* = .697). There were only two chickadees that did not meet criterion of DR ≥ 0.8 for six consecutive bins during the first phase of discrimination.. These two chickadees also did not reach the DR criterion during the second phase of discrimination. However, an analysis of performance during the first discrimination phase using independent samples *t*-tests for the first bin and the last bin of these two birds compared with those in the pure tone condition that met criterion before reaching 75 bins revealed no differences in performance (*t* = 0.602, *p* = 0.564; *t* = -1.080, *p* = 0.311, respectively). Furthermore, the response patterns of the birds that did not meet the DR performance criterion were similar to birds that did by the end of discrimination training (see Figure 3.1).

Figures 3.1 shows Vincentized acquisition curves of each S- stimulus across both phases for chickadees in the P5 S+ (top panel) and TT S+ (lower panel) conditions. On average, chickadees learned to discriminate all S- stimuli from the S+ significantly above chance by the end of each phase regardless of reinforcement condition. However, chickadees that were reinforced for P5 had the lowest DRs specifically for the TT across all eight blocks and chickadees reinforced for TT had the lowest DRs for the P5 across all eight blocks. All other

S- stimuli for chickadees in both reinforcement conditions were discriminated at a similar level of expertise with the exception of m2 in the TT S+ condition. An elaboration of these differences in rate and proficiency of learning between the two reinforcement conditions are presented below.

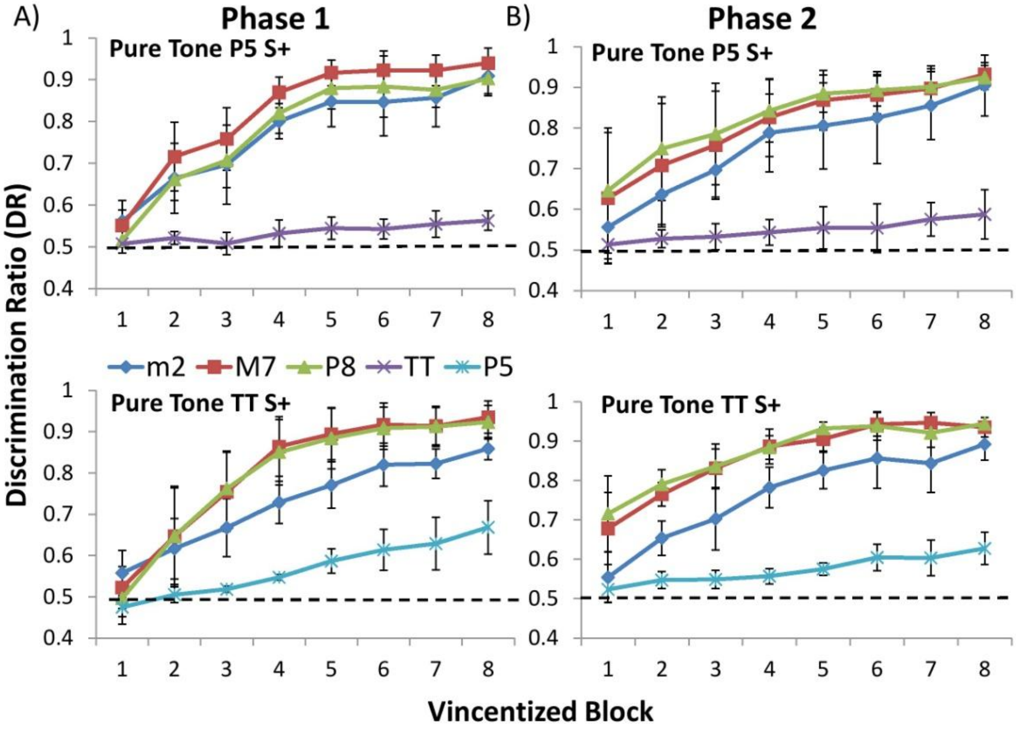


Figure 3.1 Average Discrimination Ratio (DR) for each S- stimulus across the eight Vincentized blocks for chickadees in the synthetic condition (see text for details). Panel (A) shows phase 1 and phase 2 chickadee discrimination for the P5 S+ group and panel (B) shows phase 1 and phase 2 discrimination for the TT S+ group. Error bars show 95% confidence intervals.

P5 S+ Group

Acquisition. In the ANOVA comparing DRs of birds for Phase × Vincentized block × S- stimulus, we found a main effect of block such that DRs increased as Vincentized block increased (F (7, 105) = 33.691, $p < .001$, partial $\eta^2 = .871$). Furthermore, there was a significant interaction of S- stimulus × Vincentized block (*F* (21, 105) = 15.401, $p < .001$, partial $\eta^2 = .755$) and main effect of S-stimulus (*F* (3, 105) = 53.329, $p < .01$, partial $\eta^2 = .914$), suggesting that there were differences in performance among the non-reinforced stimuli and that these differences varied over time. Pairwise comparisons of Vincentized data of DRs revealed that chickadees' discrimination of TT (*M* = .54) was significantly poorer than m2 (*M* = .77, $p < .01$), M7 (*M* = .82, $p < .01$), and P8 (*M* = .80, $p < .01$) over time. Indeed, panel A of Figure 3.1 shows that the TT was by far the most difficult for birds to discriminate across the eight Vincentized blocks. Differences among the other S- stimuli were not significant, though it appears that M7 was learned the fastest and performance for m2 and P8 were about equal.

Savings. To examine effects of stimulus generalization between phase 1 and phase 2, we compared the intercepts of each bird's Vincentized DR data between Phase 1 and Phase 2 using a Stimulus × Phase ANOVA. The intercepts between the S- stimuli seemed to vary significantly (*F* (3, 15) = 3.693, $p = .036$, partial $\eta^2 = .425$) but the effect of phase was not significant (*F* (1, 15) = .554, $p = .490$). In Figure 3.1, one can detect that the chickadees were responding significantly above chance in the first Vincentized block but only for certain stimuli; savings effects seem to be most substantial for P8, M7, and m2 (ordered from strongest to weakest), while discrimination for TT was only slightly better than chance level.

TT S+ Group

Acquisition. Chickadees in the TT S+ group also showed increasing DRs with increasing Vincentized block (*F* (7, 105) = 63.956, $p < .001$, partial $\eta^2 = .927$). There was also a significant interaction of S- stimulus × Vincentized block (*F* (21,

105) = 7.297, $p < .001$, partial $\eta^2 = .593$) and main effect of S- stimulus (*F* (3, 105) = 95.259, $p < .001$, partial $\eta^2 = .950$), indicating similar performance differences over time. The pairwise comparisons between the various S- stimuli showed that chickadees' discrimination of P5 ($M = .57$) was significantly poorer than m2 ($M = .75$, $p = .012$), M7 ($M = .83$, $p < .022$), and P8 ($M = .83$, $p < .022$) over time. Furthermore, m2 was significantly more difficult for birds to discriminate than P8 and M7 over time. These performance differences among the S- stimuli held for both phases, as indicated by panel B of Figure 3.1.

Savings. The intercepts of the Vincentized data among S- stimuli also varied significantly (*F* (1, 15) = 6.994, $p < .01$, partial $\eta^2 = .583$). The main effect of phase was significant (*F* (1, 15) = 5.667, $p < .01$, partial $\eta^2 = .456$), such that performance at the beginning of the second phase was better than at the beginning of the first phase. Figure 3.1 shows much variability in performance in the first Vincentized block during the second phase, but shows that savings were greatest for P8 followed by M7, m2, and TT. Notably, DRs for both m2 and TT are not significantly above chance during the first Vincentized block of phase 2.

3.4 Experiment 2b: Discrimination of Piano Musical Intervals

Similarly to human Experiment 1b, chickadee Experiment 2b aimed to elucidate whether similar patterns of response would emerge when the spectral complexity of stimuli was manipulated. Therefore, the same procedure, apparatus, and design from chickadee Experiment 2a were employed in this experiment. The only difference was that chickadees were exposed to the piano stimuli.

Method

Animals

Thirteen black-capped chickadees (six female and seven male) were used in this experiment. Additionally, one bird completed the first phase of discrimination but failed to complete the second phase of discrimination due to an equipment failure; the data from phase one only was included in the analysis. Therefore, there were seven chickadees (three female, four male) in the P5 S+ group and six chickadees (three female, three male) in the TT S+ group, though the seventh chickadee in the P5 S+ group only contributed to phase 1 discrimination data.

Statistical Analyses

A Vincentized analysis was again employed in order to directly compare responses to chickadees from Experiment 2a and humans from Experiment 1b. All statistical analyses conducted in Experiment 2a were also conducted in Experiment 2b (see Experiment 2a for details of analyses and Vincentizing procedure). All analyses were conducted using PASW v.18 and an alpha level of 0.05 was used to determine statistical significance in all tests.

Results

The ANOVA comparing overall performance again showed no significant differences for sex ($F(1,19) = .141$, $p = .711$) or root note ($F(1,19) = .156$, $p = .697$). All chickadees met the DR criterion for both phases before reaching 75 blocks of trials.

Figures 3.2 shows Vincentized acquisition curves of each S- stimulus across both phases for chickadees in the P5 S+ (top panel) and TT S+ (lower panel) conditions. As in Experiment 2a, chickadees learned to discriminate all S- stimuli from the S+ significantly above chance by the end of each phase regardless of reinforcement condition. Figure 3.2 also shows that the rate at which the different S- stimuli were learned across Vincentized blocks varied, but this variation appears negligible. The pattern of results across both reinforcement conditions suggest that all S- stimuli were discriminated at a similar level of expertise by the end of each phase with the exception of phase 1 for the P5 S+ group, which showed a lower overall DR for TT. These overall trends are supported by statistical analyses for each reinforcement group presented below.

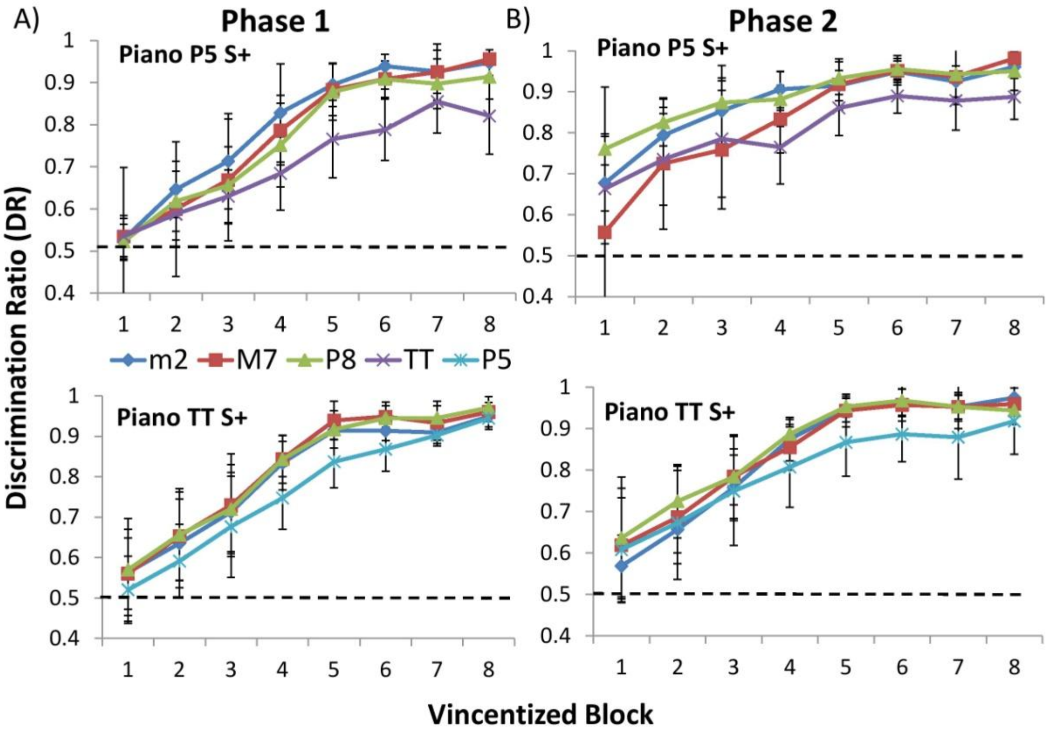


Figure 3.2 Average Discrimination Ratio (DR) for each S- stimulus across the eight Vincentized blocks for chickadees in the piano condition (see text for

details). Panel (A) shows phase 1 and phase 2 chickadee discrimination for the P5 S+ group and panel (B) shows phase 1 and phase 2 discrimination for the TT S+ group. Error bars show 95% confidence intervals.

P5 S+ Group

Acquisition. All birds in the P5 S+ condition showed a significant main effect of block such that DRs increased as Vincentized block increased (*F* (7, 105) = 61.152, $p < .001$, partial $\eta^2 = .924$). A significant interaction of S- stimulus × Vincentized block (*F* (21, 105) = 3.294, $p < .001$, partial $\eta^2 = .397$) was also observed, suggesting variation in performance among the S- stimuli during acquisition. The main effect of S- stimulus was also significant (*F* (3, 105) = 4.763, $p = .016$, partial $\eta^2 = .488$) which indicates that performance differences among the S- stimuli differed overall. However, pairwise comparisons revealed no significant differences in discrimination performance of TT ($M = .76$) compared with m2 ($M = .85$, $p = .321$), M7 ($M = .82$, $p = .337$), and P8 ($M = .84$, $p = .541$). Panel (A) of Figure 3.2 shows that performance among the S- stimuli seemed fairly consistent across acquisition, but TT was lowest overall.

Savings. The intercepts of Vincentized data among S- stimuli did not vary significantly (*F* (3, 15) = 2.571, $p = .093$, partial $\eta^2 = .340$), but there was a significant effect of phase (*F* (1, 15) = 6.752, $p < .05$, partial $\eta^2 = .575$) such that DRs were higher overall in the second phase compared to the first. As indicated by Figure 3.2, savings effects seem to be most substantial for P8, m2, TT, and M7 (ordered from strongest to weakest). Critically, all stimuli were discriminated above chance during the first block of acquisition for the second phase, which was not apparent during the first phase. P5 reinforcement consequently resulted in great variation for savings effects among S- stimuli.

TT S+ Group

Acquisition. Chickadees in the TT S+ group also showed increasing DRs with increasing Vincentized block (*F* (7, 105) = 46.720, $p < .001$, partial $\eta^2 = .903$). There was also a significant main effect of S- stimulus (*F* (3, 105) = 4.093, $p = .026$, partial $\eta^2 = .450$), indicating performance differences among S- stimuli. However, the interaction of S- stimulus × Vincentized block was not significant (*F* (21, 105) = 0.837, $p = .669$, partial $\eta^2 = .143$). Pairwise comparisons showed no statistical performance differences between P5 (*M* = .78) and m2 (*M* = .82, $p = 1.00$), M7 (*M* = .83, $p = .193$), and P8 (*M* = .84, $p = .529$) over time. This consistency in performance among the S- stimuli is illustrated in panel (B) of Figure 3.2.

Savings. The intercepts among S- stimuli did not vary significantly (*F* (1, 15) = 1.099, $p = .380$, partial $\eta^2 = .180$). The main effect of phase was also not significant (*F* (1, 15) = 3.479, $p = .121$, partial $\eta^2 = .410$). However, Figure 3.2 shows that overall performance was better during the first few blocks of phase 2 compared with the first few blocks of phase 1. It additionally demonstrates that, even at block 1 of phase 2, there is little variability among discrimination performance for S- stimuli. Thus, it appears that there were minimal savings effects from phase 1 to phase 2 and that these effects were equal for all stimuli.

3.5 Discussion

Experiments 2a and 2b tested chickadees using a Go/No-go operant discrimination procedure directly comparable to that employed for human participants in Experiments 1a and 1b. The same stimuli were used and two of the four reinforcement groups from the human experiments were employed in the chickadee experiments. All chickadees showed significant levels of discrimination performance, though there was considerable variation among chickadees, especially those discriminating synthetic intervals in Experiment 2a. Overall, chickadees discriminating piano intervals learned the discrimination much faster than chickadees exposed to synthetic intervals (see Figure 3.3). However, both groups of chickadees generally had the highest proportion of errors to S- stimuli that had an upper note close in pitch to the upper note of the S+ stimulus. This trend was more pronounced for chickadees that discriminated the synthetic intervals. Chickadees that discriminated the piano intervals seemed to show more consistent patterns of learning to all S- stimuli. Again, the specific comparisons among the individual stimuli between both groups suggest a more complex story.

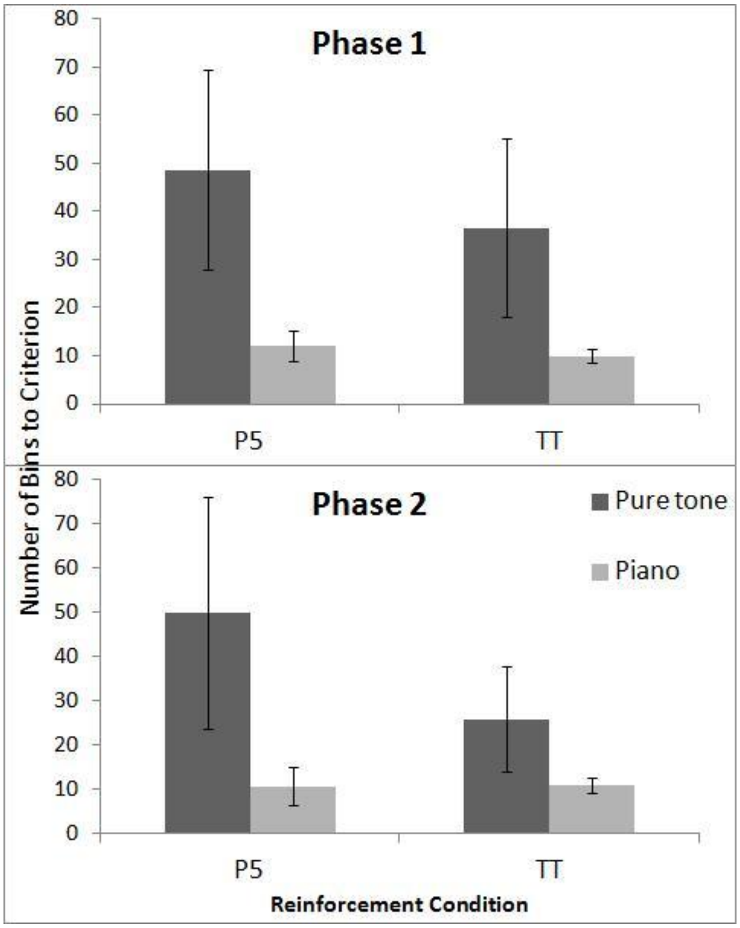


Figure 3.3 Number of bins to criterion for chickadees in both timbre groups (synthetic and piano) and both reinforcement groups (P5 and TT). Error bars are 95% CI.

Experiment 2a

On average, chickadees trained to discriminate synthetic intervals took more than twice as long to reach criterion than chickadees trained to discriminate piano intervals. Though discrimination performance was quite good for most intervals by the end of each phase, overall performance was hindered by the relatively poor discrimination of one interval. For example, chickadees in the P5 S+ condition and chickadees in the TT S+ condition could discriminate the TT S- and P5 S- to a level just above chance, respectively, by the end of each phase (see

Figure 3.1). This result is intriguing and somewhat contradictory to previous research investigating discrimination of non-biologically relevant sounds.

Songbird species (including chickadees) have been shown to discriminate triads with relative ease (e.g. Hoeschele et al., 2012; Watanabe et al., 2005). Moreover, songbirds have extremely acute processing abilities for discriminating tones that differ by less than a semitone (e.g., Weisman et al., 1998; Weisman et al., 2004). It was therefore not expected that chickadees would have so much difficulty discriminating non-reinforced intervals with an upper note closely related in pitch to the reinforced interval. However, because the stimuli used consisted of two pitches as opposed to a single pitch, a plausible explanation is that chickadees were comparing the relative pitch relations among the notes. Specifically, black-capped chickadees have been shown to rely on absolute pitch more heavily than relative pitch in discrimination tasks (Weisman & Ratcliffe, 1989), suggesting that their relative pitch processing abilities are poorer. Therefore, the decreased rate of learning for these intervals may partially be explained by the poor relative pitch processing abilities of chickadees. The absolute frequency differences between the upper notes of these difficult comparisons may explain the rest of this result.

Another interesting finding is that chickadees in the TT S+ group found the m2 interval fairly difficult to discriminate while chickadees in the P5 S+ group did not (see Figure 3.1). Though performance for m2 was considerably better than performance for P5 for chickadees in the TT S+ group, it was still significantly worse than other stimuli. The m2 and TT are the two most dissonant within-octave intervals as defined by their frequency ratios (16:15 and 45:32, respectively). Given this, it seems possible that dissonance led the chickadees to perceive these two intervals as more similar, thus hard to discriminate. However, m2 is unique within the entire stimulus set as it is the only interval composed of two tones within a critical bandwidth (see Plomp & Levelt, 1965). That is, the m2's sensory dissonance should have elicited an easier discrimination relative to the harmonically dissonant TT. Given that the intervals were composed of pure tones, harmonic dissonance should not have been a factor in determining ease of discrimination. The fact that this was not the case for the TT S+ group suggests that dissonant intervals are difficult to discriminate only when the rewarded stimulus is also dissonant.

Finally, chickadees were able to successfully show savings effects between Phase 1 and Phase 2, suggesting that the relative pitch relations that comprise these intervals served as a significant discriminative cue. Naturally, the savings effects seemed to be more pronounced for stimuli that were discriminated to a higher level of expertise. For example, the first Vincentized block of Phase 2 indicates that savings seemed somewhat insignificant for the TT in the P5 S+ group and the P5 in the TT S+ group. Though performance for these intervals in their respective groups was worse during Phase 1 overall. But because these particular stimuli did not have harmonic frequencies, it shows that chickadees can generalize the pitch relations among isolated fundamental frequencies, showing further support of their ability to use relative pitch to solve these types of tasks (Hoeschele et al., 2012).

Experiment 2b

Similarly to chickadees exposed to synthetic intervals, chickadees exposed to piano intervals were also able to discriminate the intervals with significantly high levels of performance. As noted before, they were able to reach criterion in less than half as many trials as chickadees exposed to synthetic intervals. This makes sense because the presence of higher harmonics provides relatively more information in a sound with which to discriminate. Songbirds have been shown to use the presence of specific harmonics as a discriminative cue (Cynx, Williams, & Nottebohm, 1990). Other work has also shown that black-capped chickadees, in general, perform worse when discriminating synthetic triads compared with triads that possess harmonic frequencies and that discrimination transfer between these types of stimuli is relatively poor (Hoeschele et al., unpublished data). Therefore, the finding that increased discrimination performance for piano tones relative to synthetic tones was not surprising.

Discrimination performance also varied significantly among the stimuli but to a much less extent than the chickadees exposed to synthetic intervals. However, the same general trend emerged: one particular interval proved more difficult than the rest. These intervals, again, were the TT for the P5 S+ group and the P5 for the TT S+ group. This basic finding suggests, again, that the proximal fundamental frequencies of the upper notes of these comparison intervals made the discrimination more difficult. That is, the pitch of the upper notes of the TT and P5 for P5 S+ and TT S+ groups served as a more salient cue. This recurring result may occur because chickadees use the absolute pitch values of fundamental frequencies in a harmonic vocalization quite heavily to determine incoming signals from conspecifics (Hoeschele, Guillette, & Sturdy, 2012). In the case of these discriminations, it seems as though chickadees primarily use absolute pitch to solve the task but effects of savings from Phase 1 to Phase 2 indicate that relative pitch is used as well.

Chapter 4

General Discussion

In a general sense, these sets of studies are important in that they contribute valuable information into the perception of musical sounds with and without harmonics in humans and non-human animals. Much recent research has aimed to disentangle the origins of music as it is now generally understood that music may be closely related to the development of language and may have well served significant communicative function in its evolutionary history. Examining consonance/dissonance, timbre, and the pitch relations among musical intervals, only a few aspects of music perception among many, provides only one path to painting a coherent picture of the evolutionary origins of music. The results from these experiments understood in tandem with many others, however, allow us to form concrete generalizations in regard to the relative importance of music.

Our most pertinent finding is that humans and chickadees categorize musical intervals differently, though both species are able to discriminate these intervals and generalize their discriminations to the same intervals on a different root note. The level of difference depends, at least in part, as to whether the discrimination taking place involves musical intervals with harmonic frequencies or musical intervals without harmonics. For humans, timbre modulates whether the absolute pitch of the upper note or the relative consonance/dissonance of a comparison interval is a more salient feature for categorization: synthetic intervals elicits attenuation to pitch while piano intervals with harmonics elicits categorization by means of consonance/dissonance. Chickadees seem to rely on the absolute pitch of the upper note of comparison intervals regardless of timbre. However, chickadees perform significantly better when discriminating piano intervals compared with synthetic intervals. Examining these differences in performance in relation to what is known about these properties of sound provides a more thorough explanation.

4.1 Comparative Interval Discrimination

The experiments performed with the chickadees aim to replicate directly the procedures utilized in the human experiments. Aside from apparatus differences, there were a few critical differences that were unavoidable in replicating the experiments between species. One difference was that the two species were rewarded and punished during the task differently. Humans received positive reinforcement for making a correct “Go” response with flashing text on a computer screen that said “correct” while chickadees received access to food. For making an incorrect “Go” response, humans received a time-out period of five seconds in which a new trial could not be initiated and chickadees received a time-out period of 30 seconds with the house lights off. However different, these procedures were part of our standard operating protocol and have been used in previous comparative experiments. Another key difference between the human and chickadee experiments was that all humans, regardless of reinforcement or timbre condition, received the same number of trials (96). Chickadees varied in the number of trials they took to complete the experiment because there was a set criterion in which chickadees could advance to the second phase and finish. The Vincentized analysis allowed a comparison between chickadees in all conditions because it standardized the number of blocks regardless of the actual number of trials each chickadee took to reach criterion.

In order to make the results from the chickadees and humans in our studies comparable, we assigned a rank to each stimulus for each individual. The rank for each stimulus was obtained by ranking the total number of responses for each S- stimulus relative to the other S- stimuli. Values closer to 1 indicate less responding and values closer to 4 indicate more responding. The mean of these ranks across both phases of the experiments were then compared between humans and chickadees in the same timbre and reinforcement conditions (see Figure 4.1). Figure 4.1 shows that humans and chickadees were relatively similar in responding to the major seventh and octave across all conditions except the piano P5 S+: humans, on average, found the octave the most difficult interval to discriminate whereas chickadees found the tri-tone the most difficult. The major seventh was relatively easy for humans to discriminate but was more difficult for chickadees.

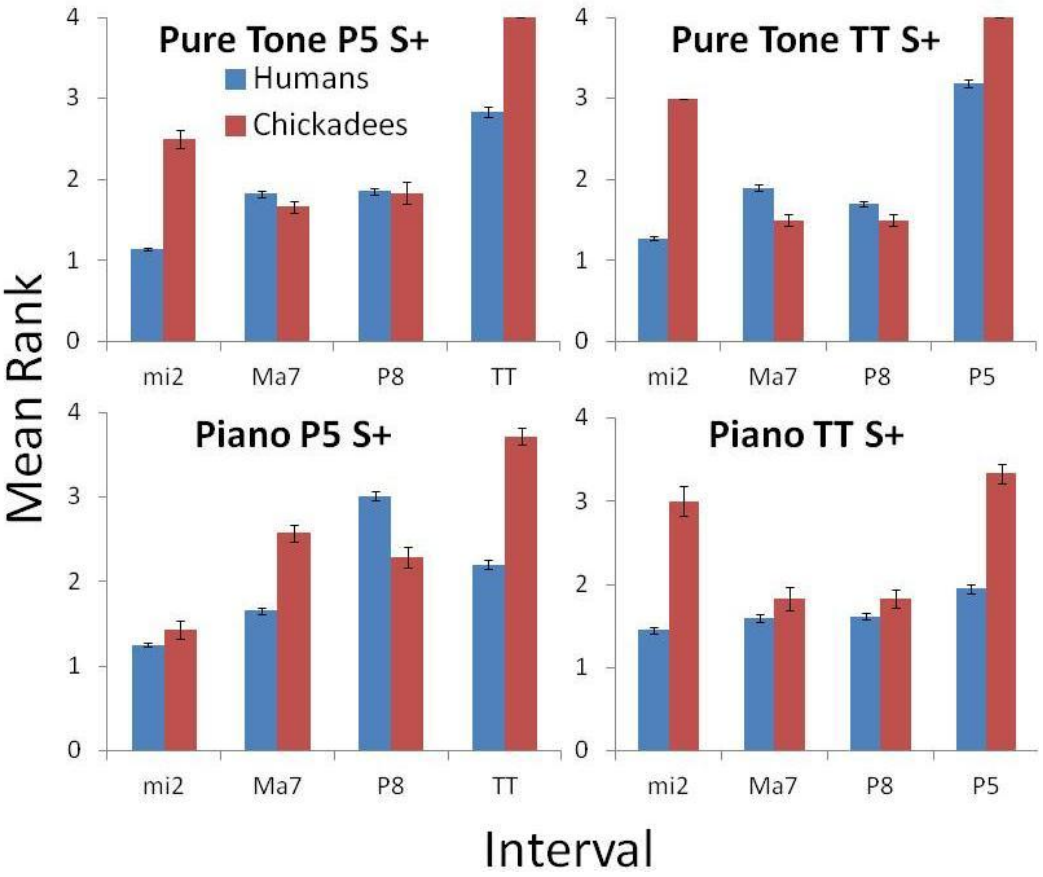


Figure 4.1 Mean rank of responses to each S- stimulus across timbre conditions and reinforcement conditions for humans and chickadees. Error bars are 95% CI.

Another key difference illuminated by Figure 4.1 is that chickadees consistently made the most errors to the tri-tone when the perfect fifth was the S+ and to the perfect fifth when the tri-tone was the S+. This trend was prevalent in both timbre conditions for chickadees but only in the pure tone condition for humans. Humans in the piano TT S+ condition made a marginally greater number of most errors to the perfect fifth compared with the minor second, major seventh, and octave. Finally, chickadees and humans substantially differed in the number of errors they made to the minor second. Humans found this interval the easiest to discriminate in every condition while chickadees found it the second most difficult interval in both pure tone reinforcement conditions and the piano TT S+ condition. Interestingly, the mean rank of errors for the minor second for chickadees in the piano TT S+ condition nearly matched that for the perfect fifth.

4.2 Pitch Perception

Pitch is an extremely salient feature of sound and is used as a primary cue for communication purposes in many hearing species (see Doupe & Kuhl, 1999). The fact that black-capped chickadees rely on the pitch (F_0) of conspecific vocalizations to determine genetic quality in a mate (Christie, Mennill, & Ratcliffe, 2004; Hoeschele, Guillette, & Sturdy, 2012) and identify individuals (Guillette et al., 2010) makes this feature important for their survival. Moreover, humans most readily use pitch to determine meaning and structure in language; some languages even have different pitches assigned for the same syllables to differentiate certain words (Deutsch et al., 2006). In the current experiments, the pitch of the upper note of the musical intervals varied in order to manipulate consonance and dissonance. The magnitude of the difference in pitch between the upper notes of reinforced intervals and non-reinforced intervals therefore varied as well. For example, humans and chickadees that were reinforced for the perfect fifth (P5) had to inhibit responses to the minor second (m2), major seventh (M7), octave (P8), and tri-tone (TT), the latter of which contained an upper note that differed by only one semi-tone from P5 while the other three differed by four or more semitones (see Figure 2.1). Our general finding, for both humans and chickadees exposed to synthetic intervals, was that non-reinforced intervals with an upper note close in pitch to the reinforced interval elicited higher error rates.

These results support our expectation in that the pitch of the upper note is more salient than the relative consonance/dissonance of these intervals.

However, humans seemed to reach relatively high DRs for these intervals by the end of both phases while chickadees were still performing slightly above chance. In a comparative sense, this facilitates previous findings suggesting that humans are better at relative pitch discriminations than chickadees (e.g., Hoeschele et al., 2012). Though Hoeschele et al. used a very similar task, the experiment was designed such that the first phase consisted of triads only on a C root note and the second phase was the same triads composed on a D root note. They posited that because of this design, the absolute pitch of the notes contributed to the chickadees' differential responses. Counterbalancing the root note on which the intervals were composed across phases in the current study suggests that it is the relative pitch that determines these errors among interval comparisons and not the absolute pitch. This is facilitated by the general finding that chickadees can use relative pitch to solve discrimination tasks but that they use absolute pitch with greater accuracy (Weisman & Ratcliffe, 1989). However, the proportion of errors both humans and chickadees made to the non-reinforced interval with an upper note close in pitch within each phase suggests that absolute pitch comparisons were being made as well. Small changes in frequency, such as that between the upper notes of our tri-tone (370 Hz) and perfect fifth (392 Hz) on C, are easy to detect when presented as isolated tones in both humans and songbirds, though some species of songbirds are more accurate (Weisman et al., 2004). Thus, we suggest that the relative proportion of errors to these intervals is likely a function of absolute pitch and relative pitch comparisons.

One issue in this explanation, however, is the proportion of errors that chickadees made to the synthetic minor seconds. If chickadees were solely relying on the differences in absolute pitch of the upper notes to solve the task, they should not have made as many errors to the minor second when discriminating it from the perfect fifth or tri-tone in their respective S+ groups. The upper note of the minor second is just one semitone above the root note, but the absolute frequency difference between that note and the upper note of the reinforced interval (perfect fifth or tri-tone) is roughly equivalent to the octave or major seventh, which chickadees had no problem discriminating. This finding is somewhat surprising when compared with humans, who found the minor second to be the easiest interval to discriminate. The exact reason for this increased error rate is not known, but a plausible explanation would be that chickadees also attended to the absolute pitch of the root note when making comparisons, therefore making an interval with an upper note close in pitch to the root note more difficult to discriminate. Another songbird species (Java sparrows) exposed to triads could not show generalization to inverted chords (i.e., when the root note is moved up one octave; Watanabe et al., 2005), suggesting that discriminations among musical stimuli with common root notes may be more difficult than discriminations among stimuli with different root notes. Another possibility might be that the minor second interval, compared with all other intervals used in the study, most closely resembles the interval used in black-capped chickadee's courtship song (Christie, Mennill, & Ratcliffe, 2004). Thus, the biological significance of the absolute pitch values of the minor second may have led the chickadees to categorize the interval differently and learn the discrimination more slowly.

4.3 Consonance/Dissonance Perception

According to Helmholtz's (1954) theory and more conventional theories of consonance and dissonance, a musical interval with harmonic frequencies will increase a dissonant sound's roughness as opposed to a musical interval without these harmonics. The simple frequency ratios by which consonant intervals are related cause their harmonic overtones to overlap exactly, producing an absence of roughness. Therefore, we expected that response patterns for synthetic intervals and piano intervals would be different for both species such that synthetic intervals related by consonance or dissonance would result in a fewer number of errors when compared with intervals with harmonic frequencies. This was true for both species, however, there were also differences. Humans reinforced for consonant piano intervals (perfect fifth and octave) showed a higher error rate to non-reinforced consonant piano intervals, while humans reinforced for dissonant piano intervals (tri-tone and major seventh) discriminated all non-reinforced piano intervals relatively equally. Chickadees in the piano condition, regardless of reinforcement, discriminated all intervals almost equally, though a trend similar to chickadees in the synthetic condition emerged: the non-reinforced interval with an upper note close in pitch to the reinforced interval tended to create the highest proportion of errors.

Comparatively, these results suggest that humans use the relative consonance/dissonance of piano intervals as a cue for discrimination purposes and chickadees do not. Though consonance/dissonance has been shown to be a relatively salient feature of musical sounds for humans, this experiment provides evidence to suggest that its relative salience can be manipulated with the absence or presence of timbre. Thus, consonance/dissonance is a more salient discriminatory cue than pitch when harmonics are present as opposed to when they are absent. Chickadees, however, still tend to rely on the relative pitch of the upper note of the intervals to learn to discriminate whether harmonics were present or not, though this effect was mitigated when harmonics were present.

Chickadees have been shown to use consistent, consonant frequency ratios in some of their own vocalizations (Weisman & Ratcliffe, 1989; Weisman et al., 1990). Moreover, our observation that chickadees were able to reach discrimination criterion in many fewer trials when exposed to piano intervals, compared with synthetic intervals, is in line with other work showing that timbre is an important cue for communication among songbird species, including black-capped chickadees (Cynx, Williams, & Nottebohm, 1990; Hoeschele et al., 2012). However, it is somewhat unsurprising that chickadees did not use the relative consonance/dissonance of the piano intervals as the primary discrimination cue because frequency ratios correlated with consonance and dissonance may be simply a human contrivance. Comparing these response patterns for musical intervals with and without harmonics suggests that experiments aiming to disentangle consonance/dissonance perception in non-human species need to address the nature of the stimulus more carefully.

Humans and chickadees also differed in their respective abilities to learn to discriminate the dissonant minor second interval and the consonant octave interval. Humans found the minor second the easiest to discriminate across all reinforcement and timbre conditions whereas chickadees found it more difficult than the octave, perfect fifth, and tri-tone. Given that the minor second is the only interval in the stimulus set that is composed of two notes within a critical bandwidth (see Plomp & Levelt, 1965), it is possible is that humans more readily identify the minor second by its sensory dissonance and are thus able to identify the stimulus more easily. However, this still leaves the question why chickadees found the minor second particularly difficult. Because chickadees tended to focus on the pitch of the upper notes in general, future studies may address this deviance in discrimination for the minor second by incorporating other intervals that have an upper note that is relatively close in pitch, such as a major second or a minor third.

The octave, an interval of special interest in its own right, was most difficult to discriminate only for humans exposed to piano intervals in the perfect fifth S+ condition. Chickadees in the perfect fifth S+ condition actually found the octave the easiest to discriminate. The octave is the only interval in the stimulus set in which the upper note of the interval shares the same tone chroma as the root note. Given this, and the octave’s nearly universal usage in music of all cultures, the octave should have been easily recognized and therefore easily discriminated by humans. Our result that shows a decreased accuracy for octave discrimination when exposed to piano intervals compared with synthetic intervals suggests that it is indeed the presence of harmonics in the piano intervals that causes this increased error rate. Tone chroma may be especially important for humans compared with chickadees as transposition across octaves is extremely prevalent in human music (Burns, 1999) and speech (Titze, 2000) and is sometimes dependent on the natural frequency range of one's voice. However, because this result was only present in the perfect fifth S+ condition, it is presumed that the relative consonance of the perfect fifth is at play. Other research suggests that the concept of "octaveness" is developed through experience and is not of perceptual origin (Sergeant, 1983). This would therefore explain the difference between humans and chickadees.

4.4 Experiential Factors

A huge component to the differences in music perception between human and non-human species may be attributed to differences in experience. The rules that govern aspects of Western music are devised by humans and this often leads to the quick and misleading assumption that music itself is a human construct. However, this undermines the possibility that rules for music may have developed through some evolutionary processes shared with other species. For example, the processing of pitch combinations is essential to the experience of music in humans but is also essential for songbird vocal communication. The physical properties of sound may have also helped shape the way Western music is made and therefore perceived. This is exemplified with findings such that unequal (i.e.

logarithmic) step sizes of frequency, representative of pitch, are accurately encoded by the auditory system (Trehub, Schellenberg, & Kamenetsky, 1999). Our finding that the relative consonance/dissonance of comparison intervals is a more salient feature of sound than pitch when harmonics are present was present only in humans and, as such, may be a product of experiential factors.

All of the human subjects in our experiments had experience with the intervals they had to discriminate through enculturation to Western music whereas chickadees did not. Chickadees do not produce or experience Western music and the birds used in these experiments were naïve to these particular intervals. This difference alone may contribute to some of the variance between the two species. However, our finding that the chickadees possessed some sensitivity to a change in timbre suggests that experience with synthetic stimuli in general may have elicited these differences. From an ecological standpoint, it makes sense that chickadees took a much greater number of trials to reach criterion for synthetic intervals simply because pure tones do not occur in nature. Humans may also have little experience with pure tones but presumably more than chickadees. One way to address these differences in future research is to use human subjects with little to no experience with Western music (i.e. infants).

Experience also varies within each subject and thus may affect accuracy for the task as well. In humans, we did not find any differences in overall performance between musicians and non-musicians. We also did not find differences in the pattern of responding between musicians and non-musicians across all reinforcement conditions. However, this may be due to a relatively small number of subjects in our experiments with a high degree of musical training. Future studies should undoubtedly aim to have equal numbers of subjects with and without training. In sum, a plausible explanation for the "naturalness" of simple ratios in musical intervals is that it is largely due to experience and also tuned by local culture (e.g., Dowling & Harwood, 1986).

4.5 Uniqueness/Innateness

Perhaps most importantly, our results suggest that consonance/dissonance perception is different for human and non-human species. Though previous research indicates that other songbird species can learn differences between consonant and dissonant stimuli (e.g., Watanabe, Uozumi, & Tanaka, 2005) and show relatively similar patterns of response as humans (e.g., Hoeschele et al., 2012), these studies addressed the relative salience of consonance and dissonance and its variance with timbre. The fact that chickadees exposed to synthetic or piano intervals did not show drastic differences in errors for particular stimuli suggests that consonance/dissonance may not be a property of sound that is perceived innately. But because chickadees *can* discriminate sounds based on their relative consonance/dissonance, it rules out the idea that consonance/dissonance in a Western musical sense is a completely unique property. Thus, the roughness that occurs with dissonant sounds and the absence of roughness in consonant sounds may be perceived by non-human species, but their perception may be different. The differences in overall performance between chickadees exposed to synthetic intervals and piano intervals further suggests that timbre is not a perceived quality unique to humans. This finding has implications for future studies to address multiple aspects of sound when trying to understand how non-human species perceive musical sounds.

There is some reason to expect that chickadees might be able to represent musical stimuli the way humans do because pitch intervals play a role in song recognition (e.g., Hurly et al., 1992). Although the possibility remains that song recognition relies on modular acoustic components that might not apply to biologically irrelevant stimuli such as musical intervals. The results of the chickadee experiments show support for the memorization of absolute frequency values of specific notes, a finding that has been shown in other species of birds using notes presented sequentially rather than simultaneously (e.g., Hulse, Cynx, & Humpal, 1984). Absolute pitch perception is of huge importance to chickadee communication and of markedly less importance for human communication. This difference alone may explain the differences present in many comparative studies aiming to understand the perception of certain qualities of music across species. Further training with more stimulus sets within a phase could help parse out the variance related to pitch memorization versus categorization of intervals based on their harmonic structure.

Speculation into the establishment of the fixed ratios associated with consonance and dissonance in music reveals that humans and non-human species may have a similar preference for some intervals as opposed to others based on similar constraints on auditory perception and memory (Hauser & McDermott, 2003). For example, some of the more “natural” intervals (e.g. perfect fifth, major second) may be found in animal vocalizations for this reason. Although the design of our experiments facilitated only two different types of timbre across humans and chickadees, further investigations with other types of timbre may indeed show differences from the results presented here. Some chickadee vocalizations possess pitch ratios that are very similar to some intervals defined by Western music (Horn et al., 1992; Weisman et al., 1990). As such, using intervals with a timbre closer to that of a chickadee's vocalization may facilitate increased performance for these natural intervals. Using a timbre that is more unfamiliar to humans may also eliminate biases due to experience.

In summary, the ability to discriminate harmonic structures associated with Western music varies in humans and chickadees with spectral distribution. Simple ratios have important roles in music and are common in the music of numerous cultures in adult and infant listeners (e.g. Koon, 1979; Lentz, 1965; Crickmore, 2003; Schellenberg & Trehub, 1996). There is still much to uncover in relation to the origin of these ratios and whether or not they are unique to Western music or common with non-human animals. At the very least, our results present a framework for probing interspecies differences through careful consideration of stimulus properties and the behavior of humans and chickadees. Humans across all cultures make and enjoy music with certain common features, including simple frequency ratios. These ratios associated with consonant and dissonant sounds are also found in vocalizations of black-capped chickadees and are related to perceptual sensitivities across more than two. Studying other vocal learning species, including primates, and nonvocal learners using similar paradigms could illuminate a more exact evolutionary basis for consonance and dissonance perception. Because of the similarities and differences in responses between both synthetic and piano timbre groups in both species, our results suggest that timbre is an ecologically relevant component of sound for both species and that the presence or absence of harmonic frequencies influences each species’ perception of musical sounds differently.

References

Bachem, A. (1950). Tone height and tone chroma as two different pitch qualities. *Acta Psychologia 7*, 80-88.

Bachem, A. (1955). Absolute pitch. *Journal of the Acoustical Society of America, 27*, 1180–1185.

Braaten, R.F. & Hulse, S.H. (1991). A songbird, the European starling (*Sturnus vulgaris*), shows perceptual constancy for acoustic spectral structure. *Journal of Comparative Psychology, 105,* 222-231.

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

Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music 2nd ed.,* 215–264. San Diego, CA: Academic Press.

Butler, J.W. & Daston, P.G. (1968). Musical consonance as musical preference: A cross-cultural study. *Journal of General Psychology, 79,* 129-142.

Cazden, N. (1959). Musical intervals and simple number ratios. *Journal of Research in Music Education, 7,* 197-220.

Cazden, N. (1980). The definition of consonance and dissonance. *International Review of the Aesthetics and Sociology of Music, 11,* 123-168.

Chiandetti, C. & Vallortigara, G. (2011). Chicks like consonant music. *Psychological Science, 22*, 1270-1273.

Christie, P. J., Mennill, D. J., & Ratcliffe, L. M. (2004). Pitch shifts and song structure indicate male quality in the dawn chorus of black-capped chickadees. *Behavioural Ecology and Sociobiology, 55,* 341–348.

Crickmore, L. (2003). A re-valuation of the ancient science of harmonics. *Psychology of Music, 31,* 391-403.

Cynx, J., Hulse, S. H., & Polyzois, S. (1986). A psychophysical measure of pitch discrimination loss resulting from a frequency range constraint in European starlings (*Sturnus vulgaris*). *Journal of Experimental Psychology,12,* 394-402.

Cynx, J., Williams, H., Nottebohm, F. (1990). Timbre discrimination in Zebra finch (*Taeniopygia guttata*) song syllables. *Journal of Comparative Psychology, 104,* 303-308.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

Deutsch, D., Henthorn, T., Marvin, E., & Xu H-S (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *Journal of the Acoustical Society of America*, 119, 719-722.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience, 22*, 567-631.

Dowling, W. J. & Harwood, D. L. (1986). *Music cognition*. Orlando: Academic Press.

Durand, S.E., Tepper, J.M., & Cheng, M-F (1992). The shell region of the nucleus ovoidalis: A subdivision of the avian auditory thalamus. *Journal of Comparative Neurology, 323,* 495-518.

Fitch, W.T. (2006). The biology and evolution of music: A comparative perspective. *Cognition, 100,* 173-215.

Galileo Galilei (1638/1954). Dialogues Concerning Two New Sciences. Translated by H. Crew and A. de Salvio, 1914. New York: Dover Publications.

Garner, W. R. (1970). The stimulus in information processing. *American Psychologist, 25,* 350-358.

Gibson, J.J. (1966). The Senses Considered as Perceptual Systems. Boston: Houghton Mifflin.

Goldstein, E.B. (1999). Sensation and Perception. Brooks/Cole Publishing, Stamford, Connecticut, USA.

Greenwood, D.D. (1961). Auditory masking and the critical band. *Journal of the Acoustical Society of America, 33,* 484-501.

Grey, J.M. (1976). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America, 61,* 1270-1277.

Grey, J.M. (1978). Timbre discrimination in musical patterns. *Journal of the Acoustical Society of America, 64,* 467-472.

Griffiths, R., Double, M. C., Orr, K., & Dawson, R. J. G. (1998). A DNA test to sex most birds. *Molecular Ecology, 7,* 1071–1075.

Guillette, L.M., Bloomfield, L.L., Batty, E.R., Dawson, M.R.W., & Sturdy, C.B. (2010). Black-capped (*Poecile atricapillus*) and mountain chickadee

(*Poecile gambeli*) contact call contains species, sex, and individual identity features. *Journal of the Acoustical Society of America, 127,* 1116-1123.

Guillette, L. M., Farrell, T. M., Hoeschele, M., Nickerson, C.M., Dawson, M. R. W., & Sturdy, C., B., (2010). Mechanisms of call note-type perception in black-capped chickadees (*Poecile atricapillus*): Peak shift in a note-type continuum. *Journal of Comparative Psychology, 124,* 109-115.

Hauser, M.D. & McDermott, J. (2003). The evolution of the music faculty: A comparative perspective. *Nature Neuroscience, 6,* 663-668.

Helmholtz, H. (1954). On the sensations of tone as a physiological basis for the theory of music. New York : Dover (Original work published 1885).

Hoeschele, M., Cook, R.G., Guillette, L.M., Brooks, D.I. & Sturdy, C.B. (2012). Black-capped chickadee (*Poecile atricapillus*) and human (*Homo sapiens*) chord discrimination. *Journal of Comparative Psychology, 126,*57-67.

Hoeschele, M., Cook, R.G., Guillette, L.M., Hahn, A.H., & Sturdy, C.B. (2012). Auditory same/different concept learning and generalization in black-capped chickadees (*Poecile atricapillus*). *PLoS ONE, 7: e47691. doi:10.1371/journal.pone.0047691.*

Hoeschele, M., Guillette, L.M., & Sturdy, C.B. (2012). Biological relevance of acoustic signal affects discrimination performance in a songbird. *Animal Cognition, 15,* 677-688.

Hoeschele, M., Weisman, R.G., & Sturdy, C.B. (2012). Pitch chroma discrimination, generalization and transfer tests of octave equivalence in humans. *Attention, Perception, & Psychophysics, 74,* 1742-1760.

Horn, A.G., Leonard, M.L., Ratcliffe, L., Shackleton, S.A., & Weisman, R.G. (1992). Frequency variation in songs of black-capped chickadees (*Parus atricapillus*). *The Auk, 109,* 847–852.

Hughes, M., Nowicki, S., & Lohr, B. (1998). Call learning in black-capped chickadees (*Parus atricapillus*): The role of experience in the development of ‘chick-a-dee’ calls. *Ethology, 104,* 232–249.

Hulse, S. H., Bernard, D. J., & Braaten, R. F. (1995). Auditory discrimination of chord-based spectral structures by European starlings (*Sturnis vulgaris*). *Journal of Experimental Psychology: General, 124,* 409–423.

Hulse, S. H., Cynx, J., & Humpal, J. (1984). Absolute and relative pitch discrimination in serial pitch perception by birds. *Journal of Experimental Psychology: General, 113,* 38-54.

Hurly, T. A., Ratcliffe, L., Weary, D. M., & Weisman, R., (1992). White-throated sparrows (*Zonotrichia albicollis*) can perceive pitch change in conspecific song by using the frequency ratio independent of the frequency difference. *Journal of Comparative Psychology, 106,* 388-391.

Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. In *Behavioral Neurobiology of Birdsong* (Vol. 1016, pp. 749-777). New York: New York Academy of Sciences.

Lejeune-Dirichlet, P. (1829) . Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données. (In French), transl. On the convergence of trigonometric series which serve to represent an arbitrary function between two given limits. *Journal f ̈ur die reine und angewandte Mathematik, 4*, 157–169.

Kameoka A. & Kuriyagawa, M. (1969a). Consonance theory part 1: Consonance of dyads. *Journal of the Acoustical Society of America, 45,* 1451-1459.

Kameoka, A. & Kuriyagawa, M. (1969b). Consonance theory part 2: Consonance of complex tones and its calculation method. *Journal of the Acoustical Society of America, 45,* 1460-1469.

Kling, J. W., & Riggs, L. A. (Eds.) (1971). *Woodworth and Schlosberg's experimental psychology,* New York: Holt, Rinehart & Winston.

Koon, N.K. (1979). The five pentatonic modes in Chinese folk music. *Chinese Music, 2,* 10-13.

Krohn, K. I., Brattico, E., Valimaki, V., & Tervaniemi, M. (2007). Neural representations of the hierarchical scale pitch structure. *Music Perception, 24*, 281–296.

Krumhansl, C.L. (1991). Music Psychology: Tonal structures in perception and memory. *Annual Review of Psychology, 42,* 277-303.

Krumhansl, C.L. & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 739-751.

Lentz, D.A. (1965). The gamelan music of Java and Bali. Lincoln: University of Nebraska Press.

Levitin, D. J. (2010). Why music moves us. *Nature, 464,* 834–835.

Licklider, J.C.R. (1951). A duplex theory of pitch perception. *Cellular and Molecular Life Sciences, 8,* 128-134.

Luce, D.A. (1963). Physical correlates of nonpercussive musical instrument tones. Ph.D. dissertation, Cambridge: Massachusetts Institute of Technology.

Lundin, R.W. (1947). Toward a cultural theory of consonance. *Journal of Psychology: Interdisciplinary and Applied, 23,* 45–49.

Marler, P. (1970). A comparative approach to vocal learning: Song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology, 71,* 1-25.

Marr, D. (1982) Vision: A Computational Investigation into The Human Representation and Processing of Visual Information. N.Y.: W.H. Freeman.

Melara, R.D., Marks, L.E. (1990). Interaction among auditory pitch dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics, 48,* 169-178.

McAdams S., Cunibile, J.C. (1992). Perception of timbre analogies. *Philosophical Transactions of the Royal Society, London, Series B, 336,* 383-389.

McDermott J. H., Hauser M. (2004) Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition, 94,* 11-21.

McDermott, J.H. & Hauser, M. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception, 23,* 29-59.

McDermott, J.H., Lehr, A.J., & Oxenham, A.J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, *20*, 1035-1041.

McDermott, J.H. & Oxenham, A.J. (2008). Music perception, pitch, and the auditory system. *Current Opinion in Neurobiology, 601,* 1-12.

Mooney, R. (2009). Neural mechanisms for learned birdsong. *Learning & Memory*, *16*, 655–669.

Njegovan, M., Hilhorst, B., Ferguson, S., & Weisman, R. (1994). A motor driven feeder for operant training in songbirds. *Behavior Research Methods Instruments & Computers*, *26,* 26–27.

Nottebohm, F. (1972). The origins of vocal learning. *The American Naturalist, 106*, 116-140.

Palya, W. L., & Walter, D. E. (2001). Document set for the high-performance experiment controller. Retrieved February 22, 2010 from http://www.jsu.edu/depart/psychology/sebac/Exp-Ctl.html.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*, 674-681.

Patel, A.D. (2008). Music, language, and the brain. Oxford University Press, Oxford, U.K.

Pinker, S. (1997). How the mind works (1st ed.). New York: Norton.

Pitt, M.A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology: Human Perception and Performance*, *20,* 976-986.

Plack C.J., Oxenham A.J., Fay R.R., Popper A.N. (2005). Pitch: neural coding and perception. Springer handbook of auditory research. New York: Springer Science.

Plomp, R. & Levelt, W. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America, 38,* 548-560.

Pyle, P. (1997). Identification Guide to North American Birds. Slate Creek Press, Bolinas, CA.

Russo, F.A. & Thompson, W.F. (2005). An interval size illusion: The influence of timbre on the perceived size of melodic intervals. *Perception & Psychophysics, 67,* 559-568.

Schellenberg, E.G., Bigand, E., Poulin-Charronnat, B., Garnier, C., & Stevens, C. (2005). Children's implicit knowledge of harmony in Western music. *Developmental Science, 8*, 551-566.

Schellenberg, E.G. & Trehub, S.E. (1996). Natural musical intervals: Evidence from infant listeners. *Psychological Science, 7,* 272-277.

Sergeant, D. (1983). The octave – Percept or concept. *Psychology of Music, 11,* 3-18.

Shepard, R.N. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review, 89,* 305-333.

Sidtis, J. (1980). On the nature of the cortical function underlying right hemisphere auditory perception. *Neuropsychology, 18,* 321-330.

Singh, P.G. (1987). Perceptual organization of complex-tone sequences: A tradeoff between pitch and timbre? *Journal of the Acoustical Society of America, 82*, 886-899.

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

Sturdy, C.B. & Weisman, R.G. (2006). Rationale and methodology for testing auditory cognition in songbirds. *Behavioral Processes, 72,* 265-272.

Sundberg, J. (1991). The Science of Musical Sounds, Academic Press, San Diego, CA.

Takeuchi, A.H. & Hulse, S.H. (1993). Absolute pitch. *Psychological Bulletin, 113,* 345-361.

Terhardt, E., Stoll, G., Seewann, M. (1982). Pitch of complex signals according to virtual-pitch theory: Test, examples, and predictions. *Journal of the Acoustical Society of America, 71,* 671-678.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., Schröger, E. (2005). Pitch discrimination accuracy in musicians vs. nonmusicians: an event-related potential and behavioral study. *Experimental Brain Research, 161,* 1-10.

Thorpe, W. (1954). The process of song-learning in the chaffinch as studied by means of the sound spectograph. *Nature, 173,* 465–469.

Tierney, A. T., Russo, F. A., Patel, A. D. (2011). The motor origins of human and avian song structure. *Proceedings of the National Academy of Sciences, 108*, 15510-15515.

Tirovolas, A.K. & Levitin, D.J. (2011). Music perception and cognition research from 1983 to 2010: A categorical and bibliometric analysis of empirical studies in Music Perception. *Music Perception, 29,* 23-26.

Titze, I.R. (2000). *Principles of voice production.* Iowa City, IA: National Center for Voice and Speech.

Trainor, L.J., Tsang, C.D., & Cheung, V.H.W. (2002). Preference for sensory consonance in 2- and 4-month old infants. *Music Perception, 20,* 187-194.

Trehub S.E., Schellenberg E., Kamenetsky S.B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance 25,* 965-975.

Watanabe, S., Uozumi, M., & Tanaka, N. (2005). Discrimination of consonance and dissonance in Java sparrows. *Behavioral Processes, 70,* 203-208.

Weisman, R. G., Balkwill, L.-L., Hoeschele, M., Moscicki, M. K., Bloomfield, L. L., & Sturdy, C. B. (2010). Absolute pitch in boreal chickadees and humans: Exceptions that test a phylogenetic rule. *Learning and Motivation, 41*, 156-172.

Weisman, R., Njegovan, M., & Ito, S. (1994). Frequency ratio discrimination by zebra finches (*Taeniopygia guttata*) and humans (*Homo sapiens*). *Journal of Comparative Psychology, 108,* 363-372.

Weisman, R. Njegovan, M., Sturdy, C., Phillmore, L., Coyle, J., & Mewhort, D. (1998). Frequency-range discriminations: Special and general abilities in zebra finches (*Taeniopygia guttata*) and humans (*Homo sapiens*). *Journal of Comparative Psychology, 112,* 244-258.

Weisman, R.G., Njegovan, M.G., Williams, M.T., Cohen, J.S., & Sturdy, C.B. (2004). A behavior analysis of absolute pitch: Sex, experience, and species. *Behavioral Processes, 66,* 289-307.

Weisman, R. & Ratcliffe, L. (1989). Absolute and relative pitch processing in black-capped chickadees, *Parus atricapillus*. *Animal Behavior, 38,* 685-692.

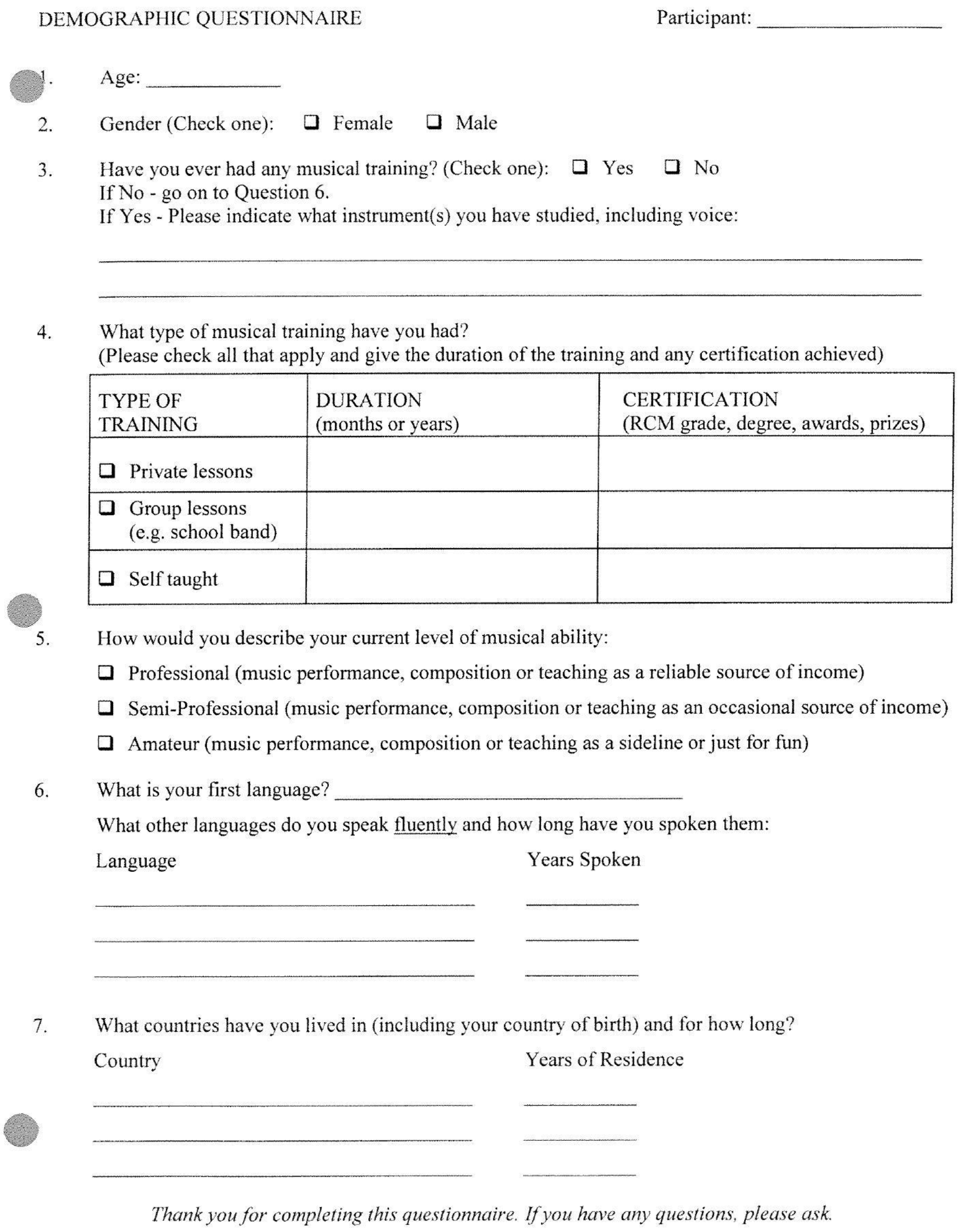
Weisman, R., Ratcliffe, L., Johnsrude, I., & Hurly, A.T. (1990). Absolute and relative pitch production in the song of the black-capped chickadee. *The Condor, 92,* 118-124.

Woolley, S. M. N. & Casseday, J. H. (2003). Response properties of single neurons in the zebra finch auditory midbrain: Response patterns, frequency coding, intensity coding, and spike latencies. *Journal of Neurophysiology, 91,* 136-151.

Zatorre, R.J. & Halpern, A.H. (1979). Identification, discrimination, and selective adaptation of simultaneous musical intervals. *Perception and Psychophysics, 26,* 384-395.

Zentner, M.R. & Kagan, J. (1996) Perception of music by infants. *Nature. 383,* 29.

Appendix A



DEMOGRAPHIC QUESTIONNAIRE Participant: __________________

1. Age: _____________

2. Gender (Check one): ☐ Female ☐ Male

3. Have you ever had any musical training? (Check one): ☐ Yes ☐ No
If No - go on to Question 6.
If Yes - Please indicate what instrument(s) you have studied, including voice:
____
____

4. What type of musical training have you had?
(Please check all that apply and give the duration of the training and any certification achieved)

TYPE OF TRAINING	DURATION (months or years)	CERTIFICATION (RCM grade, degree, awards, prizes)
☐ Private lessons		
☐ Group lessons (e.g. school band)		
☐ Self taught		

5. How would you describe your current level of musical ability:
☐ Professional (music performance, composition or teaching as a reliable source of income)
☐ Semi-Professional (music performance, composition or teaching as an occasional source of income)
☐ Amateur (music performance, composition or teaching as a sideline or just for fun)

6. What is your first language? ___________________________________
What other languages do you speak fluently and how long have you spoken them:

Language	Years Spoken
______________________________	__________
______________________________	__________
______________________________	__________

7. What countries have you lived in (including your country of birth) and for how long?

Country	Years of Residence
______________________________	__________
______________________________	__________
______________________________	__________

Thank you for completing this questionnaire. If you have any questions, please ask.

Background questionnaire used in human Experiments 1a and 1b.