University of Alberta

Pitch perception in vocal learners: Fundamental shared components of pitch processing and biological relevance

by

Marisa Hoeschele

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

Doctor of Philosophy

Psychology

Dedication

This thesis is dedicated to my grandfather, Wolfgang Arnold, who I lost earlier this year. His excitement about the accumulation of knowledge, his insight into working with others, and his predominant optimism and generosity continues to influence me as I pursue my goals.

Abstract

Pitch perception is important in both music and language perception in humans. Vocal learners, including humans, share the property of learning their vocalizations from a tutor when young. Similar behavioural and neural mechanisms underlying vocal production among different species of vocal learners makes it possible to study biological origins of language and music. Here the goal was to understand whether a songbird vocal learner, the black-capped chickadee, has similar pitch perception to humans. Humans generally rely more on relative pitch, the ability to detect the relationship among notes, when judging the pitch of stimuli (e.g., "minor 3rd"). Black-capped chickadees, however, generally rely more on absolute pitch, or the ability to detect a pitch without an external referent (e.g., "A" note), but are able to learn relative pitch rules. At the same time, this songbird has a relative pitch cue contained in their *fee-bee* song. In chapter 2, using a go/nogo operant paradigm, I assessed whether the relative pitch cue from the *fee-bee* song is an important biologically relevant stimulus that influences the perception of pitch in this species. I found that the pitch interval from the *fee-bee* song was discriminated more quickly and with a higher level of expertise than other pitch intervals. Then, in chapters 3-5, using a similar operant paradigm in humans, I compared first the relative and then the absolute pitch strategies of humans to chickadees. In chapter 3, I found that both humans and chickadees showed similar response patterns to the relative pitch of chords. Both species also transferred the discrimination to novel pitches. In chapters 4 and 5, I

compared the absolute pitch strategies of the two species by pitting pitch height perception (log-linear assessment of frequency) against octave perception (treating notes separated by a doubling in frequency, e.g., two "A" notes in Western music, as being similar). I found that, while human participants showed strong evidence of octave perception, black-capped chickadees seem to rely on pitch height perception alone. Despite both species having successful pitch perception strategies, the strategies used to accomplish this appear to be quite different. The implications of these results are discussed.

Table of Contents

List of Tables

List of Figures

lower complex tone of 250 cps and a variable higher one depends on the

frequency of this tone. Both complex tones consist of 6 harmonics. The

vertical lines represent interval width after an equally tempered scale.

Reprinted with permission from Plomp, R. & Levelt, W. J. M. (1965).

Tonal consonance and critical bandwidth. *Journal of the Acoustical*

Society of America, 38, 548-560. Copyright 1965, Acoustical Society of

Chapter 1 General introduction

Vocal Learning and Comparative Music Perception

Vocal learners are animals that learn their vocalizations through experience with vocalizing adult members of their species. These groups of vocal learners are relatively rare: to date only humans, oscine songbirds, hummingbirds, parrots, bats, cetaceans, pinnipeds, and elephants are thought to be vocal learners (Tyack, 2008; Fitch, 2005). Vocal learning species, especially songbirds, are often used as models for human language because, in terms of auditory perception and production, they have comparable vocal behavioural stages as well as neural and genetic architecture (Doupe & Kuhl, 1999; Haesler et al., 2007).

More recently, researchers have begun to argue that the human musical faculty should also be evaluated using cross-species comparisons (e.g. Fitch, 2005; Justus & Hustler, 2005; Hauser & McDermott, 2003); a field named by some as biomusicology (see Fitch, 2005). These researchers argue that cross-species comparisons can help to paint a clearer picture of the mechanisms that support music processing. In addition, because music, like language, is a human universal (i.e. a trait shared by all human populations regardless of their isolation from one another; Fitch, 2006), the much-debated evolutionary purpose of music may also be addressed by making comparisons among homologous and analogous species. Vocal learners make interesting analogies in biomusicology, both because there are parallels between human language and music, and because of additional musical analogies that can be made by comparing nonhuman vocal

learners to humans.

The human literature suggests that music and language are controlled by overlapping brain regions. Music and language are also arguably analogous, because they are both primarily auditory, but have non-auditory components — such as gestures or facial expressions in language and dance in music. Music and language also both depend on structural rules and have semantics attached to them (although debatably language has more complex grammar and more concrete meanings; the above is reviewed in Patel, 2003). As well, because language and music are both human universals, many researchers have argued that the two may have co-evolved, although it is debatable whether music is a spandrel (a byproduct rather than a direct adaptive product of evolution) of language (e.g. Pinker, 1997) or whether music functions to accomplish such tasks as mate attraction, child rearing, group cohesion, et cetera (Fitch, 2005). The idea that music and language may be inherently linked in humans lends support to the idea that model species for human language may be especially suitable for biomusicology.

In addition, some vocal learners have vocal and perceptual auditory traits that seem more akin to music production and perception than language. For example, birdsong was termed "song" because of its perceived resemblance to music. Birdsong is used for mate attraction (McGregor, 1991), which is one of several proposed purposes of human music that was originally theorized by Darwin (Fitch, 2005). Additionally, at least some vocal learners, including avian

species, are able to attend to and move to the rhythm of a song (Schachner, Brady, Pepperberg & Hauser, 2009). Furthermore, it has been proposed — also originally by Darwin (see Darwin, 1871) — that music and language may both have evolved from a primitive version of modern-day music (Fitch, 2005; Masataka, 2009). It is possible that the songs of nonhuman vocal learners may be similar to this song-like "protolanguage" which is thought to have existed in human ancestors.

Pitch Perception

The focus in my dissertation is on the comparative processing of pitch. Pitch perception is an important component of both language and music. In language, pitch plays a role in the prosody of speech (the intonation, stress, and timing of speech; Pinker, 1994). It is also a cue for dominance (Puts, Gaulin & Verdolini, 2006) and distinguishes males and females (Vuorenkoski et al., 1978). In music, pitch is fundamental. Pitch is interesting as a psychological phenomenon because it is the name given to the percept of how high or low in frequency a sound is heard. Typically, pitch corresponds to the frequency of a sound; however, this is not always the case (Dowling & Harwood, 1986).

Pitch can be evaluated in an absolute or relative fashion. Evaluating the absolute pitch of a sound means identifying a sound without an external referent (e.g., to know that a note is an "A" note from Western music without having heard another note with which to compare it), while evaluating the relative pitch of a sound means comparing the difference in pitch of two or more notes. Both

4

these strategies may be used when evaluating a speaking voice. For example, the absolute pitch of a voice can lead one to make judgments about whether the speaker is male or female (Vuorenkoski et al., 1978), while the relative pitch between parts of utterances may give us information about the speaker's emotions (Curtis & Bharucha, 2010).

In music, absolute and relative pitch have similar effects as in language. In terms of absolute pitch, high-pitched melodies sound more submissive while low-pitched melodies sound more dominant (Huron, Kinney & Precoda, 2006). At the same time, the relative pitch, along with other properties such as tempo, set the mood for the instrument that is playing (Curtis & Bharucha, 2010).

An additional wrinkle to this story involves the nature of absolute pitch. Absolute pitch can be broken down into two different types: pitch height and pitch chroma (Shepard, 1982). Pitch height is a log-linear perception of pitch. According to pitch height, the higher a pitch is in frequency, the higher the perceived pitch. Pitch chroma, on the other hand, is sometimes described as the quality of a note. Pitch chroma repeats each time frequency is doubled. This doubling relationship is recognized in all human musical systems and the interval between a note and its double is referred to as an octave (Crickmore, 2003). In the mid-nineteenth century, Drobisch proposed a helix as a model of human absolute pitch perception (see Figure 1-1). The helix combines the linear pitch height perception and the circular pitch chroma perception into one continuous pitch space. Although there is work that suggests that both pitch height and chroma are

used by humans (Shepard, 1982), it is not clear whether this model could be applied to other species.

Model species

In my dissertation I focus on a comparison between humans and wild-caught black-capped chickadees. I chose to work with chickadees because they represent an ideal group for my research; they are a common species of vocal learning oscine songbird that can be easily captured from wild populations. Being wild-caught makes their behaviour in the laboratory more likely to reflect their natural behaviour than is typical for laboratory-strain animals and allows us to have a suitable sample size on which to form interpretations.

Additionally, black-capped chickadees have been studied fairly extensively in terms of their absolute pitch height. They are no exception to the rule that songbirds tend to have heightened absolute pitch sensitivity compared to humans. Black-capped chickadees can accurately discriminate eight ranges of frequencies, while humans (excluding the rare humans with absolute pitch) are less accurate with only three ranges (Lee et al. 2006; Weisman et al. 2010; Weisman et al., in press). Their use of relative pitch, and how it ties in with their fine-tuned absolute pitch abilities is therefore of interest. I was also interested in whether black-capped chickadees perceive absolute pitch chroma when assessing the absolute pitch of stimuli.

So far I have noted the advantages of studying pitch in black-capped

chickadees are the fact that they are wild-caught, and seem to have strong pitch height abilities. Although all these points are relevant, black-capped chickadees are not the only species that match these criteria. What makes black-capped chickadees especially interesting to study with regards to their pitch, however, is their song. The songs of songbirds are vocalizations that are used for mate attraction (McGregor, 1991). Song is typically one of the more complex vocalizations of a given species (Catchpole and Slater 2008), but on the surface this does not appear to be the case for the black-capped chickadee (Dixon and Stefanski 1970; Mennill and Otter 2007). Except for a couple of isolated populations (Kroodsma et al. 1999; Gammon and Baker 2004), the song is always composed of the same two notes: the *fee* note, and the *bee* note. Unlike songs of many other songbird species, black-capped chickadees in all other regions sing a *fee* note followed by a *bee* note in every iteration of the song. The notes themselves are also fairly simple in structure, being tonal in nature with no harmonics. There is a slight glissando, or decrease in pitch, from the beginning of the *fee* note (first note) to the end of the *fee* note. The *bee* note is lower in pitch than the *fee* not, and is fairly constant in pitch throughout its duration. What is interesting is that while the absolute pitch of the two notes can vary even within individuals, the relative pitch of the two notes remains constant. As this pitch ratio also carries over long distances unlike other stable features (such as duration), it has often been proposed to be an important sexual signal (Horn et al. 1992; Weisman & Ratcliffe 1989; Weisman et al. 1990). In fact, some work has found

that dominant males produce this pitch ratio with higher fidelity (Christie, Mennill, & Ratcliffe, 2004). Other work has shown that female chickadees attend to this ratio (Weisman & Ratcliffe, 2004).

This relative pitch relationship in the *fee bee* song of the black-capped chickadee is unusual because songbirds are known to have adept absolute pitch abilities compared to the majority of humans (excluding rare humans with musical absolute pitch; Weisman et al., 2010). Songbirds also seem to more readily use absolute pitch information over relative pitch information, although they can learn to use either (Cynx, Hulse, & Polyzois, 1986; Hulse et al. 1984; Hulse & Cynx 1985). The use of relative pitch in the *fee bee* song is therefore surprising, and may implicate black-capped chickadees as more adept relative pitch processors than the average songbird.

Current studies

My goal in this dissertation was to understand both specific adaptations and fundamental shared components of pitch perception in humans and chickadees by using operant paradigms. I aimed to look in more depth at the two kinds of pitch that have been less explored in chickadees to date: relative pitch and absolute pitch chroma. Chapters 2 and 3 examine relative pitch perception and chapters 4 and 5 look at absolute pitch chroma perception. Chapter 2 is a study I conducted to understand what role the *fee bee* song plays in black-capped chickadee relative pitch perception. The birds were trained to discriminate

different relative pitch intervals by minimizing the opportunity for the birds to use absolute pitch. These relative pitch intervals varied in terms of how similar they were to natural *fee-bee* song by adjusting the relative pitch interval and the structure of the actual note types by using either sinewave tones or shifted natural notes. Chapter 3 compares human and chickadee relative pitch perception. In this study I trained both humans and chickadees to discriminate different musical chords and transfer their discrimination to a novel absolute pitch. Chapters 4 and 5 compare humans and chickadees in terms of their absolute pitch chroma perception. Chapter 4 is an in depth study of human absolute pitch chroma perception. The goal here was to design an operant paradigm that shows human chroma perception that can be used to test other species. Chapter 5 applies the operant paradigm from chapter 4 to black-capped chickadees to determine whether this species perceives chroma. Finally, chapter 6 discusses the preceding chapters as a group and interprets the findings.
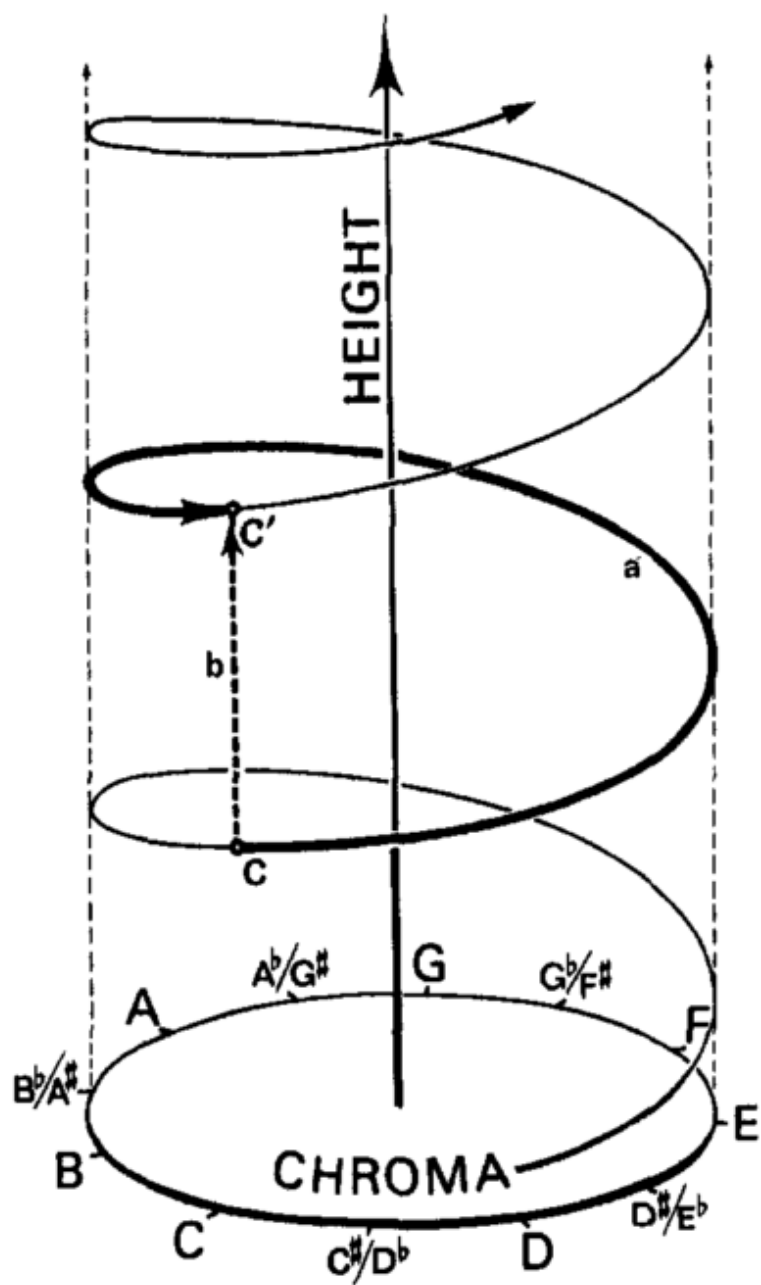
Figure 1-1.

References

Catchpole C. K., & Slater P. J. B. (2008). *Bird song: Biological themes and variations.* Cambridge: Cambridge University Press.

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.

Curtis, M. E. & Bharucha, J. J. (2010). The minor third communicates sadness in speech, mirroring its use in music. *Emotion, 10*(3), 335-348.

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.

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

Dixon, K. L., Stefanski, R. A. (1970). An appraisal of the song of the black-capped chickadee. *The Wilson Bulletin, 82*, 53-62.

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, FL, USA:

Academic Press.

Fitch, W. T., (2006). On the biology and evolution of music. *Music Perception, 24*(1), 85-88.

Fitch, W.T., (2005). The evolution of music in comparative perspective. *Annals of the New York Academy of Sciences, 1060*, 29-49.

Gammon, D. E., Baker, M. C. (2004). Song repertoire evolution and acoustic divergence in a population of black-capped chickadees, *Poecile atricapillus*. *Animal Behavior, 68*, 903-913.

Haesler, S., Rochefort, C., Georgi, B., Licznerski, P., Osten, P., & Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockout of FoxP2 in songbird basal ganglia nucleus area X. *PLoS Biology, 5*(12), 2885-2897.

Hauser, M.D. McDermott, J. (2003). The evolution of the music faculty: a comparative perspective. *Nature Neuroscience, 6*(7), 663-668.

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.

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*(1), 38-54.

Hulse, S. H., & Cynx, J. (1985). Relative pitch perception is constrained by absolute pitch in songbirds (*Mimus, Molothrus*, and *Sturnus*). *Journal of Comparative Psychology, 99*, 176-196.

12

Huron, D., Kinney, D. & Precoda, K. (2006). Influence of Pitch Height on the

    Perception of Submissiveness and Threat in Musical Passages. *Empirical*

    Musicology Review, 1(3), 170-177.

Justus, T. & Hutsler, J. J. (2005). Fundamental issues in the evolutionary

    psychology of music: assessing innateness and domain specificity. *Music*

    Perception 23(1), 1-27.

Kroodsma, D. E., Byers, B. E., Halin, S. L., Hill, C., Minis, D., Bolsinger, J. R.,

    Dawson, J., Donelan, E., Farrington, J., Gill, F. B., Houlihan, P., Innes, D.,

    Keller, G., MacAulay, L., Marantz, C. A., Ortiz, J., Stoddard, P. K.,

    Wilda, K. (1999). Geographic variation in black-capped chickadee songs

    and singing behavior. *The Auk, 116*, 387-402.

Lee, T. T. Y., Charrier, I., Bloomfield, L. L., Weisman, R. G., & Sturdy, C. B.

    (2006). Frequency-range discriminations and absolute pitch in black-

    capped chickadees (*Poecile atricapillus*), mountain chickadees (*Poecile*

    gambeli), and zebra finches (*Taeniopygia guttata*). *Journal of comparative*

    psychology, *120*(3), 217–28.

Masataka, N. (2009). The origins of language and the evolution of music: a

    comparative perspective. *Physics of Life Reviews, 6*, 11-22.

McGregor, P. K. (1991). The singer and the song: On the receiving end of bird

    song. *Biological Reviews, 66,* 57-81.

Mennill, D. J., & Otter, K. A. (2007). Status signaling and communication

    networks in chickadees: Complex communication with a simple song. In:

13

Otter KA (Ed.), *The ecology and behavior of chickadees and titmice: an integrated approach* (pp. 215-233). New York: Oxford University Press.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*(7), 674-681.

Pinker, S. (1994). *The Language Instinct*. New York, NY: HarperCollins Publishers Inc.

Pinker, S. (1997). *How the mind works*. New York, NY: W.W. Norton & Company Inc.

Puts, D. A., Gaulin, S. J. C. & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior, 27*, 283–297.

Schachner, A., Brady, T.F., Pepperberg, I.M., & Hauser, M.D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology, 19*(10), 831-836.

Shepard, R.N. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review, 89*(4), 305-333.

Smith, S. M. (1991). *The Black-Capped Chickadee: Behavioral Ecology and Natural History*. Ithaca, New York: Cornell University Press.

Tyack, P. L. (2008). Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *Journal of Comparative Psychology, 122*(3), 319-331.

Vuorenkoski, V., Lenko, H.L., Tjernlund, P., Vuorenkoski, L., & Perheentuppa, J. (1978). Fundamental voice frequency during normal and abnormal growth, and after androgen treatment. *Archives of Disease in Childhood, 53*, 201-209.

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

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

Weisman, R. G. & Ratcliffe, L. (2004). Relative pitch and the song of black-capped chickadees. *American Scientist, 92*(6), 532-539.

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.G., Mewhort, D.J.K., Hoeschele, M., & Sturdy, C.B. (in press). New perspectives on absolute pitch in birds and mammals. In E.A. Wasserman & T.R. Zentall (Eds), Handbook of Comparative Cognition. Oxford: Oxford University Press.

Chapter 2 Biological relevance of acoustic signal affects discrimination performance in a songbird[1]

Introduction

Pitch can be classified based on absolute or relative features. Absolute pitch is the ability to classify a note without an external referent. In the human literature, absolute pitch is typically used to refer to members of a very small proportion of the population that are able to recognize a note type without a comparison (e.g., an "A" note; Baharloo, Johnston, Service, Gitschier, & Freimer 1998). However, a level of absolute pitch that is less fine-grained occurs in the general population (e.g., distinguishing male and female voices; Bachorowski & Owren 1999). In comparative work with humans, pigeons, rats, and songbirds, it has become clear that songbirds out-perform the other species at absolute pitch tasks (Weisman et al. 1998; Weisman et al. 2004; Friedrich et al. 2007; for review see Weisman et al. 2006). Potentially because of this strong absolute pitch ability, there is some evidence to suggest that songbirds tend to rely on absolute pitch when solving relative pitch tasks (Cynx et al. 1986; Hulse et al. 1984; Hulse & Cynx 1985). Relative pitch refers to evaluating relations among notes. These relations are evaluated using the "interval" or frequency ratio between two or more notes, as frequency is perceived on a logarithmic scale. Although most research demonstrates the importance of absolute pitch cues for songbirds, there are some cases where relative pitch may be important to songbirds in nature.

One such case is the song of the black-capped chickadee (*Poecile atricapillus*), a common North American songbird and my study species. Black-

capped chickadees have a tonal, two-note song called the *fee-bee* song (Ficken et al. 1978). The *fee-bee* song is used for mate attraction and territory defense (e.g., Otter & Ratcliffe, 1993; Mennill et al. 2002). Importantly, although the two notes in the song can be sung at different absolute pitches within an individual, the relative pitch interval between the notes remains highly consistent both within and across individuals (Horn et al. 1992; Weisman et al. 1990; see Figure 2-1). Moreover, dominant birds are more adept at maintaining this pitch interval consistency (Christie et al. 2004), and pitch interval information is conserved over long transmission distances (Christie et al. 2003), making it putatively an important signal of dominance status for the receiver. Because of the potential importance of the pitch interval contained within the *fee-bee* song, relative pitch assessment (i.e., interval comparisons among notes) seems to be an important ability in black-capped chickadees. This is in contrast to studies showing other species of songbirds' typical reliance on using absolute pitch when solving relative pitch based discrimination tasks (e.g., Cynx, et al. 1986; Hulse et al. 1984).

Inspired by the idea that relative pitch may be important for black-capped chickadees, I set out to understand whether pitch discrimination abilities fluctuate (i.e., whether two stimuli are treated the same through generalization or different through discrimination) based on the biological relevance of the stimuli being compared. Therefore, I wondered whether discriminations are learned in fewer trials by animals when discriminations involve stimuli that have properties shared

by naturally-occurring signals to which animals attend. I examined whether pitch interval discriminations can be influenced by the species-specific biological relevance of different pitch interval types. I trained male and female black-capped chickadees to discriminate sine-wave tone pitch intervals (hereafter referred to as synthetic stimuli) in a Go/NoGo task. In addition to using synthetic stimuli, birds were also trained with actual *fee-bee* songs that were shifted to contain both species-typical and atypical pitch intervals (*fee-bee* song stimuli hereafter referred to as natural stimuli). Using natural stimuli allowed us to explore whether the spectral qualities of song notes, as opposed to using synthetic sine-wave tones, would facilitate discrimination learning in black-capped chickadees. By transferring birds from discriminating synthetic stimuli to natural stimuli (and vice versa), we were able to look at whether pitch was likely to control the discrimination.

We had three main predictions that lead us to pursue this research. (1) The species-typical black-capped chickadee song pitch interval would be easier to discriminate than other similarly structured intervals whether by using absolute or relative pitch. (2) Natural stimuli would be easier for the birds to discriminate than synthetic stimuli. (3) There would be no difference between male and female chickadees completing the task based on similar performance between male and female chickadees in most operant experiments (e.g., Guillette et al. 2009; Bloomfield et al. 2008; Charrier et al. 2005).

Method

Subjects

Thirty-six black-capped chickadees (18 males, 18 females) naïve to the
experimental stimuli were used in this experiment. Some birds had prior operant
conditioning experience with chick-a-dee call note stimuli (e.g., Guillette et al.
2009), and past experience was counter-balanced across training groups and
sexes. Birds were captured in and around Edmonton, Alberta, Canada (Stony
Plain, 53˚46N, 114˚01W; North Saskatchewan River Valley, 53˚53N, 113˚53W;
Mill Creek Ravine, 53˚52N, 113˚47W) or at Barrier Lake Field Station,
Kananaskis, Alberta, Canada (51˚02N, 115˚03W). All birds were determined at
the time of capture to be at least one year of age by examining the colour and
shape of their outer tail retrices (Pyle 1997). Birds were tested between October
2008 and March 2010. The order birds were tested was counter-balanced such
that all groups and sexes were run across all seasons.

Apparatus

During the experiment, birds lived in modified colony room cages (30 ×
40 × 40 cm). Each cage was placed inside a ventilated, sound-attenuating chamber
and contained three perches and dispensers for water and grit. The chambers were
illuminated by a 9-W, full spectrum fluorescent bulb. An opening on the side of
the cage (11 × 16 cm) gave each bird access to a motor-driven feeder (Njegovan

et al. 1994). Both the feeder and the perch closest to the feeder (request perch) had infrared cells to monitor the position of the bird. A personal computer connected to a single-board computer (Palya & Walter 2001) scheduled trials and recorded responses to stimuli. Stimuli were played from CD through either a Cambridge A300 Integrated Amplifier (Cambridge Audio, London, England) or a NAD310 Integrated Amplifier (NAD Electronics, London, England) and then through a Fostex FE108 Sigma full-range speaker (Fostex Corp., Japan; frequency range 200–16000 Hz) located beside the feeder. See Sturdy and Weisman (2006) for a detailed description of the apparatus.

Stimuli

Based on Weisman et al. (1994) and Njegovan and Weisman (1997), we constructed an array of synthetic and natural note pairs with particular descending pitch intervals described below. The note pairs within the array are henceforth referred to as pitch intervals because this is the manipulated feature in our study. In line with previous work, we used the same pitch for Note 2 (i.e., the second note in each pair called the *bee* note in *fee-bee* song; see Figure 2-1) in each pitch interval, for all pitch interval categories. Therefore, birds could not use the absolute frequency of Note 2 to determine the pitch interval type to which a particular interval belonged. We designed an array with 18 unique Note 2 frequencies between 2519 and 4116 Hz with steps of 21/24 (by quarter tones or half semitones in Western music) between each Note 2 value. Note 1 was then

synthesized at a consistent pitch interval from Note 2 for each pitch interval category. We had five different pitch interval ratios in our study: 1.07, 1.13, 1.19, 1.25, 1.31, thus this set of pitch interval ratios combined with the 18 Note 2 frequencies generated an 18 (Note 2) × 5 (Note 1) array of pitch intervals. These pitch intervals were chosen because they increased at equal steps in ratio (i.e., a ratio difference of 0.06). The pitch interval ratio of 1.13 is the species-typical interval we chose because it matches the average ratio of the two notes in black-capped chickadee *fee-bee* song (Weisman et al. 1990; Horn et al. 1992), and the accuracy at which the pitch interval is produced is an indicator of male quality (Christie et al. 2004).The full stimulus array is displayed in Table 1. However, because of pilot testing, a subset of the total array of stimuli was used for training to improve the chances that a high number of individuals could complete the task (see highlighted rows in Table 1). Specifically, we used a Note 2 of 2996 Hz and 3362 Hz because they are separated by a reasonably large difference in absolute pitch, but both remain roughly within one standard deviation of natural *bee* notes (Weisman et al. 1990). Stimuli were generated using synthetic stimuli, and again using natural stimuli, both of which are described separately below. See Figure 2-2 for an example of the stimuli used.

Synthetic Stimuli

Tones 440 ms in duration were synthesized using SIGNAL version 4.04.15 (Engineering Design, Berkeley, USA). This duration is roughly the same

duration as natural *fee* and *bee* notes in black-capped chickadee song (X = 357

ms, Ficken et al. 1978; X = 407 ms *fee* notes, Christie et al. 2004; X = 406 ms,

Lohr 2008; X = 474 ms, based on measurements taken by Hoeschele et al. 2010)

and is the duration used in prior studies upon which we were basing our work

here (Weisman et al. 1994; Njegovan & Weisman 1997). All tones were tapered

for the first and last 5 ms to avoid transients. Tone pairs from Table 1 were then

concatenated such that the internote interval was kept at a constant 100 ms which

also approximates natural song (X = 135 ms, Ficken et al. 1978; X = ~118 ms

based on based on measurements taken by Hoeschele et al. 2010).

Natural Stimuli

We used 10 *fee-bee* songs, recorded by Proppe et al. (in prep) in Elk Island

National Park (53°37N, 112°53W). By using songs from this location, we could

ensure that none of the birds used in our experiment were familiar with the

individuals used for discriminative stimuli generation. Selection of 10 songs was

done randomly except that recorded songs had to be high quality (i.e., little to no

visible noise observed in a sound spectrogram or audible interference on the track

monitored using a −35 dB below peak amplitude cutoff using SIGNAL) and no

more than one song was taken from each site within the park (separated by 300 m

to 14 km X +SEM = 5.82+0.98 km) to ensure we did not use the same bird twice.

The selected songs were edited into their note components (i.e., *fee* and

bee notes) using SIGNAL and then band pass filtered using Goldwave version

5.12 (Goldwave, Inc., St. John's, Canada) to remove any background noise outside of the song frequencies. The individual notes were then shifted to appropriate frequencies as described above and displayed in Table 1 using Goldwave. *Fee* notes were used as Note 1 and *bee* notes as Note 2 as in their natural song. All notes (10 *fee* notes and 10 *bee* notes) were pseudorandomly selected for use in the array such that all notes were used equally (nine times each), all were used evenly across the five pitch interval ranges (either for one or two pairings for each pitch interval type), over a wide span of absolute pitches (*fee* note range X = 2067 Hz and *bee* note range X =1060 Hz), and notes were never paired with their natural partner (i.e., the *fee* note from song 1 was never paired with the *bee* note from song 1), but were paired with all other possible notes. The notes were concatenated with an internote interval of 100 ms, the same interval used for the synthetic stimuli to approximate the natural duration of *fee-bee* song.

To further control for and ensure that the amplitude of the individual stimuli did not influence how the birds perceive frequency, all pitch intervals (synthesized and natural stimuli) were standardized at two different amplitudes (70 dB and 80 dB sound pressure level [SPL] with a reference value of 20 microPA) to reduce the effects of amplitude on frequency judgments (Moore 1989). All pitch intervals therefore had two versions: one where the first note was 80 dB and the second was 70 dB, and vice versa. The amplitude of each note was standardized in the sound attenuating chamber at the approximate position of the

24

bird's head using a Brüel & Kjær Type 2239 (A weighting, slow response; Nærum, Denmark) dB meter to ensure that all frequencies were presented at the same relative level through the speaker (± 0.3 dB) and then were measured periodically during the experiment with a Radio Shack Sound Level Meter (A weighting, slow response; Radio Shack, Fort Worth, USA) to ensure the absolute amplitude of the stimuli was consistent across birds.

Procedure

Nondifferential Training

Once a bird had learned to use the request perch and feeder to obtain food, Nondifferential Training began. This stage was designed to promote high, uniform responding to all stimuli before discrimination training. During this stage, the bird received food for responding to all future discrimination training stimuli (both future S+ and future S- stimuli), and a 30 s timeout with the houselights off if the bird left the request perch before a stimulus had completed playing (interrupted trial). The 36 birds were divided into three groups of 12 birds each (six males, six females; 1.13 Group, 1.19 Group and 1.25 Group named according to which pitch interval ratio was the future S+ for each group). Each group was trained with three neighbouring pitch interval types: a small (future S-), a medium (future S+), and a large (future S-) pitch interval that were separated by pitch interval ratio step sizes of 0.06. (1.13 Group; 1.07, 1.13, 1.19 ratio pitch intervals, 1.19 Group; 1.13, 1.19, 1.25 ratio pitch intervals, and 1.25 Group; 1.19, 1.25, 1.31

pitch interval ratios, see Table 1) and heard all six of the exemplars for their assigned pitch intervals: both the synthetic and natural versions, each at two different amplitudes, for a total of 24 stimuli. Further details on the nondifferential procedure can be found in Hoeschele et al. (2012).

Discrimination 1 Training

This stage was identical to Nondifferential Training except that each bird only heard half the stimuli it had been trained with in Nondifferential Training (either all synthetic or all natural), and some of the stimuli that had previously resulted in food now resulted in a 30-s intertrial interval with the houselights off if the bird entered the feeder (S- stimuli). Within each group (1.13 Group, 1.19 Group, and 1.25 Group, see Table 1) half of the birds (three males, three females) were trained with synthetic stimuli and half were trained with natural stimuli. For each group, pitch intervals with the middle ratio (i.e., the pitch interval ratio for which the group was named e.g., 1.13 for the 1.13 Group) served as S+ stimuli, while pitch intervals with the small and large ratios (e.g., 1.07 and 1.19 in the 1.13 Group) served as S- stimuli. Thus the birds in the 1.13 Group were trained to respond to pitch intervals with a ratio matching the species-typical interval (1.13), while withholding responding to species-atypical pitch intervals (with ratios of 1.07 and 1.19). Birds in the 1.25 Group received no training with the species-typical interval as they were only trained with the 3 larger intervals. The criterion

to complete this stage was six blocks at a discrimination ratio (DR) ≥ 0.80 with

the last two blocks being consecutive. See response measures for DR calculations.

Discrimination 2 Training

This stage was identical to Discrimination 1 Training except that the birds

were trained with the stimulus type not previously used in Discrimination 1

Training (i.e., synthetic if originally trained with natural in Discrimination 1

Training). The criterion for this stage was the same as Discrimination 1 Training.

Response Measures

To determine whether the birds had successfully learned to discriminate

the S+ pitch interval from the two S- pitch intervals we calculated a

discrimination ratio (DR). The DR is a standard measure that falls between 0 and

1, where 0 means all visits to the feeder followed S- stimuli, 0.5 represents chance

with equal responding to S+ and S- stimuli, and 1 represents perfect

discrimination of S+ from S-. Importantly, however, we adjusted this measure by

excluding interrupted trials (see Guillette, Farrell, Hoeschele, Nickerson, Dawson

& Sturdy 2010 for details on the calculations).

Statistical Analyses

Blocks to Criterion

We conducted this analysis to determine which groups learned the discriminations in the fewest trials and to test our three initial predictions: (1) The species-typical black-capped chickadee song pitch interval would be easier to discriminate than other similarly structured intervals. (2) Natural stimuli would be easier for the birds to discriminate than synthetic stimuli. (3) There would be no difference between male and female chickadees completing the task.

We conducted a repeated-measures MANOVA that included group (1.13 Group, 1.19 Group, and 1.25 Group) and sex as categorical predictors. The dependent measures for this analysis were the number of blocks required to reach criterion for each of the two stimulus types (synthetic and natural). This analysis allowed us to compare how many trials it took birds to complete the task irrespective of the slope of the learning curves.

Acquisition of Pitch Interval Categories

While blocks to criterion is an overall measure of trials taken to complete the different discriminations, the current analysis was conducted to determine whether there was any variation in learning curves among S- pitch intervals. Given that we were interested in how the species-typical song pitch interval affected performance, it seemed especially prudent to evaluate this S- interval

separately in the 1.19 Group (the only group where the species-typical song interval is one of the S- intervals).

To complete this analysis, we conducted two repeated measures MANOVAs: The first assessing the discrimination using synthetic stimuli, and a second assessing the discrimination using natural stimuli. We included group (1.13 Group, 1.19 Group, and 1.25 Group) and sex as categorical predictors. For our dependent variables, we calculated separate DRs for each S- pitch interval type (smaller and larger than S+) compared to the S+ pitch interval type for each block of training. This DR was calculated by using only the percent response for one of the S- pitch interval types. The birds learned the discrimination in differing numbers of trials, therefore acquisition curves were Vincentized to seven blocks of data (the minimum number of blocks required by birds to learn the discrimination) so that acquisition could be compared across birds. Vincentizing was calculated by dividing the total number of actual blocks a bird performed to reach criterion (e.g., 31) by seven Vincentized blocks into whole numbers (in this example, four remainder three). Vincentized blocks were then computed by grouping the total number of blocks a bird completed into groups the size of the quotient (in this example, four) and placing the any remaining blocks into the middle block (4th block, in this example, the 4th block would have $4 + 3 = 7$ actual blocks grouped in) and then averaging the DR across grouped blocks (see Kling & Riggs 1971). After Vincentizing, there were 14 dependent variables for each of the two (synthetic and natural) analyses: 7 Vincentized blocks $\times$ 2 pitch

interval types (smaller and larger pitch intervals from the S+). These analyses allowed us to compare performance across acquisition irrespective of the number of trials.

Final Performance

While the previous analysis studied performance across the discrimination, all birds were required to reach the criterion of a DR $\geq$ 0.8 for six blocks with the last two being consecutive. However, variation in DRs was still possible beyond this criterion (i.e., a bird could maintain a DR of 0.80 and another bird could be near perfect with a DR of 0.96). While Vincentizing the data allows assessment of the overall shape of the learning curves, it gives little information about final level of discrimination performance. In this analysis, we looked at discrimination performance during the final block. A repeated measures MANOVA using group (1.13 Group, 1.19 Group, and 1.25 Group) and sex as categorical predictors and the DR for the final actual block for each discrimination (synthetic and natural) and each pitch interval (small and large) as four dependent variables. This allowed us to examine the last 500 trials that each bird needed to reach criterion irrespective of how long it took to do so and any changes over blocks of trials.

Savings

This analysis assesses whether learning transferred from one discrimination (e.g., synthetic) to the other (e.g., natural). Because the natural stimuli contained spectral differences among notes rather than just pitch, we wondered whether the pitch information would transfer from one task to the other. To complete this analysis, we calculated one-tailed z scores for each individual bird using binomial tests for dichotomous data to determine whether the proportion of responses that were directed to S+ stimuli were significantly above chance (1/3 chance because 1/3 of the trials were S+, note that this is different from the DR where S+ responses were weighted equally to S- responses) during the first block of each of the two discriminations (synthetic and natural). This allowed us to assess whether any birds responded significantly above chance during the first block and if so how many birds did so during which discrimination.

Results

Blocks to Criterion

Figure 2-3 shows the average blocks to criterion for each group (i.e., 1.13 Group, 1.19 Group, 1.25 Group) during each of the two discriminations (i.e., synthetic and natural). The discrimination with synthetic stimuli took more trials to solve the less similar the discriminated S+ interval was to the species-typical

interval, but the discrimination using natural stimuli was solved in a similar number of trials for all discrimination groups. The MANOVA for blocks to criterion revealed a significant main effect of group (i.e., 1.13 Group, 1.19 Group, 1.25 Group; $F_{2,30} = 4.13$, $P = 0.026$, partial $\eta^2 = 0.216$) and discrimination (i.e. synthetic or natural; $F_{1,30} = 81.78$, $P < 0.001$, partial $\eta^2 = 0.732$) and a significant interaction between group and discrimination ($F_{2,30} = 4.70$, $P = 0.017$, partial $\eta^2 = 0.239$). The main effect of discrimination revealed that the natural stimuli ($X = 11.81$) were learned in significantly fewer blocks than the synthetic stimuli ($X = 27.03$). We explored the interaction between group and discrimination with further Tukey's post-hoc tests and found that while the groups did not differ significantly when the discrimination was natural (1.13 Group $X = 11.58$, 1.19 Group $X = 12.08$, Irrelevent $X = 11.75$; all Ps = 1.000), the 1.25 Group ($X = 32.67$) took significantly more trials to learn than the 1.13 Group ($X = 20.00$; $P = 0.002$) when the discrimination was synthetic. The 1.19 Group ($X = 28.42$) was not significantly different from either other group when the discrimination was synthetic (all Ps ≥ 0.083). This suggests that the main effect of group was driven entirely by the discrimination using synthetic stimuli.

Acquisition of Pitch Interval Categories - Synthetic Stimuli

Figure 2-4 shows how birds were performing across Vincentized blocks. In the repeated-measures MANOVA for synthetic Vincentized blocks, we found a significant main effect of Vincentized blocks ($F_{6,180} = 194.20$, $P < 0.001$, partial

$\eta2 = 0.866$) such that, as Vincentized blocks increased, DRs increased. We also found a significant interactions of a pitch interval by group ($F2,30 = 10.44$, $P <$ 0.001, partial $\eta2 = 0.410$). Tukey's post-hoc tests exploring the pitch interval by group interaction showed that the 1.19 Group learned the smaller S- pitch interval in their discrimination (the species-typical pitch interval; $X = 0.72$) to a significantly higher DR than the larger S- pitch interval (a species-atypical pitch interval; $X = 0.67$; $P = 0.014$). Both other groups (1.13 Group small $X = 0.70$; 1.13 Group large $X = 0.72$; 1.25 Group small $X = 0.67$; 1.25 Group large $X =$ 0.71) learned the two S- pitch intervals at DRs that were not significantly different (both groups had species-atypical pitch intervals as both of their S-s; all Ps $\geq$ 0.127).

Acquisition of Pitch Interval Categories - Natural Stimuli

In the repeated-measures MANOVA for natural Vincentized blocks, we found a significant main effect of Vincentized blocks ($F6,180 = 147.46$, $P <$ 0.001, partial $\eta2 = 0.831$) such that, as Vincentized blocks increased, DRs increased. In addition, there was a significant main effect of pitch interval (smaller and larger S-; $F1,30 = 7.15$, $P = 0.012$, partial $\eta2 = 0.192$). We also found two significant interactions: a pitch interval by group interaction ($F2,30 =$ 4.23, $P = 0.024$, partial $\eta2 = 0.220$) and a pitch interval by block by sex interaction ($F6,180 = 2.27$, $P = 0.039$, partial $\eta2 = 0.070$). Tukey's post-hoc tests exploring the pitch interval by group interaction revealed that the 1.25 Group

learned the larger S- pitch interval (X = 0.79) in their discrimination to a significantly higher DR than the smaller S- pitch interval (X = 0.74; both species-atypical pitch intervals; P = 0.011). Both other groups (1.13 Group and 1.19 Group) learned the two S- pitch intervals to DRs that were not significantly different from one another (all Ps ≥ 0.742). This suggests that the main effect of pitch interval was driven entirely by the 1.25 Group. Tukey's post-hoc tests that explored the pitch interval by block by sex interaction showed that, during the first two Vincentized blocks, males but not females responded at a significantly higher DR to the larger pitch interval (X = 0.60) than the smaller pitch interval (X = 0.54; P ≤ 0.011).

Final Performance

Figure 2-5 shows the DR during the final block averaged across both discriminations to each pitch interval. While the 1.13 Group had similar DRs for both pitch intervals, the 1.19 Group had a significantly higher DR for their smaller S- pitch interval (species-typical pitch interval), and the 1.25 Group had a higher DR for their larger pitch interval (furthest from the species-typical interval). In the repeated-measures MANOVA on the final blocks of discrimination, we found a significant main effect of sex ($F_{1,30} = 5.85$, $P = 0.022$, partial $\eta^2 = 0.163$) that revealed that females responded at a significantly higher DR during their final block. There was also a significant main effect of discrimination (synthetic and natural; $F_{1,30} = 26.17$, $P < 0.001$, partial $\eta^2 =$

0.466) that revealed the discrimination using natural stimuli generated a significantly higher DR during the final block than the discrimination using synthetic stimuli. Further, there was a significant interaction between pitch interval (smaller and larger S-) and group (1.13 Group, 1.19 Group, 1.25 Group; $F_{2,30} = 6.68$, $P = 0.004$, partial $\eta 2 = 0.308$). Tukey's post-hoc tests conducted on this interaction showed that the 1.25 Group had a significantly higher DR for the larger S- pitch interval than the smaller S- pitch interval ($P = 0.034$) and the 1.25 Group had a significantly lower DR for the smaller pitch interval in their discrimination than the 1.19 Group had for the smaller pitch interval in their discrimination (species-typical pitch interval; $P = 0.035$).

Savings

Table 2 contains the number of subjects within each condition that were above chance. Figure 2-6 shows graphically the overall DRs separated by sex, group, and starting discrimination, during the first block of discrimination 1 and 2. One bird responded above chance during the first block of Discrimination 1 Training ($z = 4.57$, $p < 0.001$). This bird was a female in the 1.13 Group that completed the discrimination using natural stimuli first. An additional 19 of the 36 birds responded above chance during the first block of the Discrimination 2 Training (all $zs \geq 1.93$, all $ps \leq 0.027$). Eight of these birds learned the tone discrimination second, and 11 learned the discrimination using natural stimuli second. There did not appear to be an advantage for either sex or group (i.e., 1.13

Group, 1.19 Group, 1.25 Group) as both sexes and all three groups had members that were responding above chance during the first block of the second discrimination.

Discussion

The biological relevance of stimuli appears to impact both how many trials are needed and how well stimuli are discriminated in a task with several stimulus categories. We found that birds trained with their species-typical *fee-bee* song pitch interval as the S+ category and two species-atypical pitch intervals as S- categories in a Go/NoGo discrimination learned the discrimination in fewer trials than birds that were trained with all species-atypical pitch intervals. Additionally, except for the largest pitch interval we used, which was the most dissimilar interval from the species-typical interval, the more similar pitch intervals were to the species-typical pitch interval, the more accurately they were discriminated. By conducting the discrimination with both synthetic and natural stimuli, we also observed an advantage for birds learning discriminations using natural stimuli. Over half the birds, when transferred from one stimulus type to the other, showed significant savings, which suggests they were attending to pitch cues to solve the task on both discriminations. Finally, we found that females reached a higher level of accuracy in the final stages of discrimination, but that males were more accurate at discriminating larger pitch intervals early on in the discrimination.

We also observed differences among discrimination groups that suggest that, although all groups were trained with pitch intervals that increased with the same step sizes, the absolute size of the pitch intervals affected performance. As the S+ and S- pitch intervals within a discrimination group moved further away from the species-typical song pitch interval, birds required more trials to discriminate the pitch intervals. This suggests that the chickadees may have a range of relative pitch interval types for which they are especially sensitive regardless of the absolute pitch of their components. However, these differences disappeared when they were placed in a more naturalistic context with shifted song notes instead of synthetic tones.

It is possible, because the natural discrimination was overall learned faster than the synthetic discrimination, that the task with natural stimuli was not difficult enough to show a reliable effect of group, and that a larger sample of stimuli could have promoted more diversity in responses among groups. Whether or not an effect would be seen with more stimuli, it is unclear why the natural discrimination was overall easier for the birds compared to the synthetic discrimination. One possibility is that the use of pitch intervals similar to *fee-bee* song may have promoted closer attention to the stimulus set and a similar level of attention was attained by using natural stimuli. Alternatively, the birds may have used additional cues within the shifted song stimuli (i.e., rote memorization of spectral features of each stimulus) to solve the discrimination and these additional features may have masked any differences in discrimination among groups.

However, even in the discrimination using natural stimuli, the birds appeared to be using pitch to some degree (either relative pitch or the absolute pitch of the first note), because many of them were then able to transfer what they learned in the discrimination using natural stimuli to the discrimination using synthetic stimuli as shown by above chance responding during the first block of synthetic training.

Indeed, for the synthetic discrimination, pitch was the only cue the birds could be using to solve the task. It is important to note that the discriminations could have been solved by memorizing the absolute pitch of the first note (the second note was maintained across pitch intervals). If absolute pitch was being used primarily, however, it is more difficult to find an explanation of why the species-typical pitch interval appeared to be easier to discriminate than the other pitch intervals. One possibility is that the relevance of the absolute pitches of the notes varied among pitch interval types despite the fact that chickadees sing their *fee-bee* song at variable absolute pitches. Importantly, the possible values for the second note in our stimuli, which were the same for all pitch intervals, were chosen based on being within the natural absolute pitch range of natural *fee-bee* song. Because of this, the frequency of the first note is more likely to be within the natural range of *fee-bee* song for the species-typical song pitch interval. However, the previous work by Njegovan and Weisman (1997), upon which our design was based, suggests that the birds are using, at least partially, relative pitch to solve this task. In addition, pitch intervals closer to the species-typical pitch

interval should, with the use of absolute pitch, be more difficult to discriminate because they fall more frequently into the natural absolute pitch range while we found the opposite to be the case. Thus we suggest the birds were likely using relative pitch to solve the task.

Instead of the birds having more difficulty discriminating pitch intervals within the natural absolute pitch range, discrimination was facilitated for pitch intervals that were closer to the species-typical *fee-bee* song. For the discrimination using synthetic stimuli, birds in the 1.19 Group learned to discriminate the species-typical pitch interval from the S+ pitch interval with higher accuracy compared to a species-atypical S- pitch interval. This could potentially be due to the song pitch interval being a category for the chickadees that had already been learned prior to our operant training procedure in the wild. At the same time, the 1.25 Group learned their larger S- pitch interval (the one least like the species-typical pitch interval in our study) with higher accuracy than the smaller pitch interval. This may be due to the larger pitch interval being so large that it was out of the range of normal *fee-bee* songs and was thus discriminated by being an anomaly. However, bioacoustic measurements of the *fee-bee* song indicate that all of our species-atypical pitch intervals would be outliers compared to natural song (over three standard deviations from the mean; Weisman et al. 1990). Our pitch intervals were thus far more different from one another in frequency than the ones normally encountered by the birds in nature, which is interesting because of the difficulty the birds had at solving our

discrimination and other similar discriminations (Weisman et al. 1994; Njegovan & Weisman 1997), despite the fact that the pitch interval is a marker of male quality (Christie et al. 2004). It may be that, counter-intuitively, if we had used intervals that were more similar and all within the range of natural *fee-bee* song, the birds may have had an easier time solving the discrimination since the current results suggest that the species-typical interval was discriminated with the least difficulty. It may also be that the birds have difficulty using relative pitch cues to solve a discrimination unless specifically primed with song cues. In other words, if they hear either the species-typical relative pitch interval (as in both the synthetic and natural stimuli) or if they hear specific spectral features of song (as in only the natural stimuli) their discrimination may improve. This would explain why there were differences among groups for the synthetic but not the natural discrimination since all natural discrimination stimuli contained song cues such as timbre and subtle pitch changes (e.g., the *fee* glissando, see Weisman et al. 1990).

We proposed that the sexes would perform similarly both in terms of how many trials they would need to complete the task and in terms of accuracy because of past similar operant performance between sexes for black-capped chickadees (e.g., Guillette et al. 2009; Bloomfield et al. 2008; Charrier et al. 2005). Instead, we found that the females reached a significantly higher level of accuracy when discriminating the pitch intervals. However, males were more sensitive to the larger S- pitch intervals at the beginning of the task. In the wild, males need to respond quickly to intruding males, and subordinate males typically

produce pitch intervals that are smaller, rather than larger, than the normal pitch interval when shifting their songs (Christie et al. 2004). It may be that males may have perceived the larger pitch intervals as being "super dominant" (e.g., like the preference for "supernormal" sized eggs in herring gulls; Baerends et al. 1982). However, because male black-capped chickadees tend to become more dominant with age (Smith 1991) and it has been shown in other species that consistency in song production and age are related (e.g., de Kort et al. 2009) it may be that birds are looking only for consistency over multiple renditions of song rather than absolute pitch interval size. Females, on the other hand, have a longer period of time within which to choose a mate, and choosing a mate is a long-term commitment (outside of extra pair copulations) as black-capped chickadees often mate for life (Smith 1991). Chickadee flocks that are formed in the fall typically are composed of bonded male/female pairs that often, but not always, are the same pairs that breed together in the spring. However over the winter, females may move to a higher quality male if they have the chance (Smith 1991). Unlike many other songbirds, the chickadee has periods of high song production in the winter as well as spring (Avey et al. 2008). For these reasons, it is possible that females would lean more towards accuracy than speed of assessment than a male that needs to respond to an incoming threat immediately. This fits well with our results that females had a higher level of accuracy, but only after a prolonged period of evaluation.

Researchers have shown that animals, given several stimuli from different modalities that can be used to predict an outcome, tend to learn about different predictor stimuli depending on whether the outcome is appetitive or aversive (e.g., Foree & LoLordo 1973; Kelley 1986; Schindler & Weiss 1982) and associations are more likely to form to more species-typical stimulus modalities (i.e., stimulus modalities that are more likely to be associated in nature). For example, Garcia & Koelling (1966) trained rats with a flavoured water spout that turned on an audiovisual stimulus when licked. If the rats were subsequently given an electric shock, they associated the audiovisual stimulus to the shock. In contrast, if the rats were made ill with a toxin or x-ray, they associated the flavoured water with the illness. Our results suggest that biological relevance can affect the way animals group stimuli within a single modality (in our case, auditory pitch interval stimuli). While it is known that in some cases animals will come to generalize responding normally reserved for species-typical stimuli to other similar stimuli but not dissimilar stimuli (e.g., Cusato & Domjan 1998), our results here suggest that this type of generalization may be more difficult to achieve than generalization among species-atypical stimuli. In essence, it may be that generalization is more likely to occur with a set of stimuli that, despite its members being objectively equally dissimilar to one another as members of another set of stimuli, are further from a species-typical category.

Table 2-1. Frequencies (Hz) of the pitch interval note pairs in the array. Each row represents five interval types between a single Note 2 frequency and five different Note 1 frequencies for each pitch interval ratio. The pitch intervals that were used to train each group are represented by the black bars and group name (1.13 Group, 1.19 Group, and 1.25 Group). The grey rows are the frequency values that were used for training with the birds.

| | | | | 1.25 Group | | |
| | | | 1.13 Group | | | |
Note 2	Ratio	1.07	1.13	1.19	1.25	1.31
2519		2695	2846	2998	3149	3300
2593		2774	2930	3085	3241	3397
2669		2856	3016	3176	3336	3496
2747		2939	3104	3269	3434	3599
2827		3025	3195	3365	3534	3704
2910		3114	3289	3463	3638	3813
2996		3205	3385	3565	3745	3924
3083		3299	3484	3669	3854	4039
3174		3396	3586	3777	3967	4158
3267		3495	3691	3887	4083	4279
3362		3598	3800	4001	4203	4405
3461		3703	3911	4119	4326	4534
3562		3812	4026	4239	4453	4667
3667		3923	4143	4363	4583	4803
3774		4038	4265	4491	4718	4944
3885		4157	4390	4623	4856	5089
3999		4279	4518	4758	4998	5238
4116		4404	4651	4898	5145	5392
			1.19 Group			

Table 2-2. The proportion of individuals of each group and sex that responded above chance in the first block of their first and second discrimination. The top panel shows birds that learned the discrimination using synthetic stimuli first. The bottom panel shows birds that learned the discrimination using natural stimuli first.

Synthetic First

Group	Sex	Significantly Above Chance Synthetic	Natural
1.13 Group	Male	0/3	1/3
1.13 Group	Female	0/3	1/3
1.19 Group	Male	0/3	2/3
1.19 Group	Female	0/3	3/3
1.25 Group	Male	0/3	2/3
1.25 Group	Female	0/3	2/3

Natural First

Group	Sex	Significantly Above Chance Natural	Synthetic
1.13 Group	Male	0/3	1/3
1.13 Group	Female	1/3	1/3
1.19 Group	Male	0/3	2/3
1.19 Group	Female	0/3	1/3
1.25 Group	Male	0/3	1/3
1.25 Group	Female	0/3	2/3
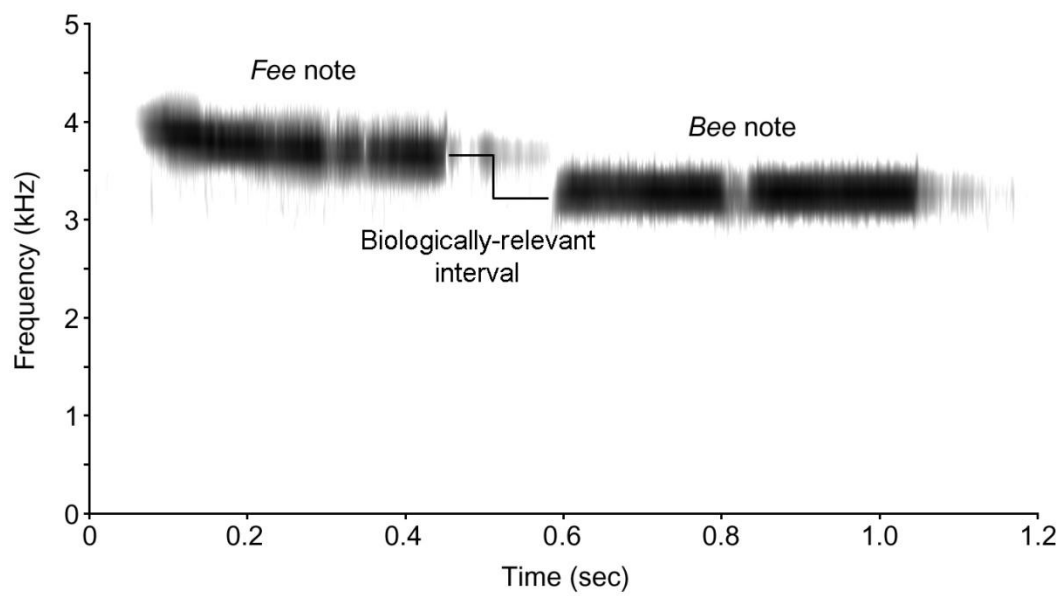
Figure 2-1.
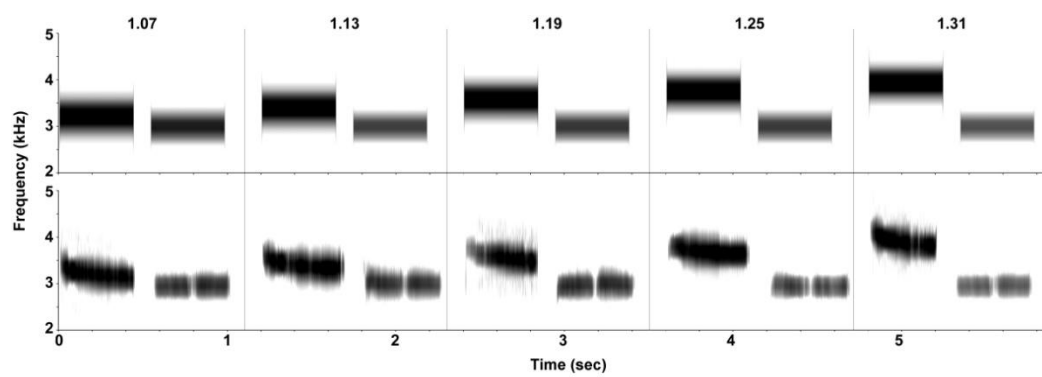
Figure 2-2.
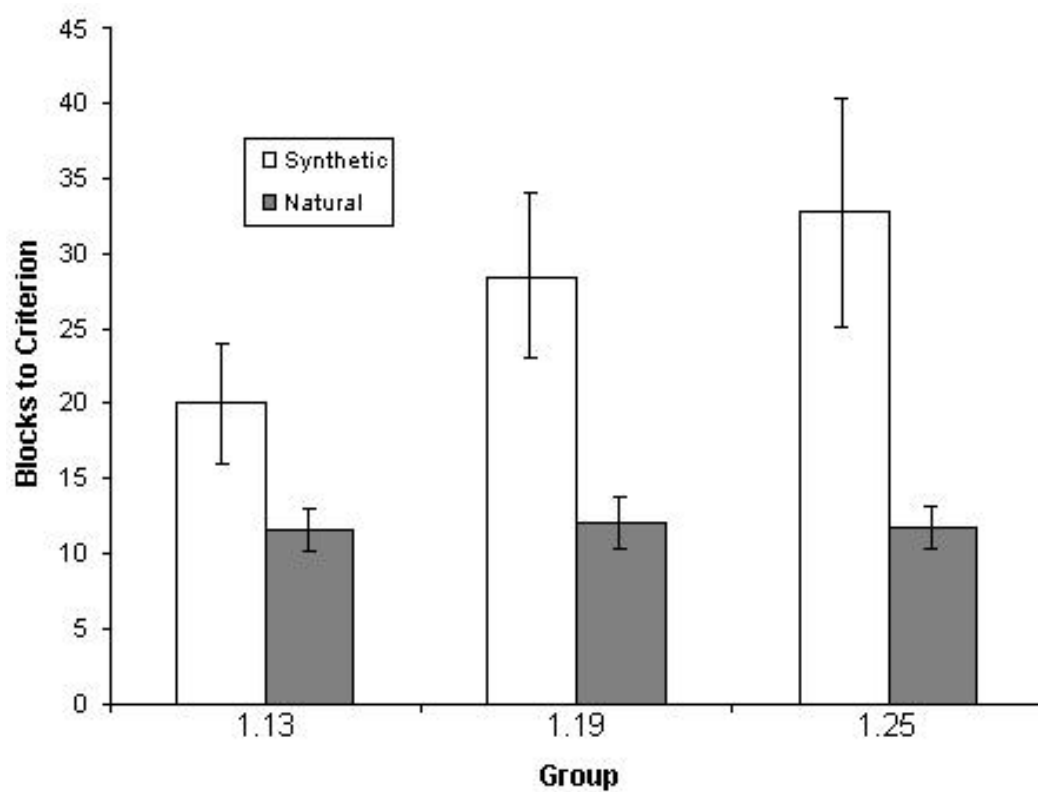
Figure 2-3.
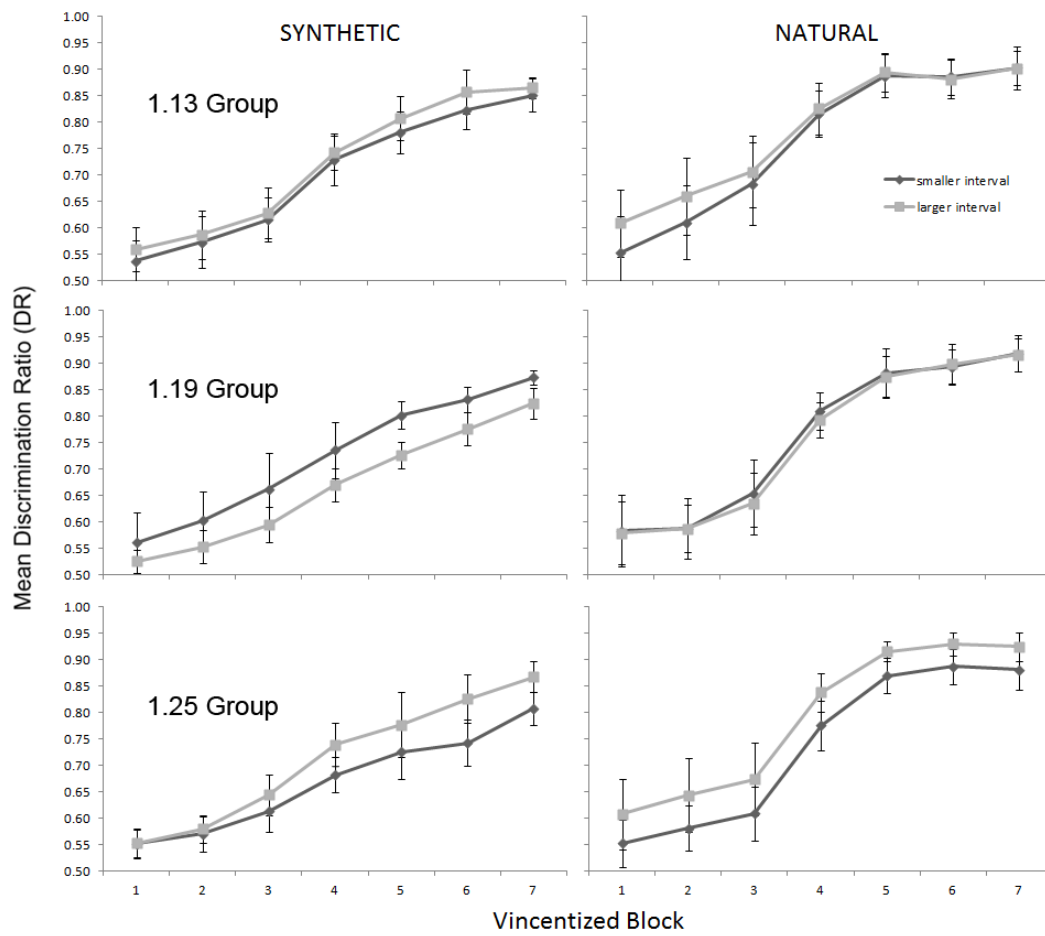
Figure 2-4.
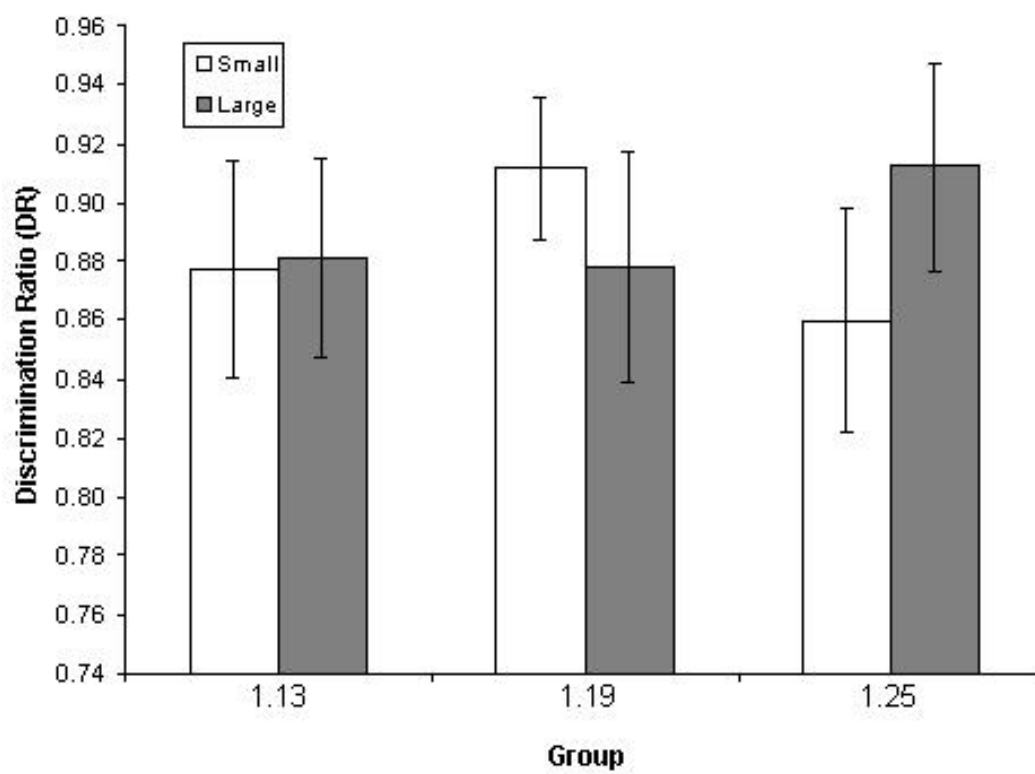
Figure 2-5.
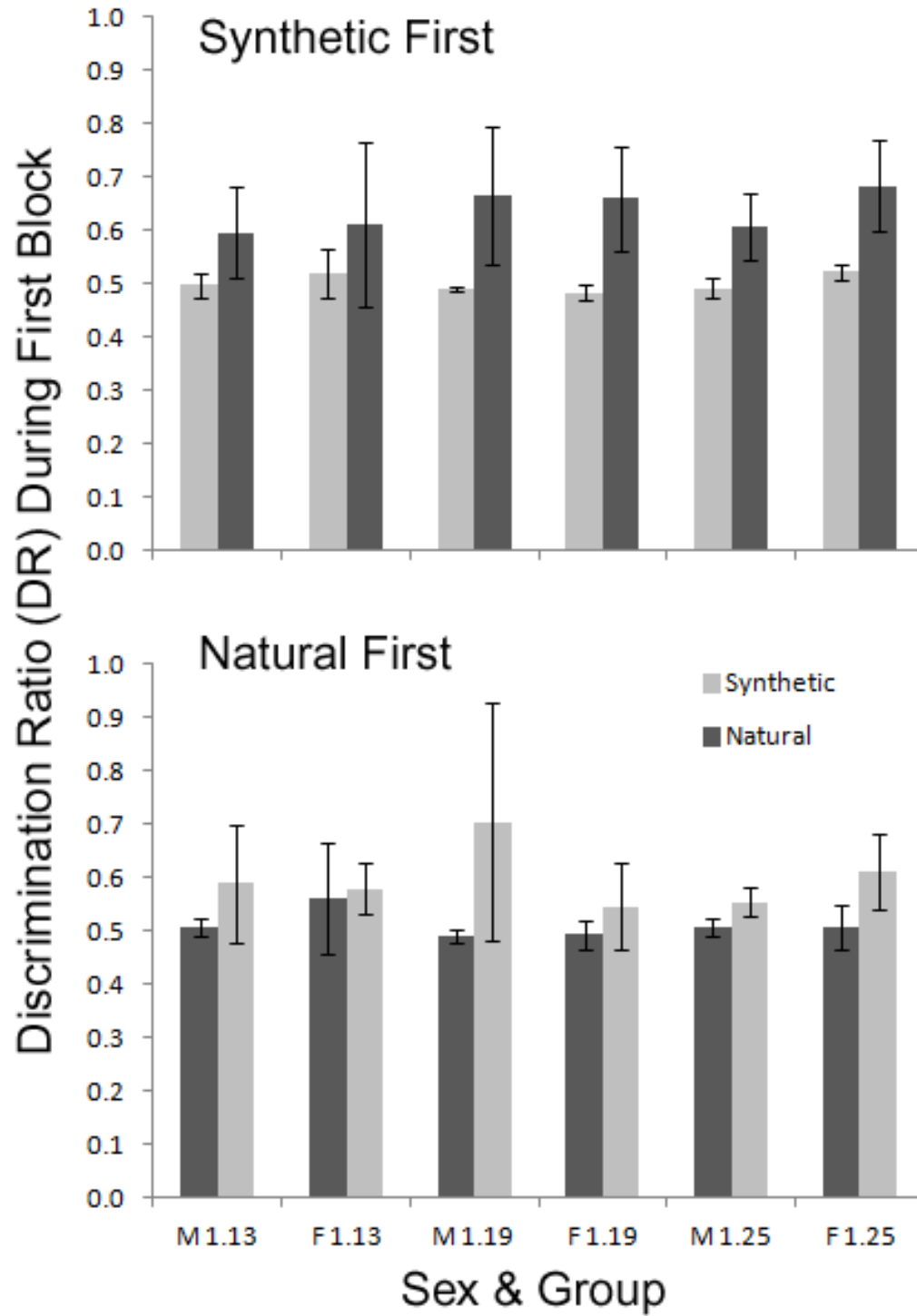
Figure 2-6.

References

Avey, M. T., Quince, A. F., & Sturdy, C. B. (2008). Seasonal and diurnal patterns of black-capped chickadee (*Poecile atricapillus*) vocal production. *Behavioural Processes, 77*, 149-155.

Bachorowski, J-A., & Owren, M. J. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *Journal of the Acoustical Society of America, 106*, 1054-1063.

Baerends, G. P., Drent, R. H., Hogan-Warburg, A. J., Blokzijl, G. J., van der Molen, A. G., de Vos, O. C., Baarda, H. J., Mudde, M. J., Kruijt, J. P., & Vodegel, R. (1982). The herring gull and its egg: Part II. The responsiveness to egg-features. *Behaviour, 82*, I-XIII+1-416.

Baharloo, S., Johnston, P. A., Service, S. K., Gitschier, J., & Freimer, N. B. (1998). Absolute Pitch: An Approach for Identification of Genetic and Nongenetic Components. *American Journal of Human Genetics, 64*, 224-231.

Bloomfield, L. L., Farrell, T. M., & Sturdy, C.B. (2008). All "chick-a-dee" calls are not created equally. Part II. Mechanisms for discrimination by sympatric and allopatric chickadees. *Behavioural Processes, 77*, 87-99.

Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychology, 16*, 681-684.

Charrier, I., Lee, T. T-Y., Bloomfield, L. L., & Sturdy, C. B. (2005). Acoustic

    mechanisms of note-type perception in black-capped chickadee (*Poecile*

    atricapillus) calls. *Journal of Comparative Psychology, 119*, 371-380.

Christie, P. J., Mennill, D. J., Ratcliffe, L. M., (2003). Chickadee song structure is

    individually distinctive over long broadcast distances. *Behaviour,*

    141,101-124.

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.

Cusato, B., & Domjan, M. (1998). Special efficacy of sexual conditioned stimuli

    that include species typical cues: tests with a conditioned stimuli

    preexposure design. *Learning and Motivation, 29*, 152-167.

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.

de Kort, S. R., Eldermire, E. R. B., Valderrama, S., Botero, C. A., Vehrencamp, S.

    L., (2009). Trill consistency is an age-related assessment signal in banded

    wrens. *Proceeding of the Royal Society B: Biological Sciences, 276*, 2315-

    2321.

Ficken, M. S., Ficken, R. W., & Witkin, S. R., (1978). Vocal repertoire of the

    black-capped chickadee. *The Auk, 95*, 34-48.

Foree, D. D., & LoLordo, V. M., (1973). Attention in the pigeon: differential effects of food-getting versus shock-avoidance procedures. *Journal of Comparative and Physiological Psychology, 85*, 551-558.

Friedrich, A., Zentall, T., & Weisman, R. G., (2007). Absolute pitch: Frequency-range discriminations in pigeons (*Columba livia*) - Comparisons with zebra finches (*Taeniopygia guttata*) and humans (*Homo sapiens*). *Journal of Comparative Psychology, 121*, 95-105.

Garcia, J., Koelling, R. A., (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science, 4*, 123-124.

Gibson, B., & Kamil, A., (2009). The synthetic approach to the study of spatial memory: have we properly addressed Tinbergen's "four questions"? *Behavioural Processes, 80*, 278-287.

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.

Guillette, L. M., Reddon, A. R., Hurd, P. L., & Sturdy, C., B., (2009). Exploration of a novel space is associated with individual differences in learning speed in black-capped chickadees, *Poecile atricapillus*. *Behavioural Processes, 82*, 265-270.

Hoeschele, M., Moscicki, M. K., Otter, K. A., van Oort, H., Fort, K. T., Farrell, T.

M., Lee, H., Robson, S. W. J., Sturdy, C. B., (2010). Dominance Signalled
in an Acoustic Ornament. *Animal Behaviour, 79*, 657-664.

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.

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.

Hulse, S. H., Cynx, J., (1985). Relative pitch perception is constrained by absolute
pitch in songbirds (*Mimus*, *Molothrus*, and *Sturnus*). Journal of
Comparative Psychology 99, 176-196.

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

Kelley, M. J., (1986). Selective attention and stimulus-reinforcer interactions in
the pigeon. *Quarterly Journal of Experimental Psychology, 38B*, 97-110.

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

Lohr, B. (2008). Pitch-related cues in the songs of sympatric mountain and black-
capped chickadees. *Behavioural Processes, 77*, 156-165.

Mennill, D. J., Ratcliffe, L. M., & Boag, P. T. (2002). Female eavesdropping on

male song contests in songbirds. *Science, 296*, 873.

Moore, B. C. J., (1989). *An introduction to the psychology of hearing, 3rd edn*. London, UK: Academic Press.

Njegovan, M., & Weisman, R. G., (1997). Pitch discrimination in field- and isolation-reared black-capped chickadees (*Parus atricapillus*). *Journal of Comparative Psychology, 111*, 294-301.

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.

Otter, K., & Ratcliffe, L., (2011). Changes in singing behavior of male black-capped chickadees (*Parus atricapillus*) following mate removal. *Behavioral Ecology and Sociobiology, 33*(6), 409-414.

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

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

Ryan, M. J., & Keddy-Hector, A. (1992). Directional patterns of female mate choice and the role of sensory biases. *American Naturalist, 139*, S4-S35.

Schindler, C. W., Weiss, S., J., (1982). The influence of positive and negative reinforcement on selective attention in the rat. *Learning and Motivation, 13*, 304-323.

Smith, S. M., (1991). *The Black-Capped Chickadee: Behavioral Ecology and Natural History.* Ithaca, New York: Cornell University Press.

Sturdy, C. B., Bloomfield, L. L., Farrell, T. M., Avey, M. T., & Weisman, R. G., (2007). Category perception as a natural cognitive activity in songbirds. *Comparative Cognition and Behavior Reviews, 2*, 93-110.

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

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

Weisman, R. G., Njegovan, M. G., Williams, M., Cohen, J., & Sturdy, C. B. (2006). The comparative psychology of absolute pitch. In E. A. Wasserman & T. Zentall (Eds.), *Comparative Cognition: Experimental Explorations of Animal Intelligence* (pp. 71-86) Oxford: Oxford University Press.

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. *Behavioural Processes, 66*, 289-307.

Weisman, R. G., Njegovan, M., Sturdy, C. B., Phillmore, L. S., Coyle, J., & Mewhort, D. J. K., (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., & Ito, S., (1994). Frequency ratio discrimination

by zebra finches (*Taeniopygia guttata*) and humans (*Homo sapiens*).

Journal of Comparative Psychology, 108, 363-372.

Weisman, R. G., Ratcliffe, L. M., Johnsrude, I., & Hurly, T. A. (1990). Absolute

and relative pitch production in the song of the black-capped chickadee.

Condor, 92, 118-124.

Zentall, T. R., Galizio, M., & Critchfied, T. S., (2002). Categorization, concept

learning, and behavior analysis: an introduction. *Journal of the*

Experimental Analysis of Behaviour, 78, 237-248.

Chapter 3 **Black-capped chickadee (*Poecile atricapillus*) and human (*Homo sapiens*) chord discrimination**[2]

Introduction

Music is a universal and potentially unique feature of our species. All cultures have some form of music, and musical systems that developed in isolation from one another share many common elements such as logarithmic pitch scales and octave equivalence (notes doubled in frequency are perceived as the same; Dowling & Harwood, 1986). Music is a large part of human culture and archaeological research suggests it has been for at least 30,000 to 50,000 years (Kunej & Turk, 2000).

As a result, interest has grown in the developing field of biomusicology, an evolutionary approach to the scientific study of music (Fitch, 2005; Hauser & McDermott, 2003; Justus & Hutsler, 2005; Trehub, 2003). Research in this field attempts to unravel the evolutionary origins of music through comparative research between human and non-human animals. This approach increases our understanding of the origins of the human musical faculty, what limitations and requirements there might be on perceiving/producing music, and what fundamental behavioural or physiological building blocks are present in non-human species that use auditory channels to listen and communicate. In many cases, songbirds have been chosen as subjects for comparative studies of music perception in animals because of the importance of acoustic communication to them (e.g. Hulse & Cynx, 1985; Watanabe, Uozumi & Tanaka, 2005).

A significant body of research has explored songbird pitch perception. Pitch (perception of the fundamental frequency) is a critical component of music

along with timbre (tone quality as perceived by the physical shape of soundwaves), amplitude (loudness) and temporal intervals. Pitch can be classified in one of two ways: relative pitch and absolute pitch. The former refers to the ability to perceive the relation between two or more notes. The latter refers to the ability to determine note pitch in the absence of any external referent. In many studies, songbirds appear to rely more on absolute rather than relative pitch to solve tasks designed with simple relative pitch relations (Cynx, Hulse & Polyzois, 1986; Hulse, Cynx & Humpal, 1984). A study that pitted absolute against relative pitch revealed that, although European starlings (*Sturnus vulgaris*) appeared to solve the task using absolute pitch as their primary strategy, the birds could use relative pitch to some extent. The switch to relying on relative pitch was especially evident when the test stimuli were from a range of frequencies that overlapped training stimuli (Hulse & Cynx, 1985). Subsequent studies have discovered that songbirds have highly developed absolute pitch perception abilities and perform much better than humans, pigeons and rats on tasks that exclusively require absolute pitch (Weisman et al., 1998; Weisman, Njegovan, Williams, Cohen & Sturdy, 2004; Friedrich, Zentall & Weisman, 2007; for review see Weisman, Williams, Cohen, Njegovan & Sturdy, 2006). In fact, fine-grained absolute pitch perception (typically measured as note naming) in humans is very rare and has been estimated to occur in less than one in ten thousand individuals (Bachem, 1955). Thus, the mechanisms underlying songbirds' perception of pitch intervals, which require the use of relative pitch, may be different from that of

60

humans.

Besides looking at sequentially presented pitch relations, the notes forming these intervals can also be presented simultaneously. In this case, they form a harmonic complex that is the basis for chord perception in music. In particular, triads (three-note chords) play an important role in musical structure and organization. Birds have been found to be sensitive to such harmonic stimuli. Both java sparrows (*Padda oryzivorai*; Watanabe, et al., 2005) and starlings (Hulse, Bernard & Braaten, 1995) that have been successfully trained to discriminate among triadic chords showed some transfer to other triadic chord stimuli presented at different absolute pitches where the relations among the notes in the triadic chord remained constant. In other words, they solved a task with harmonic chords that required the use of relative pitch.

Recently, members of our group tested a non-songbird species, the pigeon (*Columba livia*), using a discrimination task involving more musically structured chords than previous studies of chord discrimination by birds (Brooks & Cook, 2010). The pigeons were trained to discriminate a major triadic chord (1st, 3rd, and 5th note from a major scale) from four other chords that were similar, except that either the 3rd or the 5th note of the chord was raised or lowered by one semitone. This generated four musical chords: the minor chord (minor), with a lowered 3rd note, the suspended 4th chord (sus4) with a raised 3rd note, the flat 5th chord (flat5), with a lowered 5th note, and the augmented (aug) chord with a raised 5th note. The major, minor, suspended 4th, and augmented chords are

61

frequently used in various musical contexts. The flat5 chord, although not normally used in Western music without also flattening the 3rd to form a diminished chord, was tested to maintain consistency of design within the experiment (see Figure 3-1). Brooks and Cook (2010) found that pigeons and humans, when trained (or asked) to rate the similarity of these manipulated chords to the major chord, displayed similar patterns of responding. For instance, chords with manipulations of the 3rd note were rated as more similar to the major chord than chords having manipulations of the 5th note.

Here we report data from a chord discrimination study using the same stimuli and contingencies that were used by Brooks and Cook (2010) with both black-capped chickadees (*Poecile atricapillus*) and humans (*Homo sapiens*) as subjects. Black-capped chickadees are North American songbirds that, in common with other true (oscine) songbirds, have a complex, learned vocal system subserved by an interconnected set of definable brain regions (Ficken, Ficken & Witkin, 1978; Shackleton & Ratcliffe, 1993; Hughes, Nowicki & Lohr, 1998; Mooney, 2009). Thus songbirds, and their communication system, stand in sharp contrast to pigeons, whose vocalizations are unlearned and who lack a comparable neural architecture. Songbird vocal learning has often been compared to human vocal learning (e.g., Bolhuis, Okanoya & Scharff, 2010; Doupe & Kuhl, 1999). More recently it has been suggested that vocal learners may be an ideal group for studying the evolution of music (e.g., Fitch 2005; Patel, 2006) and evidence exists suggesting that vocal learners may share some essential components for music

processing that are relatively rare, such as entrainment to a musical beat (Patel, Iversen, Bregman & Schulz, 2009; Schachner, Brady, Pepperberg & Hauser, 2009).

The purpose of Experiment 1 was to use the same stimuli from Brooks and Cook (2010) to compare a songbird's chord discrimination performance with that previously reported in pigeons (Brooks & Cook, 2010). This provides a compelling comparison of chord discrimination and harmonic perception across very different orders of birds. In Experiment 2, we tested humans in an operant discrimination experiment comparable with those used to test pigeons and chickadees (i.e., with no verbal instructions) which permitted a more direct comparison between human and avian performance than the rating task used by Brooks and Cook (2010). Our goal here was to conduct a systematic comparison of chord perception across species to determine how humans and avian species perceive musical chords. We had three main expectations: (1) that chickadees should be able to solve this chord task and do so more accurately than pigeons based on previous comparative pitch processing findings (Lee, Charrier, Bloomfield, Weisman & Sturdy, 2006; Friedrich, et al., 2007), (2) that human performance should be about equivalent to that of chickadees and superior to that of pigeons (Njegovan & Weisman, 1997; Weisman, Njegovan & Ito, 1994) and (3) that chickadees, pigeons and humans have the same pattern of responding to all chords (responding that is primarily based on spectral properties of the chords rather than cultural influences e.g., modeled by N. D. Cook & Fujisawa, 2006).

63

General Method

Stimuli

Chords were generated using a French Horn synthesized MIDI timbre (Sonar 4, Cakewalk, Boston, USA). All stimuli were triads composed of notes selected from the fourth octave of the diatonic scale. For C root training, the S+ (reinforced or go) stimulus was the C major triad (C E G) while S- stimuli (non-reinforced or no-go) consisted of semitone deviations of either the 3rd or the 5th note of the C major chord in the following four variations: minor (C Eb G), sus4 (C F G), flat5 (C E Gb), and aug (C E G#). For D root training, the S+ stimulus was the D major chord (D F# A) while S- stimuli consisted of semitone deviations of either the 3rd or the 5th note of the D major chord in the following four variations: minor (D F A), sus4 (D G A), flat5 (D F# Ab), and aug (D F# A#). These chords were identical to those tested earlier with pigeons (Brooks & Cook, 2010) except that they were shortened and the edited portion tapered to 2-s total duration using SIGNAL version 5.10.24 (Engineering Design, Berkeley CA). This change made it possible to use the chords under our standard protocol (that requires chickadees listen to the entire stimulus before making a response) while maintaining the native attack and steady state portion of the stimuli.

General Procedure

Both chickadees (Experiment 1) and humans (Experiment 2) were trained using a Go/No-Go paradigm first to discriminate the C major chord from C minor, sus4, flat5 and aug chords. The C major chord was presented on 50% of the trials and was the rewarded stimulus. The other stimuli were presented approximately equally and together made up the other 50% of the trials which were unrewarded stimuli. The same discrimination was then conducted using the D root version of the same chords (discriminating the D major chord from D minor, sus4, flat5 and aug chords). The details of how this procedure was implemented in each species are described separately for each experiment below.

Response Measures

To determine whether the chickadees and humans had successfully learned to discriminate among the chords, we calculated a discrimination ratio (DR) between the S+ chord and each S- chord. To calculate the DRs, we divided the percent response for the S+ chord by the sum of the percent response for the S+ chord and the S- chord of interest. A DR of 0.5 indicates equal responding to both S+ and S-, while a higher DR means more responding to S+ and a lower DR means more responding to S-. Percent response for chickadees was calculated by first excluding any interrupted trials from the total number of trials (see Experiment 1 procedure).

Experiment 1: Black-capped Chickadees

Method

Animals

Six black-capped chickadees (three male and three female as identified by DNA analysis; Griffiths, Double, Orr & Dawson, 1998) were tested between April and August 2009. These birds were captured in Edmonton, Alberta, Canada (53∘06' N, 113∘04' W), and at the Barrier Lake Field Station in Kananaskis Country, Alberta, Canada (51∘02' N, 115∘03' W) between December 2007 and March 2008. All birds were determined at the time of capture to be at least one year of age by examining the color and shape of their outer tail retrices (Pyle, 1997). Prior to the experiment, all birds were housed individually in Jupiter Parakeet cages (30 × 40 × 40cm; Rolf C. Hagen, Inc., Montreal, Canada) in colony rooms. The birds had visual and auditory, but not physical, contact with one another.

Throughout the experiment, birds were housed within individual operant chambers (see apparatus), maintained on the natural light cycle, and had ad libitum access to water, grit and cuttle bone. During experimentation, however, food was only available as a reward for correct responding in the operant discrimination task. Each chickadee had prior experience with auditory

discriminations involving natural and synthetic stimuli (natural or synthetic fee-bee songs or chick-a-dee call note stimuli), but were naïve to the current stimulus set. For details about diet before and during the experiment, please refer to Guillette, Reddon, Hoeschele & Sturdy (2011).

Apparatus

During the experiment, the birds lived in modified colony room cages (30 × 40 × 40cm). Each cage was placed inside a ventilated sound-attenuating chamber and contained three perches and dispensers for water and grit. The chambers were illuminated by a 9-W, full spectrum fluorescent bulb. An opening on the side of the cage (11 × 16cm) gave each bird access to a motor-driven feeder (Njegovan, Hilhorst, Ferguson & Weisman, 1994). Both the feeder and the perch closest to the feeder (request perch) had infrared cells to monitor the position of the bird. A computer and single-board computer (Payla & Walter, 2001) scheduled trials and recorded responses to stimuli. Stimuli were played from CD through an amplifier and a full-range speaker located beside the feeder. For more equipment details, please refer to Sturdy & Weisman (2006). Stimuli were presented at amplitudes of roughly 76-82 dB as measured by a Radio Shack Sound Level Meter (A weighting, slow response; Radio Shack, Fort Worth, USA) from the position of each bird's head when located at the request perch.

Procedure

Nondifferential training

After a bird had learned to use the request perch and feeder, nondifferential training began. Trials continued indefinitely throughout the day cycle. Landing on the request perch, breaking the infrared beam and remaining for 1-s on average (range 0.9-1.1-s) initiated a trial. During each trial, one of the ten chords (both C and D root) was randomly chosen without replacement from the pool and played through the speaker. Once all chords had been played in a random order, they were again randomized and the procedure repeated. If the bird left the perch during the 2-s playback of the chord (i.e., an interrupted trial), the chamber lights turned off for a 30-s inter-trial interval (ITI), during which no new trial could begin. This ensured that the birds listened to the entire stimulus before making a response. If the bird flew or hopped to the feeder within 1-s after the completion of the stimulus, the bird received 1-s access to food. This was followed by a 30-s ITI with the chamber lights remaining on. If the bird left the perch upon completion of the stimulus, but did not enter the feeder, the trial ended after 1-s and the bird could then initiate a new trial. To prevent each bird from sitting continuously on the request perch and initiating trial after trial, a new trial could only be activated by either leaving and returning to the request perch or waiting for a 60-s ITI. This increased the probability that a bird would make a response on a given trial. Data from this phase were evaluated in 500-trial blocks.

Nondifferential training continued until birds were going to the feeder on a high proportion of trials (at least 60% of the trials) for at least six blocks and birds were responding with ≤ 3% difference to future S+ and S- chords for at least four blocks. This ensured there were no initial response biases to the stimuli.

C root discrimination training

C root discrimination training was identical to nondifferential training, but with several critical exceptions. During this phase, birds were only presented the C root chords. In addition, birds were rewarded only for responding, by entering the food hopper, following presentation of the S+ C major chord. If they entered the feeder after an S- stimulus (all other C root chords), the chamber lights were turned off during a 30-s ITI. Finally, the reinforced (S+) C major chord made up 50% of the trials, while the four other nonreinforced (S-) C chords (i.e., alterations of the 3rd and 5ths) made up equal proportions of the remaining trials (12.5% each). Birds were trained with C root discrimination training until they each completed six 500-trial blocks (the last two occurring consecutively) in which Discrimination Ratio (DR) for each chord was greater than or equal to 0.80 (see response measures, above, for DR calculations). This took a variable number of trials for each bird, but ensured all birds were at the same level of expertise. The number of blocks taken to reach this criterion was subsequently used in our analyses (see statistical analyses section).

69

D root discrimination training

D root discrimination training commenced for each bird after criterion at C root training had been reached. This phase was identical to the C root discrimination training, except that the D chords were used as discriminative stimuli instead of the C chords. This was to test whether the birds could generalize the discrimination they learned about the C root chords to chords centered around a novel root note. The training criterion was the same as for C root discrimination.

Statistical Analyses

All analyses were conducted using Statistica version 8.0 (Statsoft Inc.). An alpha level of 0.05 or less was used to judge significance in statistical tests.

Acquisition

There was considerable variation among the birds in the number of blocks required to reach criterion for all chords (between 11-47 blocks depending on bird and discrimination training phase, either C or D root). For this reason, acquisition curves were standardized to Vincentized curves with 10 blocks of data so that we could compare the rate of acquisition across birds. Vincentizing was calculated by dividing the total number of actual blocks a bird did (e.g. 32) by 10 Vincentized blocks into whole numbers (in this example, 3 remainder 2). Vincentized blocks were then computed by grouping the total number of blocks a bird did into groups

the size of the quotient (in this example, 3) and placing the any remaining blocks

evenly into the middle two blocks (5th and 6th block, in this example, both the

5th and 6th block would have $3 + 1 = 4$ actual blocks grouped in) and then

averaging the DR across grouped blocks (see Kling & Riggs, 1971). We

conducted an analysis of variance (ANOVA) comparing the DR for Sex $\times$

Vincentized blocks $\times$ Chord type to see whether there were any differences in

learning rates (e.g., if some chords are learned faster closer to the beginning of the

learning curve). Additionally, to make our current work comparable to the

previous pigeon work (Brooks & Cook, 2010) we conducted a similar analysis as

that just described, but collapsing; (1) minor and sus4 chords to compare

acquisition with manipulations of the 3rd, and (2) the aug and flat5 chords to

compare acquisitions with manipulation of the 5th.

Blocks to Criterion

Because we tested all birds until they reached criterion at both stages for

both C and D root training, we were able to evaluate the number of blocks to

reach criterion for each individual chord discrimination (e.g., the minor from the

major with a DR ≥ 0.8 for six blocks with the last two being consecutive). This

was evaluated using a Sex $\times$ Chord type ANOVA.

Savings. We also looked at whether there were any savings from C root training,

to D root training. To complete this analysis, we calculated one-tailed z scores for

each individual bird using binomial tests for dichotomous data to determine

whether the proportion of responses that were directed to S+ stimuli were significantly above chance (1/2 chance because 1/2 of the trials were S+) during the first block of each of the two discriminations (training and transfer).

Results

All birds reached criterion for each of the four comparison chord types in both C root and D root training. Figure 3-2 shows the average percent response across the 10 Vincentized blocks for each chord type for both C root (upper panel) and D root (lower panel) training. Figure 3-2 also shows that there is variation in the rate at which different S- chords are learned across acquisition. A more detailed depiction of the differences among S- chords is shown in Figure 3-3 where the number of blocks to reach criterion for each chord type and training phase are plotted for each individual bird. The pattern of results in Figure 3-3 suggests that the minor chord and the flat5 chords took the longest to learn during C root training, whereas the minor and sus4 chords took the longest to learn during D root training. In general, this pattern of results was supported by statistical analyses described below.

C root training

Acquisition. We found a main effect of Vincentized block in the ANOVA looking at the DRs of birds by Sex × Vincentized blocks × Chord type, both in the

separated (minor VS sus4 VS flat5 VS aug) and collapsed (3rd manipulations VS 5th manipulations) versions of the analysis ($F(9, 36) = 72.78$, $p < 0.001$, partial $\eta^2 = 0.948$) where DR increased as Vincentized blocks increased. In the separated analysis, we also found a main effect of chord ($F(3, 12) = 10.90$, $p < 0.001$, partial $\eta^2 = 0.732$) that was not found in the collapsed analysis. Tukey post-hoc tests show that the minor chord ($M = 0.70$) had a significantly lower overall DR than the aug chord ($M = 0.73$; $p = 0.015$) and the sus4 chord ($M = 0.75$; $p = 0.001$). In addition, the sus4 chord had a significantly higher DR than the flat5 chord ($M = 0.71$; $p = 0.010$).

Blocks to Criterion. Figure 3-3 shows the blocks to criterion for each bird for each chord. All birds learned the minor chord last, suggesting it took the longest to learn. Four birds learned sus4 first, one learned the aug chord first, and one learned both the aug and flat5 chords first. The analyses on blocks to criteria also showed results similar to the acquisition analysis. There was a main effect of chord ($F(3, 12) = 7.38$, $p = 0.005$, partial $\eta^2 = 0.648$). Tukey post-hoc tests showed that the minor chord ($M = 27$) took significantly longer to learn than both the sus4 ($M = 22.33$, $p = 0.011$) and the aug chord ($M = 22.16$, $p = 0.009$), but not the flat5 chord ($M = 25.17$, $p = 0.464$).

D root training

Acquisition. We found a main effect of Vincentized block in the ANOVA looking at the DRs of birds by Sex $\times$ Vincentized blocks $\times$ Chord type, both in the

separated (minor VS sus4 VS flat5 VS aug) and collapsed (3rd manipulations VS 5th manipulations) versions of the analysis (F(9, 36) = 36.96, p < 0.001, partial η^2 = 0.902) where DR increased as Vincentized blocks increased. The main effect of chord did not reach significance in the separated analysis (F(3, 12) = 3.20, p = 0.062, partial η^2 = 0.445), however it was significant in the collapsed analysis (F(3, 12) = 34.89, p = 0.004, partial η^2 = 0.897). The collapsed analysis results suggests that, during D root training, the birds learned the 5th manipulations (aug M = 0.77; flat5 M = 0.77) to a higher DR than the 3rd manipulations (minor M = 0.73; sus4 M = 0.73).

Blocks to Criterion. Figure 3-3 shows the blocks to criterion for each bird for each chord. Four of the six birds learned to discriminate the minor chord from the major chord last, and the other two learned to discriminate either the aug or the sus4 chord from the major chord last. All birds except one learned to discriminate one or both of the 5th manipulation chords from the major chord first (flat5 and aug); one bird learned to discriminate the flat5 and minor chord from the major chord first. The analyses on blocks to criteria showed that there was a main effect of chord (F(3, 12) = 4.61, p = 0.023, partial η^2 = 0.535). However, Tukey post-hoc tests showed no significant differences between the chords, although the minor chord (M = 21.67) approached a significantly longer learning time compared to the flat5 (M = 16.17, p = 0.068) and aug (M = 16.33, p = 0.078) chord, but not the sus4 chord (M = 21.17, p = 0.994). This suggests that the chords that had lower DRs may also have taken longer to learn.

Savings

Figure 3-2 shows acquisition of the initial C root discrimination (upper panel) and the subsequent D root discrimination (lower panel). A comparison of the pattern of responding to S+ and S- stimuli in the early stages of each panel suggest that training from the C root discrimination aided in solving the subsequent D root discrimination. That is, there appeared to be savings in the D root discrimination such that responding to S+ and S- chords was already diverging early on in training by comparison to that seen in early trials of C root training. On average, the DRs during the first block of D root training appeared higher (M = 0.58) compared to C root training (M = 0.51). The same was true for the individual chords (C minor M = 0.52, D minor M = 0.59; C sus4 M = 0.51, D sus4 M = 0.58; C flat5 M = 0.51, D flat5 M = 0.53; C aug M = 0.50, D aug M = 0.61). Only one bird was at criterion within the first block of D root training (i.e., had a DR ≥ 0.80) for two of the four chords (aug and sus4), which also had the best average performance the end of the C root training.

The binomial tests for dichotomous data showed that while only one bird was above chance during the C root discrimination (bird 48110 required the fewest blocks to complete the discrimination; z = 1.90, p = 0.029), three birds responded above chance during the first block of the D root discrimination (all zs ≥ 2.29, all ps ≤ 0.011). Thus it appears that there was savings of the C root discrimination to the D root discrimination, although only for half the birds which is less than what was seen in some previous work with biologically relevant

75

stimuli (e.g., Bloomfield & Sturdy, 2008; Guillette, Reddon, Hurd & Sturdy, 2009).

Experiment 2: Humans

Method

Nineteen undergraduate students (ages range 18-22 years; 15 reported private lessons in musical training, two reported no private lessons but group lessons or self-taught ability in music) at the University of Alberta completed the task for course credit.

Apparatus & Stimuli

Participants were tested using a laptop computer (Toshiba Tecra, Intel Pentium M processor and Intel 855 series chip set) equipped with Sennheiser HD 580 headphones. Computer software (programmed in Visual Basic) tested the participants in a task directly comparable to the chickadees in Experiment 1 using identical chord stimuli. Participants had the option of adjusting the volume to a comfortable level using a dial at the front of the laptop (see Weisman et al. 2010).

Procedure

After completing a survey about musical training, participants were

trained to discriminate the major chord from the four comparison chords used in Experiment 1. S+ and S- trials were selected randomly an equal number of times within each test: the major chord was played 48 times, and all other chords were played 12 times each for a total of 96 trials. Each participant completed this test, first with the C root chords and then with the D root chords.

The first ten of the nineteen participants completed each C and D root chord training twice. At that point, we found that all participants were successfully discriminating the C chords by the end of the first test. Thus, subsequently, we had participants perform each test only once. We found no statistical differences between participants who completed one or two tests. Thus, the data for all nineteen participants were collapsed and only the first test used for all analyses.

Participants initiated a trial by clicking on the "play" button on the screen to hear a chord. If the major chord played and the participant pressed a button labeled "S+" the word "correct" appeared in a box adjacent to the S+ button as visual feedback. If the participant clicked the S+ button for any other chord, the word "incorrect" appeared in a box adjacent to the S+ button and the next trial was delayed by 5-s. Participants were told that sounds had been placed in two categories, but were given no hints as to how to categorize the sounds.

Statistical analyses

All analyses were conducted in Statistica version 8.0 (Statsoft Inc.). An

alpha level of 0.05 or less was used to judge significance in statistical tests.

Overall Performance

We evaluated performance by comparing the DRs for each chord calculated for all 96 trials within subjects using sex as a between-subjects categorical variable in an ANOVA for both C root and D root training. This analysis was repeated collapsing the chords into 3rd manipulations (minor and sus4) and 5th manipulations (aug and flat5) to make it directly comparable to previous pigeon work (Brooks & Cook, 2010) and Experiment 1 with the chickadees.

Savings

We also looked at whether there were any savings from C root training for individuals by testing the proportion of responding to the first exposure of the S-chords in C and D root training against chance (0.5) using single sample t-tests.

Results

Participants learned the task extremely quickly (often within the first two exposures to each chord type, see Figure 3-4). All participants met criterion for the first test (C root) except one. This one participant expressed confusion after completing the task and pressed go on almost every trial (94/96 trials) and as a result this participant's data were excluded from the analysis. All other

participants were above 0.8 DR for at least two of the four chords during the 96 trial block and only three participants had DRs for any chords that were less than 0.7.

Overall Performance

C Root Training

During the C root training, we found a main effect of chord ($F(3,14) = 4.61$, $p = 0.007$, partial $\eta^2 = 0.224$). Mean DR for the different chords types across the 96 training trials were flat5 ($M = 0.93$), aug ($M = 0.89$), minor ($M = 0.86$), sus4 ($M = 0.84$). We conducted dependent t-tests to determine the particular differences among the chord types underlying the main effect of chord. We found that the flat5 chord had a significantly higher DR than both the sus4 ($t(17) = 3.26$, $p = 0.005$) and the minor ($t(17) = 3.20$, $p = 0.005$) chord. The collapsed analysis confirms this pattern and shows a main effect of chord type, such that 3rd manipulations ($M = 0.85$) were more difficult to discriminate compared to 5th manipulations ($M = 0.91$; $F(1, 16) = 12.53$, $p = 0.003$, partial $\eta^2 = 0.439$).

D Root Training

During the D root training, a similar pattern emerged compared to C root training. There was a main effect of chord ($F(3,14) = 6.51$, $p = 0.001$, partial $\eta^2 = 0.289$). The degree of learning appeared to be similar when comparing DRs for the flat5 ($M = 0.98$), aug ($M = 0.96$), minor ($M = 0.95$), and sus4 ($M = 0.90$)

79

chords. We conducted dependent t-tests to determine the particular differences among the chord types underlying the main effect of chord. We found that the sus4 chord had a significantly lower DR than both the flat5 ($t(17) = 3.47$, $p = 0.003$) and aug ($t(17) = 3.04$, $p = 0.007$) and minor chord ($t(17) = 2.31$, $p = 0.034$). The minor chord also had a significantly lower DR than the flat5 chord ($t(17) = 2.40$, $p = 0.028$). The collapsed analysis supported this further with a main effect of chord type ($F(1, 16) = 11.45$, $p = 0.004$, partial $\eta^2 = 0.417$) where manipulations of the 3rd ($M = 0.92$) had a lower average DR than manipulations of the 5th ($M = 0.97$).

Savings

We compared the proportion of responding during the first exposure to all S- chords by each individual (i.e., collapsing across chord) for D root training against chance, and we found a significant difference ($t(17) = 2.17$, $p = 0.045$), which we did not find when comparing the proportion of responding during the first exposure to S- chords by each individual for C root training against chance ($t(17) = 0.20$, $p = 0.842$). Indeed, in Figure 3-4, one can already see the separation of the S+ and S- chords in trial one for the D root training, which is not the case for C root training.

Comparative Analysis

In order to make the results from the chickadees and humans in our study and the previously studied pigeons more directly comparable, we plotted the proportion of errors to each S- chord for each species (see Figure 3-5). For chickadees, we used the proportion of errors in the final Vincentized block of acquisition. For humans, we used the proportion of errors across the whole session because of how quickly they learned the task. For pigeons, we used the proportion of errors once they reached a steady state of responding (see Brooks & Cook, 2010). The figure shows that pigeons and humans were very similar in both the rank order (sus4 was the most difficult, followed by the minor, aug and the flat5) and magnitude of their errors during C root training, while chickadee errors were less to the sus4 chord and more to the flat5 chord than the other two species. In D root training, pigeons and humans continued to follow the same pattern of which chords were the most difficult, but with more variable magnitudes. Chickadees continued to differ in the order of most to least errors from the other two species in that the sus4 was not as difficult as the minor chord to discriminate for the chickadees.

Discussion

Here we show that both black-capped chickadees and humans are able to discriminate triadic chords with the same root. In contrast to the previous work

with pigeons, all chickadees learned to discriminate these chords, whereas only a subset of pigeons were able to learn the discrimination (Brooks & Cook, 2010), as was also the case in other avian studies with chord stimuli (e.g., Watanabe et al., 2005; Hulse et al., 1995). For all three species (humans, chickadees and pigeons) now tested with this set of common chord stimuli, manipulations of the 3rd note (minor and sus4 chords) are more difficult to discriminate from the major chord than chords with manipulations of the 5th note (aug and flat5 chords). However, more detailed comparisons among the individual chords create a more complex picture.

We expected that humans and chickadees would both be able to solve the task and exhibit similar patterns of discrimination with one another. In a general sense, this is true: both humans and chickadees are able to discriminate among stimuli with complex harmonic structures and do so in a comparable manner. However, there are subtle differences that that may reflect different ecological niches for the different species. For example, differences in the rate of acquisition in chickadees are likely a function of the key or absolute pitch of the chords (i.e., whether they were C or D root chords). This is a possibility because songbirds are very adept at absolute pitch processing and have no trouble distinguishing tones that differ by roughly a semitone (Weisman et al., 1998; Weisman, Njegovan, Williams, Cohen & Sturdy, 2004; Friedrich, Zentall & Weisman, 2007; for review see Weisman, Williams, Cohen, Njegovan & Sturdy, 2006). For instance, chickadees found both 5th manipulations clearly easier to discriminate when the

root note was D, but not when the root was C. When the root was C, the sus4 chord (3rd manipulation) was relatively easy for them to discriminate, and the flat5 was relatively more difficult for them to discriminate. This differential responding depending on the root of the chord was not found in either humans or pigeons, who generally responded similarly to the four chord types regardless of the root note (but see Brooks & Cook, 2010, for some exceptions in pigeons). However, this effect in chickadees may be partially due to the fact that all birds were trained first with C root chords and then D root chords. This suggests that the absolute pitch of the notes may be affecting the manner in which chickadees process complex auditory stimuli, a common finding in studies of auditory perception in songbirds (Cynx et al., 1986; Hulse, et al., 1984; Njegovan & Weisman, 1997). However, it is also possible that humans are similarly affected by changes in absolute pitch, but their strong reliance on relative pitch masks any differential responding based on absolute pitch. Part of this reliance of relative pitch by the human participants in our study could be affected by the fact that Western music has an equally tempered scale (i.e., all notes are separated by the same distance). This is potentially relevant because it means that different keys (e.g., C major or D major) make use of the same relative pitch ratios and make absolute pitch less important than in a musical system with an unequally tempered scale. The exact reason for the differences among species is at the moment not clear, especially given that pigeons seem to be responding more similarly to humans than chickadees. This is somewhat surprising since humans and

chickadees, but not pigeons, are vocal learners (see Tyack, 2008 for a recent review of vocal learners) and songbirds and humans have similar vocal learning behaviors, and neural pathways underlying vocal production (Doupe & Kuhl, 1999). Songbirds are also thought to share perceptual abilities required for musical perception such as rhythm perception (see Schachner, Brady, Pepperberg & Hauser, 2009), an ability with which pigeons struggle (Hagmann & Cook, 2010).

The chickadees also consistently found the minor chord the most difficult to discriminate from the major chord, while humans consistently found the sus4 the most difficult to discriminate from the major chord. This may be due to musical experience in the humans we tested, because the minor chord more commonly occurs in Western music than the sus4 chord. Further support from this comes from the fact that the sus4 introduces novel intervals between the notes, compared to the major chord, which the minor chord does not. Specifically, the 3rd note in the major chord forms a major 3rd with the 1st note and a minor 3rd with the 5th note. In the minor chord, these relationships are swapped such that the 3rd note makes a minor third with the 1st note and a major 3rd with the 5th note. The 3rd note in the sus4 chord, on the other hand, forms a perfect 4th with the 1st note and a major 2nd with the 5th note. The perfect 4th is considered more consonant and the major 2nd less consonant than either 3rd interval (major and minor 3rds; Rasch & Plomp, 1999). To an animal that has little to no experience with western music it is possible therefore, that a sus4 would be easier to

84

discriminate. However, the pigeons also showed much more difficulty in discriminating the sus4 chord from the minor in C root chords, suggesting that there might be something else at work here. One possibility that the chickadees may have found the sus4 chord easier to discriminate because the sus4 includes a major 2nd interval, which is very close to the interval used in their song to advertise male quality (Christie, Mennill & Ratcliffe, 2004).

Despite the differences, however, when we compare overall responding to the chord types regardless of root (see bottom panel of Figure 3-5), the species do respond similarly, with 3rd manipulations being more difficult than 5th manipulations. One possible explanation for the similar overall responding across all three species may be related to the perceived consonance and dissonance of the different chord types. Helmholtz (1877) originally proposed that pitch intervals were either consonant or dissonant because of the simplicity or complexity of the relationships of the harmonics. He pointed out that dissonant intervals, such as the minor second or major seventh, have few shared harmonics and their harmonics often are close, creating a roughness or beating sound that musicians listen for when tuning an instrument. On the other hand, perfect intervals such as a perfect 5th (ratio of 2:3) and octave (ratio of 1:2), have simple relations with many shared upper harmonics and are perceived more like a single note. If this is the case, it may be that we would not have found the same results had we used pure tones for the creation of the chords because of their lack of upper harmonics. Testing with pure tones may be a potential route that might lead to further insight on this

matter. However, pure tones are thought to have a similar effect to that of harmonics because of the matching of phase information for neural firing patterns (Burns, 1999), so it may not matter whether or not pure tones are used.

A potential problem with Helmholtz' definition of consonance and dissonance is that this does not fully correspond to how humans treat the chords in many experiments, including the current study. By Helmholtz' reasoning alone, the aug chord should sound relatively consonant because of the simple harmonic relationships amount the notes, but this is not what is generally found. Other factors are thought to contribute to consonance and dissonance, but there are no agreed upon definitions. One model of consonance and dissonance that seems to accurately reflect ratings of consonance and dissonance across cultures and therefore may be of greater universal significance is Cook & Fujisawa's (2006) model. In addition to beating harmonics, they take into account the "tension" of the chord in deriving the formulas for their model. A chord is thought to be more tense if the intervals between all notes are similar than if they are different (e.g., the relations of the 3rd to the 1st and 5th note of the aug chord are both major 3rds i.e., both four semitones apart). A more "tense" chord is also more dissonant and this may be because of cognitive grouping mechanisms whereby it is more difficult to group tones based on relative distance when all distances are equal (N. D. Cook & Fujisawa, 2006). This fits well with our current results for both humans and black-capped chickadees, as well as the previous results with pigeons, who found the aug chord was quite different to the major chord, despite

86

the similar consonance according to Helmholtzs' definition. In fact, the avian

species found this chord especially easy to discriminate regardless of the absolute

pitch of the chords.

One future avenue for our research would thus be to systematically alter

the tension and relationship of the upper harmonics in intervals to examine what

controls consonance and dissonance in both humans and other animals. Because

humans from different musical cultures have variations in what they perceive as

consonant and dissonant (Carterette & Kendall, 1999), animal research may be

able to address which potential factors contributing to musical perception are

cultural influences, and which are fundamental components of complex auditory

processing.

Another important consideration for our study is that all chord changes

were being evaluated against the major chord and not directly against perfect

consonance. The 3rd in the major chord forms a major 3rd interval and a minor

3rd interval between the 1st and 5th note respectively, while the 5th note forms a

perfect 5th and a minor 3rd, with the 1st and 3rd note respectively. The frequency

of the notes forming a perfect 5th have a ratio of 2:3 and this interval is

considered the most consonant interval after unison and the octave. Major and

minor 3rds, while still considered consonant, are thought to be less consonant in

comparison to the perfect intervals (Rasch & Plomp, 1999). Thus, part of the

reason the 5th alterations were easily differentiable from the major chord is

potentially because of the greater change along a consonance/dissonance

continuum.

By using complex musical chords as stimuli, the current experiment also provided additional insight into how chickadees perceive auditory stimuli. One example of this is the observation of savings when the root changed from C to D. Half the chickadees show savings with a DR higher than chance in the first block of D root training, and one chickadee reached criterion for some chords within the first 500-trial block (sus4 and aug chord). It may therefore be possible for chickadees to more heavily rely on relative pitch rather than absolute pitch in some circumstances; however it could also be that chickadees are simply learning the second discrimination more rapidly because they learned what types of differences to attend to among chords. In other words, if we had presented a second group of birds with a D root chord task that used chords with different relative pitch relations, would this have been learned as rapidly? This is something that we are interested in investigating further. All birds in our experiment were able to learn the chords with sufficient training, which is different from the previous work with sparrows, starlings and pigeons (Watanabe et al., 2005; Hulse et al., 1995; Brooks & Cook, 2010). It may be that chickadees in particular are better at relative pitch tasks because they attend to relative pitch cues in their song (Weisman, Ratcliffe, Johnsrude & Hurly, 1990; Weisman & Ratcliffe, 2004; Christie et al., 2004). However, past research with descending intervals suggests that it can be very difficult to train black-capped chickadees on relative pitch tasks (Njegovan & Weisman, 1997). It is possible that, because our

current task presented all notes simultaneously, that the chickadees were better able to use relative pitch rather than memorizing the absolute pitch of one of the notes. Further training with more chord sets using different roots could lead to insight on the matter of whether chickadees are better able to switch to using relative pitch for pitch categorization than previously assumed. Additionally, comparisons between presentations of harmonic chords, where all notes are played simultaneously, and broken chords, where notes are played individually, may reveal that the apparent reliance on absolute pitch in songbirds may not be due to a preference for absolute pitch processing, per se, but rather could be rooted in the demands of the tasks such as attentional or memory differences among experiments. To date, most studies looking at relative mechanisms in chickadees (Njegovan & Weisman, 1997) and other songbirds (Cynx et al., 1986; Hulse et al., 1984) have used broken chords rather than harmonic chords.

There are several broad implications of our comparative perceptual work with musical stimuli. As West, King and Goldstein (2004) suggest, music may be a building block of vocal communication shared across a wide variety of taxa. Masataka (2009) further supports this claim by suggesting that a primitive form of modern human music was potentially the foundation of both modern music and language in humans. It is possible that there may be an underlying mechanism of consonance and dissonance that is present across all cultures based on properties of chords (see model by N. D. Cook & Fujisawa, 2006) and the perception of these properties appear to also be predictable by the structure of human speech

(Schwartz, Howe & Purves, 2003). This evidence, along with evidence suggesting that pitch intervals may be used to convey similar messages in speech (at least the minor interval, Curtis & Bharucha, 2010), suggests that consonance and dissonance may have an evolutionary basis rather than simply being a human cultural phenomenon. By studying both other mammals, especially other primates, and distantly related both vocal learning and non-vocal learning animals, we can begin to tease apart what parts of our musical perception are fundamental components of auditory perception in all hearing animals, and what parts of musical perception are required for complex auditory processing (analogous) or present only in closely related species (homologous). Because of how similarly pigeons and humans responded compared to chickadees, our results potentially suggest that more fundamental auditory processing components found in a non-vocal learning pigeon are found in both humans and chickadees, and a higher-level ecologically-relevant component may be influencing the further differences among chickadees (e.g., species specific vocalizations) and humans (e.g., cultural influences).

Figure 3-1.

Figure 3-2.

Figure 3-3.
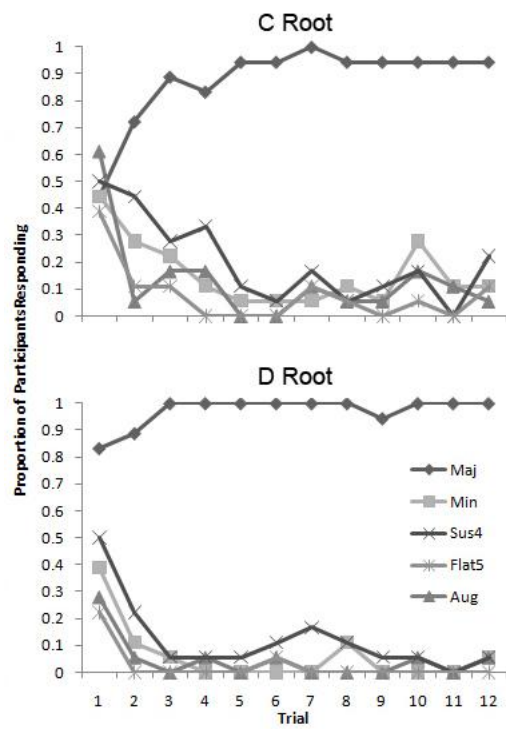
Figure 3-4.
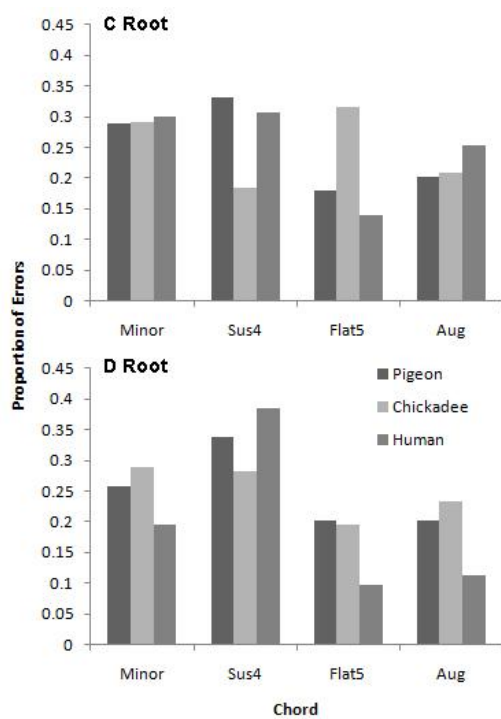
Figure 3-5.

References

Bachem, A. (1955). Absolute pitch. *Journal of the Acoustical Society of America, 27*(6), 1180-1185.

Bloomfield, L. L., & Sturdy, C. B. (2008). All "chick-a-dee" calls are not created equally. Part I: Open-ended categorization of chick-a-dee calls by sympatric and allopatric chickadees. *Behavioural Processes, 77*(1), 73-86.

Bolhuis, J. J. Okanoya, K. & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Neuroscience, 11*, 747-759.

Brooks, D. I., & Cook, R. G. (2010). Chord Discrimination by Pigeons. *Music Perception 27*(3), 183-196.

Burns, E. M. (1999). Intervals, Scales, and Tuning. In D. Deutsch (Ed.), *The Psychology of Music 2nd Ed* (pp. 215-264). San Diego, California: Academic Press.

Carterette, E. C., & Kendall, R. A. (1999). Comparative Music Perception and Cognition. In D. Deutsch (Ed.), *The Psychology of Music 2nd Ed* (pp. 725-792). San Diego, California: Academic Press.

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.

Cook, N. D., & Fujisawa, T. X. (2006). The Psychophysics of Harmony Perception: Harmony is a Three-Tone Phenomenon. *Empirical*

Musicology Review, 1(2), 106-126.

Curtis, M. E., & Bharucha, J. J., (2010). The Minor Third Communicates Sadness
in Speech, Mirroring Its Use in Music. *Emotion, 10*(3), 335-348.

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*(4), 394-402.

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 Inc.

Ficken, M. S., Ficken, R. W., & Witkin, S. R. (1978). Vocal repertoire of the
black-capped chickadee. *The Auk, 95*(1), 34-48.

Fitch, W. T. (2005). The evolution of music in comparative perspective. *Annals of
the New York Academy of Sciences, 1060*, 29-49.

Friedrich, A., Zentall, T., & Weisman, R. G. (2007). Absolute pitch: Frequency-
range discriminations in pigeons (Columba livia) - Comparisons with
zebra finches (*Taeniopygia guttata*) and humans (*Homo sapiens*). *Journal
of Comparative Psychology 121*(1), 95-105.

Guillette, L. M., Reddon, A. R., Hurd, P. L., & Sturdy, C. B. (2009). Exploration
of a novel space is associated with individual differences in learning speed
in black-capped chickadees, *Poecile atricapillus*. *Behavioural Processes,*

82, 265–270.

Guillette, L. M., Reddon, A. R., Hoeschele, M., & Sturdy, C. B. (2011).

Sometimes slower is better: slow-exploring birds are more sensitive to

changes in a vocal discrimination task. *Proceedings of the Royal Society

B, 278*, 767-773.

Griffiths, R., Double, M. C., Orr, K., & Dawson, R. J. G. (1998). A DNA test to

sex most birds. *Molecular Ecology, 7,* 1071-1075.

Hagmann, C E., & Cook, R. G. (2010). Testing meter, rhythm, and tempo

discriminations in pigeons. *Behavioural Processes, 85*, 99-110.

Hauser, M. D., & McDermott, J. (2003). The evolution of the music faculty: a

comparative perspective. *Nature Neuroscience, 6*, 663-668.

Helmholtz, H. L. F. (1877). *On the sensations of tone as a physiological basis for

the theory of music*. New York: Dover Publications.

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*(3), 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(4), 409-423.

Hulse, S. H., & Cynx, J. (1985). Relative Pitch Perception Is Constrained by

Absolute Pitch in Songbirds (*Mimus*, *Molothrus*, and *Sturnus*). *Journal of

Comparative Psychology, 99*(2), 176-196.

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*(1), 38-54.

Justus, T., & Hutsler, J. J. (2005). Fundamental issues in the evolutionary psychology of music: assessing innateness and domain specificity. *Music Perception, 23*(1), 1-27.

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

Kunej, D., & Turk, I. (2000). New perspectives on the beginning of music: Archeologicla and musicological analysis of a middle paleolithic bone "flute". In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music*. Cambridge, MA: MIT Press.

Lee, T. T. Y., Charrier, I., Bloomfield, L. L., Weisman, R. G., Sturdy, C. B., (2006). Frequency-range discriminations and absolute pitch in black-capped chickadees (*Poecile atricapillus*), mountain chickadees (*Poecile gambeli*), and zebra finches (*Taeniopygia guttata*). *Journal of Comparative Psychology 120*(3), 217-228.

Masataka, N. (2009). The origins of language and the evolution of music: a comparative perspective. *Physics of Life Reviews, 6*, 11-22.

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

Njegovan, M., & Weisman, R. G., (1997). Pitch Discrimination in Field- and

Isolation-Reared Black-Capped Chickadees (*Parus atricapillus*). *Journal of Comparative Psychology, 111*(3), 294-301.

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.

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. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception, 24*(1), 99-104.

Patel, A. D., Iversen, J. R., Bregman, M. R. & Schultz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology, 19*, 1-4.

Pinaud, R., & Terleph, T. A. (2008). A songbird forebrain area potentially involved in auditory discrimination and memory formation. *Journal of Biosciences, 33*(1), 145-155.

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

Rasch, R., & Plomp, R. (1999). The Perception of Musical Tones. In D. Deutsch (Ed.), The Psychology of Music 2nd Ed (pp. 215-264). San Diego, California: Academic Press.

Schachner, A., Brady, T.F., Pepperberg, I.M., & Hauser, M.D., (2009).

Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology, 19*(10), 831-836.

Schwatz, D. A., Howe, C. Q., & Purves, D. (2003). The Statistical Structure of Human Speech Sounds Predicts Musical Universals. *The Journal of Neuroscience, 23*(18), 7160-7168.

Shackleton, S. A., & Ratcliffe, L. (1993). Development of song in hand-reared black-capped chickadees. *Wilson Bulletin, 105*(4), 637-644.

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

Trehub, S. E., Schellenberg, E. G., & Kamenetsky, S. B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 965-975.

Tyack, P.L., (2008). Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *Journal of Comparative Psychology, 122*(3), 319-331.

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

Weisman, R .G., Balkwill, L.-L., Hoeschele, M., Moscicki, M. K., Bloomfield, L. L., & Sturdy, C. B. (2010). Frequency-range discriminations in boreal chickadees and humans: Exceptions that test a phylogenetic rule. *Learning and Motivation, 41*, 156–173.

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

Weisman, R. G., Njegovan, M., & Ito, S. (1994). Frequency ratio discriminations
by zebra finches (*Taeniopygia-guttata*) and humans (*Homo-sapiens*).
Journal of Comparative Psychology 108(4), 363-372.

Weisman, R. G., Njegovan, M., Sturdy, C. B., Phillmore, L. S., Coyle, J., &
Mewhort, D. J. K. (1998). Frequency range discriminations: special and
general abilities in zebra finches (*Taeniopygia guttata*) and humans
(*Homo sapiens*). *Journal of Comparative Psychology, 112*(3), 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. *Behavioural processes, 66*, 289-307.

Weisman, R. G., & Ratcliffe, L. M. (2004). Relative pitch and the song of black-
capped chickadees. *American Scientist, 92*(6), 532-539.

Weisman, R. G., Njegovan, M. G., Williams, M., Cohen, J., & Sturdy, C. B.
(2006). The comparative psychology of absolute pitch. In E. A.
Wasserman & T. Zentall (Eds.), *Comparative Cognition: Experimental
Explorations of Animal Intelligence* (pp. 71-86). Oxford: Oxford
University Press.

West, M. J., King, A. P., & Goldstein, M. H. (2004). Singing, socializing and the
musical effect. In P. Marler & H. Slabbekoorn (Eds.), *Nature's Music* (pp.

374-387). San Diego: Elsevier Academic Press.

Chapter 4 Pitch chroma discrimination, generalization and transfer tests of octave equivalence in humans[3]

Introduction

Two acoustic events are separated by an octave when the frequency of the second event is double or half the frequency of the first. This logarithmic relationship between acoustic events spaced an octave apart is a description of the physics of wave transmission. Human perception has evolved to grasp this unique acoustic relationship in speech and music (e.g., Burns, 1999; Patel, 2003; Peter, Stoel-Gammon, & Kim, 2008). In all cultures, production and perception of the octave are fundamental characteristics of music (Crickmore, 2003). That is, although the number of notes in an octave, their labels, and their frequencies can differ in music across cultures, all cultures recognize the similarity between notes an octave apart (the notes are said to have the same pitch chroma); this phenomenon is known as octave equivalence.

Octave equivalence is one of the two most potent determinants of pitch judgments. A second important determinant is pitch height. Pitch height is a log-linear scale of pitch in which the more two sounds differ in frequency, the more they differ in pitch. Octave equivalence and pitch height are opposing percepts. For example, a note one-third of the way between two notes separated by an octave is more similar in pitch height to the first note than the second. Whereas, the two notes separated by an octave are more similar to one another in chroma than to the note one-third of the way between them.

Perception of octave equivalence has a neural basis: for example, neurons in the auditory midbrain show preferences for harmonically related sounds and

the ventral nucleus of the lateral lemniscus has a structure reminiscent of the pitch helix (Langner & Ochse, 2006). The pitch helix is a spiraling structure that completes a circular motion once in each octave. The helix was first proposed as a theoretical spatial mapping of human pitch height and chroma perception by Moritz Wilhelm Drobisch (ca 1846), then used by Shepard (1982) much later in his well-known theory of pitch perception.

Despite all the evidence and theory supporting the importance of octaves, experimental demonstration of the perception of octave equivalence has been problematic. In the identification of simple melodies, alteration of the octaves of several of the notes reduces identification of melodies (Deutsch, 1972). However, by maintaining pitch contour (the direction of frequency change) in the melody, identification was partially restored (Dowling & Hollombe, 1977). This makes sense because contour is a form of pitch height information. Reducing the effects of pitch height should make pitch chroma a more salient feature.

Not all instances of octave equivalence testing involve music. Tests that require even limited amounts of music training are restrictive because they preclude unbiased testing of nonmusicians. But testing outside of a music context has been problematic. Allen (1967) asked participants to rate the similarity of two notes, including notes that were separated by an octave. He found only musicians showed octave equivalence. Nonmusicians rated notes that were more similar in pitch height (i.e., frequency) as more similar than notes played at the same chroma but in an adjacent octave. Krumhansl and Shepard (1979) also found that

musically untrained participants relied more on pitch height than chroma. Kallman (1982) conducted similar experiments to Allen (1967) and in most of these, the effects of octave equivalence were small or nonexistent. Only in experiments that manipulated pitch height, such that the two comparison notes were always an octave or close to an octave apart, was there some evidence of octave generalization in nonmusicians. In other words, again, it appears that only when the effects of pitch height are greatly limited is it possible to observe even modest effects for pitch chroma. Overall, because of differences in procedure, the literature has produced contradictory results concerning octave equivalence. As Burns (1999, p. 252) noted in his review: "If the results of some relevant experiments are accepted at face value, octave equivalence is shown by rats (Blackwell & Schlosberg, 1943), human infants (Demany & Armand, 1984), and musicians (Allen, 1967), but not by starlings (Cynx, 1993), 4- to 9-year-old children (Sergeant, 1983), or nonmusicians (Allen, 1967)."

In summary, notwithstanding presumed prominence of octave equivalence in music and speech, experimental evidence for octave equivalence is sparse and contradictory. The purpose of the present research was to develop a music-independent protocol for studying octave equivalence and to use that methodology to enlarge our knowledge about equivalence in humans. In addition, we hoped that our protocol might clarify contradictory past findings lost in the myriad of protocols used in the prior literature. We adopted operant go/no-go procedures because of their positive impact in our prior research on pitch height

107

perception and their immediate usefulness in studying octave equivalence in humans of all ages and verbal abilities and in other species (e.g., Weisman, Balkwill, Hoeschele, Moscicki, Bloomfield, & Sturdy, 2010a; Weisman, Hoeschele, Bloomfield, Mewhort, & Sturdy, 2010b).

Experiment 1

The starting point for the present research was Cynx's (1993) study of octave equivalence in European starlings (*Sturnus vulgaris*). Cynx (1993) trained starlings in an operant go/no-go discrimination between 2200 Hz and 1000 Hz sinewave tones. The 2200 Hz tone was S+ (the go signal for food reward) and the 1000 Hz tone was S- (the no-go signal for no food reward) in one group; the go and no-go signals were reversed in the second group. During the generalization test, starlings heard 26 probe tones ranging from 1038 Hz to 2119 Hz. The critical probe pitches for the octave generalization test were 1100 and 2000 Hz, spaced exactly an octave from the go signal in the first (S+ = 2200 Hz) and second groups (S+ = 1000 Hz), respectively. Cynx predicted that if the birds heard octave equivalence, then they should have confused the go tone and the octave generalization probe tone. In fact, responding fell in an orderly monotonic function from the go to the no-go tone without an increase at the octave equivalent tone, which is what would be expected if the birds were using pitch

height alone. Starlings showed no evidence of octave generalization; Cynx (1993) concluded that it was unlikely that they perceived octave equivalence. Before accepting Cynx's (1993) conclusions about songbirds, we sought to determine whether Cynx's procedure could be used to show octave generalization in humans. If the replication failed with humans, then we need to develop a more effective procedure for testing humans: one that takes pitch height into account and can also be adapted for training birds. Humans with absolute pitch (AP) can usually identify the chroma of a pitch across octaves, so we prescreened the participants to identify AP possessors using a note-naming test (Athos et al., 2007).

Method

Participants

Twenty-eight students at Queen's University participated for course credit. They provided their ages and the details of their music and language training in written responses to a questionnaire. Each gave informed written consent and the General Research Ethics Board at Queen's University approved our research protocols.

Participants ranged in age from 17 to 22 years old, M = 18; 7 were men and 21 were women. Twenty-six participants were enrolled in a first-year psychology course and completed the experiment for course credit. Two of the

109

original participants were AP possessors and two additional AP possessors were recruited especially for this study and paid $10 for participating. We determined participants' AP status using Athos et al.'s (2007) note naming test (see our results).

Because musical (e.g., Allen, 1967; Krumhansl & Shepard, 1979) and language training (e.g., Deutsch, Henthornn, & Dolson, 2004; Pfordresher & Brown, 2009) are sometimes factors in music perception, we have provided more information about the participants' histories. Four participants had no formal music training, 11 began their training with the piano; the remainder had training in voice or a variety of musical instruments. Among musically trained participants, the amount of training varied from 1 to 14 years, M = 7.5; 21 played at least one additional instrument, and 20 still played at least one instrument; only 5 had passed formal examinations in music. Nineteen participants learned English as their first language, 4 learned Korean first, the remaining 4 learned three other languages.

Apparatus

Training and testing were conducted on a Toshiba 149 Tecra laptop (Intel Pentium M processor and Intel 855 series chip set) using Sennheiser HD 580 headphones. Participants used a mouse to make their responses. The procedures and data collection were programmed in Visual Basic. Participants could use a rotary control on the computer to adjust the volume to the headphones at any time

during the experiment.

Stimuli and Procedures

The experiment consisted of three phases: a test for AP ability, auditory discrimination training, and an auditory generalization test. The second and third phases were adapted from Cynx's (1993) study.

AP Testing

The protocol was adapted from a procedure used by Athos et al. (2007) to test 2,213 participants: the note durations and frequencies were a direct replication of Athos et al. (2007). In the note-naming tests, we identified AP possessors using Athos et al.'s (2007) scoring protocol: 1 point for each correct identification and 0.75 points for responses to notes ±1 semitone from the correct note.

Sinewave tones presented in the test were synthesized at the frequencies of 40 notes randomly sampled from the 66 notes on the chromatic scale that spans the 51/2 octaves from C2 to G8, on the basis of A4 = 440 Hz; each note was played for 1000 ms (see Athos et al., 2007). The actual notes presented were D#2, F2, F#2, G#2, A#2, B2, C#3, D#3, E3, F3, G3, G#3, C4, C#4, D4, D#4, F4, F#4, A4, C5, C#5, D5, E5, F#5, G5, G#5, A5, A#5, C6, D6, A6, B6, C#7, D#7, F#7, B7, E8, F#8, G8, A#8. These tones and all the others presented in this study were constructed at a standard 16-bit, 44.1 kHz sampling rate and ramped at onset and offset, respectively, upwards and downwards for 5 ms. Because four of the

sinewave tones lie above the notes on the piano keyboard (in Octave 8) and

proved difficult to identify, participants rarely named them accurately. In practice,

therefore, the test consisted of 36 notes (see Athos et al., 2007).

The test began after a short practice session (8 trials) given to acquaint

participants with making mouse responses to graphics on the screen and to allow

participants to individually adjust tone amplitude to a comfortable level. During

the practice session and the test, a participant clicked on the "play" button at the

top of the screen, and heard a tone selected randomly without replacement from

the 40 test tones, which controlled for any possible predictable relative pitch

carryover effects between tones (Ward & Burns, 1982). To "name" the musical

note corresponding to a tone, the participant clicked on one of 12 black and white

piano keys shown on the screen. The test continued without feedback until the

participant heard all 40 tones. In this note-naming test and all following tests,

participants could take as much time as they liked between trials, as a trial began

only after clicking the play button. We did not record time between trials.

Operant discrimination training

Participants were asked to classify notes into two categories (go and no-go

tones) to the best of their ability, without any instructions about which notes made

up each category. Participants were told that discrimination training was a test of

their perceptual categorization ability but not that it was a test of octave

equivalence.

Only two frequencies were presented: 1000 and 2200 Hz sinewave tones. Each frequency was presented 50 times in a random order without replacement for 100 trials. Reward (positive feedback) was counterbalanced across two groups: a 1000 Hz S+/2200 Hz S- group and 2200 Hz S+/1000 Hz S- group. Participants initiated a trial by clicking the button labeled "play tone" on the screen to hear a tone. If a participant clicked on the button on the screen labeled "S+" after hearing a go tone, the word "correct" appeared in a box adjacent to the S+ button. If the participant clicked the S+ button on a no-go trial, the word "incorrect" appeared in a box adjacent to the S+ button; the next trial was delayed by 3 s. If a participant failed to click the S+ button after either a go or no-go tone, the trial terminated after 2 s without feedback, as is typical in go/no-go discrimination procedures. Two versions of each tone were played, one at 70 dB (SPL) and a second at 80 dB (SPL), each on 25 trials, to control for amplitude, and chosen at random without replacement. This strategy made pitch a more salient determinant of the discrimination by reducing the confounding of pitch with loudness (Moore, 1989). The initial sound pressure level (dB) of each tone was measured from the location of the ear with an integrating sound level meter (Type 2239 A, Brüel & Kjær Canada Ltd, Point Claire, Quebec, Canada). Each participant was allowed to adjust overall amplitude to a comfortable level during a short practice session (4 trials: one each of the four trial types presented during training) prior to the training session; this meant that the actual tone amplitudes heard in the discrimination task varied across participants.

Generalization testing

Participants were told that they would no longer receive feedback for responding but were asked to respond as they did during training to the best of their ability. Participants received no further instructions. Trials were identical to trials during discrimination training, except that probe tones were presented in addition to the training tones, and the feedback box was no longer presented. Hiding the feedback box eliminated reward for responding during the generalization test.

The 2 training tones and 26 probe generalization tones played at intermediate frequencies were presented during the test. We presented the same probe tone frequencies as Cynx (1993), 1038, 1059, 1100, 1122, 1165, 1189, 1235, 1260, 1308, 1335, 1386, 1414, 1468, 1498, 1556, 1587, 1648, 1682, 1746, 1782, 1850, 1888, 1960, 2000, 2076, and 2119 Hz, on 5 trials each. Two versions of the play list were used to control for amplitude: in one test version, all odd numbered probes in ascending order of frequency were played at 70 dB (SPL) and the others were played at 80 dB (SPL). In the other version, the amplitudes were reversed (i.e., 70 dB now played at 80 dB and vice versa). The two versions were counterbalanced across subjects and training conditions. In the same play lists, the training tones, 2200 and 1000 Hz were presented on 50 trials each, 25 trials at 80 dB and 25 trials at 70 dB. Each version of the play list included one trial at each probe frequency and five trials at each training frequency/amplitude combination sampled five times at random without replacement during the generalization tests

of individual participants.

Results and Discussion

Here, we report the results of AP testing, discrimination training, and testing for octave generalization.

AP testing

Using Athos et al.'s (2007) criterion score for highly accurate AP ($\geq$ 24.5 based on 36 tones), we separated participants into AP possessors, n = 4, (all summaries show M $\pm$ SE) 30.3 $\pm$. 2.2 and nonpossessors, n = 24, 6.8 $\pm$. 0.4.

Discrimination training

Every participant responded significantly more to S+ than to S- tones, n = 28, percent correct = 96% $\pm$ .5%, binominal test, p < .0001: where percent correct = (M percent response to S+ tones + [1 - M percent response to S- tones]). A mixed ANOVA found no significant effect for S+ frequency (high or low), $F(1,24) = .40$, p = .53, $\eta p2 = .02$, AP status (possessor or nonpossessor), $F(1,24) = .01$, p = .94, $\eta p2 > .01$, or the interaction, $F(1,24) = .11$, p = .74, $\eta p2 > .01$.

Generalization testing

To include the 1000 Hz S+ and 2200 Hz S+ groups in the same analysis, we calculated standardized frequencies relative to S+ by dividing the larger

frequency by the smaller frequency. In the 2200 Hz S+ group, we divided 2200 Hz by each probe frequency. In the 1000 Hz S+ group, we divided each probe frequency by 1000 Hz. Notice that the standardized frequency of the octave is always 2.0: twice or half the frequency of the S+. The results for AP possessors were indistinguishable from those for nonpossessors, so we merged their results.

Figure 4-1 shows percent response as a function of the standardized frequency of S+, S-, the octave probe, and the remaining 25 probe tones. The dashed line shows an exponential function fit to the results, excluding the octave probe: $Y = 738.7 \times 10^{-.99X}$, and the correlation of the function with the results is $r(26) = .96$, $p \leq .0001$. The choice of an exponential fit follows from the log linear relationship between frequency and pitch. The octave probe is shown as a darkened point, which fits precisely on the line for the function calculated without response values for the octave. Notice that responding to S+ and S- are not well predicted by the function: previous discrimination training resulted in near perfect discrimination between S+ and S-. In contrast, the probe stimuli were novel, so responding to them was pure generalization from previous training.

Personal histories

We asked whether the participants' gender, or previous music and language training, contributed to their ability to perform the discrimination, as measured by the percent correct, or to generalize between S+ and its octave equivalent, as measured by the percent response to the octave equivalent tone.

None of the correlations and point-biserial correlations between these two

performance measures and possible historical determinants (i.e., gender, AP

score, years of training, number of instruments, or number of languages)

approached significance, rs ≤ .16, p ≥ .45. Also, variation in first instrument

training or first language training had no significant effects on discrimination or

generalization, Fs(1,21) ≥ 1.41, ps ≤ .25, ηp2 ≤ .06.

Conclusions

Cynx's (1993) procedure failed to uncover evidence of octave

generalization (the perception of octave equivalence) in European starlings, but

more importantly here, the procedure failed to uncover evidence of octave

generalization in humans. Our findings agree with those of Kallman (1982) and

Allen (1967), who found that human participants preferred pitch height to pitch

chroma as a basis for judging the similarity of two tones. Generalization in this

task did not depend on prior musical or language experience.

Experiment 2

Although Cynx's (1993) procedure did not provide evidence of octave

equivalence in humans, it provided a starting point for our second experiment and

ties our procedural changes in the current experiment more directly with the

literature. By instituting a few important changes, we hoped to show octave equivalence with a similar operant procedure. We had several requirements for an effective test of octave equivalence. Most importantly, the new procedure had to succeed in measuring octave equivalence in humans with a broad range of musical training: from nonexistent to highly skilled. We chose a task that required minimal instructions, so that versions of the procedure could be developed for humans of all ages and verbal abilities, as well as other species. We used a go/no-go procedure to make the work compatible with Experiment 1. Also, we used variants of the go/no-go protocol to eliminate artifacts and establish the validity of our generalization measure of equivalence.

We began with discrimination training that provided feedback and reward for responding to a defined range of musical notes in the middle of an octave, the go notes, but withheld reward for responding to other notes in higher and lower pitch ranges within the same octave, the no-go notes (see Figure 4-2A). Although we had never trained a three-range discrimination to span exactly an octave, it seemed likely that humans and many other species could learn this discrimination as readily as they did other three-range discriminations (Weisman, Njegovan, Sturdy, Phillmore, Coyle, & Mewhort, 1998). The critical test of octave equivalence was generalization of the discrimination to the octave just above the training octave (see Figure 4-2B). Testing outside the pitch range of discrimination training promised to eliminate difficulties in observing octave equivalence reported by Kallman (1982) and in Experiment 1.

Method

Participants

Forty-six students at the University of Alberta and 33 students at Queen's University participated for course credit. The details of their written questionnaires and consents, and the conditions of their participation were the same as in Experiment 1. Research Ethics Boards at the University of Alberta and Queen's University approved our protocols.

Participants ranged in age from 17 to 25 years old, M = 19; 26 were men, 51 were women; 2 did not report their gender. We determined participants' AP status using Athos et al.'s (2007) note naming test. One participant was an AP possessor.

We have provided some information about the participants' histories here. Twenty-one participants had no formal music training, and 25 began their training with piano. The remainder began with a variety of other instruments including guitar and violin. Among musically trained participants, the total amount of training across all instruments varied from 3 months to 26 years, M = 7 years. Thirty-eight played at least one additional instrument, and 14 had passed formal examinations in music.

Fifty-two participants learned English as their first language, 8 learned Mandarin first, and 18 participants learned one of 10 other languages first. By

119

pooling the music and language history from Experiments 2 and 3, we were able to provide an extensive analysis of the contributions of the more common types of first music training and of first language training to the results of our discrimination training and generalization testing. These analyses are presented in the results section of Experiment 3.

Apparatus, Stimuli, and Procedures.

Several aspects of the method were unchanged from Experiment 1: (a) The apparatus and methods for producing notes; (b) the phases of the experiment, including practice trials and phase order: AP testing, discrimination training, then generalization testing; (c) the computer screens, responses, and instructions used in each phase; (d) the scoring of responses in each phase; and (e) details of the note-naming test of AP. The training sinewave tones differed in Experiment 2; they were synthesized at the frequencies of successive chromatic musical notes from C4 (262 Hz) to B5 (988 Hz).

Operant-discrimination training

We conducted discrimination training between the middle note range, and the lower and upper note ranges in Octave 4 (four notes per range, see Figure 4-2A) separately in two reward order groups. In one reward-order group: the middle-range S+ group, in the middle note range, responses to E, F, F#, or G were rewarded (with positive feedback); in the lower note range, responses to C, C#, D,

or D#, and in the upper note range, responses to G#, A, A#, or B, were not rewarded (no positive feedback). In a second reward order group: the middle-range S- group, the relationship between responding to the notes and reward was reversed, so that responses to the middle notes were not rewarded, but responses to notes in the upper and lower ranges were rewarded. This had the effect of counterbalancing the order of rewarded and unrewarded notes between the two groups, which eliminated artifacts that may have arisen from a predominance of go or no-go notes, or any other peculiarity due to the selection of go and no-go notes. The majority of participants heard two versions of each tone to control for the effects of amplitude on pitch perception (Moore, 1989), one at 70 dB (SPL) and another at 80 dB (SPL), as in Experiment 1.

We trained a subset of participants in each group with 60 dB (SPL) tones only to control for aural harmonics (Newman, Stevens, & Davis, 1937). The membranes and bones of the ear can introduce aural harmonics and the louder the tone, the louder the harmonics. Hence, playing tones at 60 dB greatly reduced the potential that aural harmonics could confound our explanation of the results. That is, by controlling for aural harmonics with the 60 dB group, we could rule out a potential confound: that the first harmonic of any sound is its octave equivalent, so the presence of loud harmonics in our training stimuli might explain a finding of octave equivalence.

Generalization testing

We conducted generalization testing using the 24 notes that comprise

Octaves 4 and 5 (Figure 4-2B), in the absence of reward (positive feedback). The

ranges of notes designated as the S+ and S- ranges during discrimination training

in Octave 4 and corresponding ranges of novel notes in Octave 5 comprised the

tests notes during generalization. Generalization of the pattern of responding

acquired during discrimination training in Octave 4 to novel notes in Octave 5

provided evidence for octave equivalence. It is important to understand that,

during both the training and generalization phases, notes were selected at random

and without replacement. The frequencies for all the musical notes in an equal-

tempered (chromatic) scale with A4 = 440 Hz, including those in the octaves used

for discrimination and generalization here and in Experiment 3, are widely

available (e.g., see Suits, 2012).

Results and Discussion

AP testing

Using Athos et al.'s (2007) criterion score for highly accurate AP ($\geq$ 24.5),

we found 1 AP possessor (AP score = 31) and 78 nonpossessors (7.5 $\pm$ .3).

The results of discrimination training and generalization testing averaged

for the four notes in each pitch range and over all trials are shown in Figures 4-3

and 4-4, respectively. Statistical analyses of the results are based on these

summaries.

Discrimination training

Not all participants learned the note-range discrimination. The criteria for acquisition of the discrimination: every learner (n = 64) but no nonlearners (n = 15) responded more on average to notes in the middle range than to notes in either the lower or upper S- range in the middle-range S+ group or the reverse in the middle-range S- group. By these criteria, 5 participants in the middle-range S+ group and 10 in the middle-range S- group failed to learn the discrimination. Here, we analyzed results for discrimination learners; later, we compared the performance of learners and nonlearners during the generalization test.

We conducted separate mixed 2 (amplitude) x 3 (note range) ANOVAs in the two reward-order groups. In the middle-range S+ group (see Figure 4-3, upper panel), we observed a significant effect of note amplitude, $F(1,26) = 18.67$, $p < .001$, $\eta p2 = .42$, a significant effect of note range, $F(2,52) = 91.13$, $p < .001$), $\eta p2 = .77$, and no significant interaction, $F(2,52) = .65$, $p = .52$, $\eta p2 > .01$. Participants in the 60 dB group responded more overall than participants in the 70/80 dB group. In planned one-tailed comparisons pooled over the two amplitude subgroups, we determined that participants responded significantly more to notes in the middle range than to notes in the lower or upper ranges, $ts(27) \geq 10.17$, $ps < .001$.

In the middle-range S- group (see Figure 4-3, lower panel), we observed

no significant effect of note amplitude, $F(1,34) = .21$, $p = .65$, $\eta p2 = .02$, a

significant effect of note range, $F(2,68) = 58.57$, $p < .001$, $\eta p2 = .60$, and a

significant interaction, $F(2,68) = 5.27$, $p = .007$, $\eta p2 = .05$. In tests of simple

effects, $p \leq .05$, we determined that the amplitude subgroups differed at a

marginal level of significance, $p = .06$, only in the lower note range, but both

amplitude subgroups responded more in the lower and upper note ranges than in

the middle-range. As the pattern of accurate discrimination was the same, we

pooled over amplitude subgroups in planned one-tailed comparisons. We

determined that the participants responded significantly less to notes in the middle

S- range than to notes in the lower S+ or upper S+ ranges, $ts(34) \geq 7.20$, $ps <$

.001. In summary, both reward-order groups showed solid evidence of

discriminating notes in the middle range from notes in the upper and lower ranges

with minimal intrusion of amplitude effects.

Generalization testing

As for discrimination training, we report results here only for

discrimination learners. The analyses are similar to those reported for

discrimination training, except that instead of comparing three note ranges, these

analyses compared six ranges: three in Octave 4 and three in Octave 5 (see Figure

4-4). We conducted separate mixed ANOVAs 2 (amplitude) x 6 (note range) in

the two reward-order groups.

In the middle-range S+ group, we observed no significant effect of note

amplitude, $F(1,26) = 1.88$, $p = .18$, $\eta p2 = .07$, a significant effect of note range,

$F(5,130) = 14.29$, $p < .001$, $\eta p2 = .36$, and no significant interaction, $F(5,130) =$

$.86$, $p = .51$, $\eta p2 = .02$. As in the analysis of discrimination training, we pooled

results across the two amplitude groups. Here, we conducted separate sets of

planned one-tailed comparisons in Octave 4, the former training octave, and

Octave 5, the novel generalization octave. In both octaves, participants responded

significantly more to notes in the middle range than to notes in either the lower or

upper ranges, $ts(27) \geq 2.92$, $ps \leq .003$.

In the middle-range S- group, we observed no significant effect of note

amplitude, $F(1,34) = .21$, $p = .65$, $\eta p2 < .01$, a significant effect of note range,

$F(5,170) = 22.09$, $p < .001$, $\eta p2 = .37$, and a significant interaction, $F(5,170) =$

2.44, $p = .036$, $\eta p2 = .04$. In tests of simple effects, $p \leq .05$, we determined that

the amplitude subgroups differed significantly in only the lower range of Octave 4

and not in any range in Octave 5. Both amplitude subgroups responded more to

notes in the lower and upper ranges than notes in the middle range in both

octaves, so we pooled results across the two groups. Again, we conducted

separate sets of planned one-tailed comparisons in Octaves 4 and 5. In both

octaves, participants responded significantly less to notes in the middle range than

to notes in the lower or upper ranges, $ts(34) \geq 3.72$, and $ps < .001$.

In summary, we have reported that participants in both reward-order

groups generalized the pattern of responses from note-range discrimination to the

note-range generalization test. Here, we determined whether the predicted

patterns of responding during generalization were observed in significant numbers of participants in each reward-order groups using binomial tests. We tested the middle S+ range group in Octave 4, 18 of 28 participants showed the predicted pattern (more responding to notes in the middle range than in either the lower or higher note ranges), chance = 1/3, p < .0001. In Octave 5, 15 of 28 participants showed the predicted pattern, chance = 1/3, p = .022. We also tested the middle S- range group in Octave 4, 25 of 36 participants showed the predicted pattern (more responding to notes in the middle range than either the lower or higher note ranges), chance = 1/3, p < .0001. In Octave 5, 24 of 36 participants showed the predicted pattern, chance = 1/3, p < .0001. Significantly, more participants than expected by chance showed the predicted pattern without feedback or reward in both octaves in both reward-order groups.

Reporting note-by-note comparisons

We have just presented the results of Experiment 2 using averages over the four-note pitch ranges. This greatly simplified statistical analyses and provided clarity in the understanding of the results. However, some readers may want a more detailed report: one that shows the results note-by-note across ranges in discrimination and generalization; this is shown in Figures 4-5 and 4-6, respectively. These figures illustrate that a finer grained presentation of the results confirms the coarser grained analysis by note ranges in Figures 4-3 and 4-4. Finer grained analyses, shown in Figures 4-5 and 4-6, found that the note ranges

were not categorical: responding gradually increased across an S+ range then declined and the reverse over S- ranges. Furthermore, generalization (see Figure 4-6) showed correspondence between note-by-note patterns of responding in Octaves 4 and 5. All that said, note-by-note analysis revealed nothing that contradicted our statistical analyses based on averages for each note range.

Comparisons with nonlearners

We asked whether learning the note-range discrimination was necessary for generalization to Octave 5. We compared the percent predicted response scores of learners and nonlearners, where percent predicted response = percent response to notes in Octave 5 that correspond to rewarded notes in Octave 4 + (100% - percent response to the notes in Octave 5 that correspond to unrewarded notes in Octave 4). Higher predicted response scores indicate better octave generalization. We compared the predicted response scores of learners (58.9% ± 1.5%), and nonlearners (51.5% ± 2.6%). Learners, t(63) = 5.99, p < .0001, but not nonlearners, t(14) = .59, p = .58, n = 15, scored significantly above chance (50%) and learners scored significantly higher than nonlearners, t(77) = 4.33, p < .0001. In summary, learners but not nonlearners were able to generalize the note-range discrimination from one octave to the next.

Experiment 3

We conducted a third experiment with two purposes in mind: to replicate the results of Experiment 2 and to extend our results with a transfer test. The logic was identical to a demonstration of positive transfer from real objects to pictures or the opposite when the contingency remained the same during transfer (Spetch & Friedman, 2006).

This additional phase introduced reward into the testing octave. For approximately half the participants, reward was chroma-matched; that is, the contingencies in the testing octave were the same as in the training octave, which we refer to as positive transfer. For the remaining participants, the reward was chroma-reversed; that is, the contingencies in the testing octave were the opposite as in the training octave, which we refer to as negative transfer. The goal was to assess whether participants would transfer more easily in the positive transfer group than the negative transfer group because of mediation by octave generalization from the original discrimination.

Method

Participants

Forty-three students at the University of Alberta participated. The details

of their written questionnaires and consents, and the conditions of their participation were the same as in Experiment 1 and 2. The Research Ethics Board at the University of Alberta approved our protocols.

Participants ranged in age from 18 to 27 years old, M = 20; 17 were men, 26 were women. We determined participants' AP status using Athos et al.'s (2007) note naming test. No participant in this experiment possessed AP.

We have provided information about the participants' music and language histories here. Twenty-three participants had no formal music training, and 10 began their training with piano. The remainder began with a variety of other instruments including voice and violin. Among musically trained participants, the amount of training varied from 1 to 13 years, M = 6 years. Fourteen played at least one additional instrument, and 7 had passed formal examinations in music. Twenty-six participants learned English as their first language, 6 learned Mandarin first, and 11 participants learned another language first. For the instruments and languages that were more common in our sample of participants from Experiments 2 and 3, we were able to perform statistical analyses to assess whether these differences impacted our results (see the personal history section of results).

Apparatus, Stimuli, and Procedures

Most aspects of the method were unchanged from Experiment 2 with the exception of the inclusion of a transfer phase at the end of the experiment.

Because both the middle-range S+ groups and the middle-range S- groups trained and tested with 60 dB (SPL) tones or with 70 and 80 dB (SPL) tones showed the same pattern of results, we conducted Experiment 3 using the middle-range S+ discrimination and played 70 and 80 dB (SPL) notes only.

Transfer testing

We conducted transfer testing with the 24 tones that comprise Octaves 4 and 5. Participants received reward in this phase depending on whether they were in the positive or negative transfer group. For both transfer groups, responses to the tones from Octave 4 were rewarded (positive feedback) following the same contingencies as training; that is, responses to E, F, F#, or G were rewarded but responses to C, C#, D, or D#, the lower tone range and G#, A, A#, or B in the upper tone range were not rewarded (no positive feedback). However, the transfer groups differed in that they had opposite contingencies in Octave 5. The positive transfer group had the same contingencies in Octave 5 as in Octave 4. The negative transfer group had the reversed contingencies in Octave 5. During transfer trials, notes were selected and played at random and without replacement, as they were in training and generalization.

Results and Discussion

AP testing

Using Athos et al.'s (2007) AP criterion score, we found only nonpossessors (7.4 ± 0.4).

Operant discrimination training

Not all participants learned the note-range discrimination. Every learner (n = 39) but no nonlearners (n = 4) responded more to notes in the middle S+ range than to notes in either the lower or upper S- range. Here, we analyze results for discrimination learners across all discrimination trials and then training block-by-block. Later, we compared the performance of learners and nonlearners during the generalization test.

We conducted an ANOVA comparing percent response across ranges and observed a significant effect of note range, see Figure 4-7, $F(2,76) = 156.58$, $p < .001$, $\eta p 2 = .80$. In planned one-tailed comparisons, we determined that participants responded significantly more to notes in the middle range than to notes in the lower or upper ranges, $ts(38) \geq 14.89$, $ps < .001$.

Discrimination acquisition

Here, we asked about the speed of acquisition of the discrimination in Octave 4. Figure 4-8 shows the course of discrimination in successive ten-trial blocks over 16 blocks of trials. The choice of ten-trial blocks was a compromise;

it ensured that at least two trials with notes in each pitch range were presented in each block and that early learning was well represented. Learning here consisted mainly of reduced responding to notes in the lower- and upper-range S- ranges with almost no change in responding to notes in the middle S+ range. Notice that extinction of responding to notes in the S- ranges was incomplete: responding remained at about the same level from the 4th trial block onward. We used the percent correct responding to assess learning: the percent correct improved significantly, $t(38) = 5.68$, $p < .001$, from $58\% \pm 1.6\%$ in the 1st trial block to $69\% \pm 1.1\%$ in the 16th block. Discrimination improved over the first 3 or 4 blocks of training, but analyses based on the first vs. last half of the training session yielded the same pattern of significant differences as the analyses we report here for the entire session. Humans who acquire the discrimination do so remarkably quickly.

Generalization testing

As for discrimination training, we report results here only for discrimination learners. The analyses are similar to those reported for discrimination training, except that instead of comparing the three note ranges in Octave 4, these analyses compared six ranges: three in Octave 4 and three in Octave 5.

We observed a significant effect of note range, $F(5,190) = 53.02$, $p < .001$, $\eta p2 = .58$ (see Figure 4-9). We then conducted separate sets of planned one-tailed

comparisons in Octave 4, the former training octave, and Octave 5, the novel

generalization octave. In both octaves, participants responded significantly more

to notes in the middle S+ range than to notes in either the lower or upper S-

ranges, $ts(38) \geq 4.08$, $ps < .001$.

We also determined whether the predicted patterns of responding during

generalization were observed in significant numbers of participants. As in

Experiment 2, we used binomial tests. In Octave 4, 37 of 39 participants showed

the predicted pattern (more responding to notes in the middle range than in either

the lower or higher note ranges), chance = ⅓, $p < .00001$, and in Octave 5, 21 of

39 participants showed the predicted pattern, chance = ⅓, $p = .0007$.

Significantly, more participants than expected by chance showed the predicted

pattern without feedback or reward in either octave.

Transfer testing

As for discrimination training and generalization testing, we report results

here only for discrimination learners. The analyses are similar to those reported

for generalization, except that we included transfer (either positive or negative

transfer) group as a between-subjects variable.

We observed a significant effect of transfer group, $F(1,37) = 6.37$, $p =$

$.016$, $\eta p2 = .17$, and note range, $F(5,185) = 10.20$, $p < .001$, $\eta p2 = .20$ (see Figure

4-10). We also found an interaction between group and range, $F(5,185) = 4.96$, p

< 0.001, $\eta p2 = .13$. Simple effects tests conducted to analyze the interaction

showed that responding in the lowest two ranges in the training octave were not significantly different between the transfer groups, $ts(37) \leq 0.19$, $ps \geq 0.854$, and in the highest range in the training octave, responding was marginally significantly higher in the negative transfer group, $t(37) = 1.99$ $p = 0.054$. In contrast, in the transfer octave, responding was significantly higher in the negative transfer group in all three ranges, $ts(37) \geq 2.86$, $ps \leq 0.007$, which suggests that participants were sensitive to the higher overall frequency of reward in the negative transfer group.

To understand how the transfer groups differed in the transfer octave, we conducted an ANOVA separately for each transfer group using the three note ranges and two octaves as within-subjects variables. In the positive transfer group, we observed a significant main effect of octave, $F(1,20) = 7.13$, $p = .015$, $\eta p2 = .21$ and of note range, $F(2,40) = 26.41$, $\eta p2 = .36$, $p < 0.001$; and no significant interaction $F(2,40) = 1.59$, $p = .217$, $\eta p2 = .07$. In the negative transfer group, we did not observe a significant main effect of octave, $F(1, 17) = 2.30$, $p = .148$, $\eta p2 = .11$, but we did observe a significant main effect of note range, $F(2,34) = 13.60$, $p < 0.001$, $\eta p2 = .12$, and no significant interaction $F(2,34) = 2.42$, $p = .104$, $\eta p2 = .12$. In other words, both the level and the pattern of responding in the transfer octave differed significantly between the positive and negative transfer groups.

We then determined whether these patterns of responding as measured by the percent predicted responding (i.e., responses that followed the pattern of responding predicted during discrimination training) were observed in significant

numbers of participants. We used a separate set of binomial tests in each group. In the positive transfer group in Octave 4, 18 of 21 participants showed the predicted pattern (more responding to notes in the middle range than in either the lower or higher note ranges), chance = 1/3, $p < .001$, and in Octave 5, 12 of 21 participants showed the predicted pattern, chance = 1/3, $p = .02$. In the negative transfer group in Octave 4, 13 of 18 participants showed the predicted pattern (more responding to notes in the middle range than in either the lower or higher note ranges), chance = 1/3, $p < .001$, and in Octave 5, only 8 of 18 participants showed that pattern, chance = 1/3, $p = .222$. Here, the most important finding is that, in the positive transfer group, more participants than expected by chance showed the expected pattern in Octave 5, whereas in the negative transfer group, the number of participants who showed the pattern predicted from original discrimination (i.e., octave generalization, or more responding to notes in the middle range than in either the lower or higher note ranges) did not differ from chance. The positive transfer group showed the predicted pattern in Octave 5; the negative transfer group showed no significant pattern. In fact, only 3 of 18 participants responded to Octave 5 with a reversed pattern (i.e., less responding to notes in the middle range than either the lower or higher note ranges) which is not significantly below chance, chance = 1/3, $p = .102$. The negative transfer group was affected by the reversed contingencies but retained enough memory of the original discrimination to retard acquisition of the reversed discrimination.

Comparisons with nonlearners

We asked whether learning the note-range discrimination was necessary to generalization to Octave 5. We compared the percent predicted response scores of learners and nonlearners, where percent predicted response = percent response to notes in Octave 5 that correspond to rewarded notes in Octave 4 + (100% - percent response to the notes in Octave 5 that correspond to unrewarded notes in Octave 4). Higher predicted response scores indicate better octave generalization. We compared the predicted response scores of learners (57% ± 1%), and nonlearners (47% ± 3%). Learners, $t(38) = 6.16$, $p < .0001$, but not nonlearners, $t(3) = 0.90$, $p = .43$, $n = 4$, scored significantly above chance (50%),and learners scored significantly higher than nonlearners, $t(41) = 2.64$, $p = .012$ in Octave 5. In summary, learners showed greater resistance to the change in the pattern of reinforcement in the negative transfer group.

Personal history

Previous research suggested that octave equivalence is enhanced in experienced musicians (Allen, 1967) and that absolute pitch (Deutsch et al, 2009) and relative pitch (Hove, Sutherland, & Krumhansl, 2010) are enhanced in people who first learned a tonal language. Here, we asked whether the participants' previous music or language histories contributed to their performance during original discrimination in Octave 4 (measure 1), during the generalization tests in Octave 4 and 5 (measures 2 and 3), or during the transfer tests in Octaves 4 and 5

136

(measures 4 and 5), as measured by the percent predicted responding (i.e., responses that followed the pattern of responding predicted during discrimination training). We term these measures, considered as a group, the octave equivalence performance quintet.

To assess claims about music and language experience, we pooled history questionnaire results from learners and nonlearners in Experiments 2 and 3 to obtain a sizable sample, n = 121 and to include the full range of variability in quintet scores. We excluded results for the sole AP possessor, as his performance affected correlations and means out of proportion to his role. Results for transfer were available only from Experiment 3, n = 43.

Our main tool for evaluating the contributions of music and language histories to octave equivalence was a correlational analysis. Measures of musical experience were internally consistent: years of music training and absolute pitch scores (within the range of AP nonpossessors), $r(119) = .22$, $p = .015$, and years of training and number of instruments played, $r(119) = .70$, $p < .0001$, were significantly correlated.

The measures in the octave equivalence performance quintet were internally consistent: performance during generalization in Octaves 4 and 5 and transfer in Octave 4 correlated significantly with performance during original discrimination training, $rs(119) \geq .34$ $ps \leq .00014$, and performance during transfer of training in Octave 5 correlated significantly with performance during generalization in Octave 5, $r(41) = .42$, $p =. 005$.

Early learning of a tone language, e.g., Mandarin, or a keyboard instrument, e.g. piano, appears to contribute to absolute and relative pitch perception (Deutsch et al., 2009, Hove et al., 2010). But first learning of Cantonese or Mandarin (tonal), or Korean (pitch-accented) languages did not result in significantly higher predicted discrimination or generalization responses than first language learning of English, $Fs(3,98) \leq 1.06$, $ps \geq .37$, $\eta p2 = .07$. Although some participants learned three or four languages, the number of languages did not correlate significantly with performance in the quintet, $rs(119) \leq .08$, $ps \geq .88$ and $r(41) = .13$, $p = .42$. Likewise, having first learned to play the piano did not result in significantly higher performance during discrimination or generalization than first learning violin, voice, or guitar, $Fs(3,18) \leq .43$, $ps \geq .73$, $\eta p2 = .02$. The current results do not provide evidence that participants' first language or first instrument experience affect octave equivalence. However, given that Experiments 2 or 3 were not designed to provide such evidence, any final conclusion regarding this issue should await further research.

Most importantly, measures in the performance quintet correlated significantly with measures of musical experience. We observed significant negative correlations between AP scores (all within the AP nonpossessor range), years of musical experience and performance during discrimination, and generalization in Octaves 4 and 5, $r(119) \geq -.22$, $p \leq .015$. In contrast, we observed a significant positive correlation between years of musical experience and performance during transfer (pooled for the positive and negative transfer

groups) in Octave 5, r(41) = .44, p = .003. Experiments 2 and 3 were not designed to provide tests of between group differences for participants with and without music training. That said, all the untrained participants learned the discrimination in Octave 4 and all but one generalized the discrimination to Octave 5.

General Discussion

In one sense, it is unnecessary to ask whether humans perceive the similarity between notes spaced an octave apart because we already know the answer. From the earliest times, human cultures have recognized that the octave is formed at a ratio of 2:1 between notes and used that fact to tune their musical instruments (Crickmore, 2003). Later, after written languages developed, diverse cultures including those of India, Babylon, and Egypt provided lasting records of how they used the octave to tune instruments, write music, and conceptualize mathematics (McClain, 1978).

The questions for an experimental scientist are: how to measure octave equivalence and under what conditions humans and other animals can be seen to perceive it. We conducted three experiments to study octave equivalence in humans. Each followed a similar path, using discrimination and generalization procedures, to search for octave generalization as a measure of octave equivalence.

139

Negative Findings in Experiment 1

In Experiment 1, humans learned a discrimination between an S+ and an S- tone with high accuracy. But during generalization testing, their responses to a tone spaced exactly an octave from S+ fit on the function predicted by pitch height generalization to other tones in the test, without recourse to octave generalization. The task proved exceptionally resistant to the perception of octave equivalence. Neither extensive music training nor absolute pitch perception improved participants' octave generalization.

Experiment 1's failure to observe octave equivalence does not mean its findings are unimportant. At the very least, the experiment showed that Cynx's (1993) procedure, which demonstrated the dominance of pitch height over octave equivalence perception in starlings, yields exactly the same result in humans – a finding consistent with previous research with humans (e.g., Allen, 1967; Kallman, 1982; Sergeant, 1983) and inconsistent with the conclusion that Cynx's (1993) experiment demonstrated that starlings fail to perceive octave equivalence. More importantly, the findings of Experiment 1 lead us to seek effective modifications to Cynx's (1993) procedures. Clearly, we needed to look elsewhere to find a simple procedure for studying octave generalization successfully, which we did in Experiments 2 and 3.

Octave Generalization and Transfer in Experiments 2 and 3

The procedures of Experiments 2 and 3 differed from those of Experiment

1 in several ways. (a) We tuned our stimulus tones to notes on the chromatic scale, which meant they increased from lowest to highest on a log-linear scale familiar to humans tutored in music. (b) During discrimination training, we divided the training octave into three ranges of four notes each and provided the same feedback and reward for responding to all four notes in each range. That is, we required much less accurate chroma identification than if a separate response was required for each of the 12 notes. (c) To reduce the influence of pitch height perception, octave generalization testing and explicit transfer of training were conducted in an adjacent octave beyond the pitch height ranges of the training notes.

The procedure used here was adapted from one we have used extensively to test for pitch height perception across songbird species (Weisman et al., 2010b). In most of these experiments, sinewave tonal stimuli increased in frequency on a simple linear scale and thus, were mistuned to human musical scales. Songbirds consistently discriminated between ranges more accurately than either rats or humans (Weisman, Njegovan, Williams, Cohen, & Sturdy, 2004). An important feature of this procedure was that rewarded and unrewarded ranges of tones alternated across frequencies, with at least three ranges (e.g., S-, S+, S-) and sometimes eight ranges (e.g., S-, S+, S-, S+, S-, S+, S-, S+) presented during discrimination training.

In Experiments 2 and 3, using similar procedures but with stimuli tuned to the musical scale, we observed strong evidence of octave generalization in the

141

next highest octave. In Experiment 2, we observed generalization whether the middle range was S+ or S-, eliminating possible artifacts introduced by either condition. However, octave generalization was influenced by whether notes in the middle range were S+s or S-s because discrimination in the middle range S+ condition depended mainly on inhibition of responding to notes in the higher and lower ranges, whereas discrimination in the middle range S- condition depended mainly on excitation of responding to notes in the higher and lower ranges.

In Experiment 3, the transfer paradigm contributed important information about the durability of octave generalization. The discrimination learned during original training not only generalized, but also octave equivalence promoted positive transfer to a like discrimination and persistent negative transfer to the opposite discrimination in the next octave. These experiments promote generalization and transfer as powerful tools for evaluation of octave equivalence. Future research should explore the retention of simple patterns of note-range discriminations over much longer periods as continuing to probe their durability in transfer tests.

Failure to Learn the Note-Range Discrimination

Several participants failed to learn original note-range discriminations. More of these participants had difficulty learning the middle range S- than middle range S+ discrimination, especially when the stimuli were played at lower amplitude. We are uncertain why the middle S- discrimination was more difficult,

but it seems likely that more participants had trouble resolving the pitches out of the lower amplitude notes. Of course, young adult human participants present more general issues because of low perceived reward, inattention, minor illness, all of which tend to interfere with accurate discrimination. It is possible that increased difficulty of the discriminations in Experiments 2 and 3 contributed to attention to the octave during generalization.

Participants who failed to learn the initial discrimination provided an interesting control. Participants who accurately discriminated the middle range in Octave 4 showed a similar pattern of responding in Octave 5. Participants who failed to learn did not show any consistent pattern of responding across pitch ranges in Octave 5. This finding helped convince us that our observations of octave equivalence in successful learners were products of the generalization of successful discrimination training – not flukes or artifacts but a solid palpable phenomenon.

Tests for artifacts of the interaction of pitch and loudness

In Experiment 2, we played the notes at different amplitudes between groups as a check for aural harmonics (Newman et al., 1937). If the participants heard loud harmonics during discrimination training, in effect they may have heard the generalization stimuli during training and would then be expected show more octave generalization than participants who heard quieter harmonics. We found no evidence of confounding by aural harmonics, i.e., we observed about the

same pattern of generalization in Octave 5 whether the notes were relatively quiet, 60 dB, or relatively louder, 70 and 80 dB. We are uncertain why the louder aural harmonics of 70 and 80 dB notes had no more effect than the much quieter harmonics of 60 dB notes. One possibility is that introducing harmonics simultaneously with stimulus notes may have muted effects when testing involves notes presented successively, as it did here.

Presenting tones that include harmonics in octave equivalence research would make the stimuli more realistic but would also introduce potential confounds (see Burns, 1999), because the octave would be present in every training stimulus as has been suggested to have occurred in Blackwell and Schlosberg's (1943) report of octave equivalence in rats using sinewave tones of uncertain purity (Burns, 1999). Another source of confounding of pitch and loudness arises from the interaction of the two, such that louder sounds can be perceived as higher in pitch (Moore, 1989). We removed the correlation between pitch and loudness by presenting each pitch at two distinct amplitudes. Generally in this research, as shown in the note-by-note presentation shown in Figures 4-5 and 4-6, the percent response was an orderly function of pitch and its correlation with reward.

Pitch Perception Is Not Unitary

When searching for evidence of octave equivalence perception, one needs to bear in mind that pitch perception is not a unitary ability, but is instead a suite

of perceptual abilities. Of course, the suite includes the perception of octave equivalence, but it also includes pitch height and relative pitch perception. To complicate matters further, these abilities function simultaneously and without reference to one another. For example, MacDougall-Shackleton and Hulse (1996) found that in starlings, pitch height perception obscured relative pitch perception within the range of the training tone sequences, but not outside of that range. Similarly, when we tested humans in the range between the S+ and S- tones in Experiment 1, pitch height perception blocked octave equivalence. Experiments 2 and 3 found blended octave equivalence and pitch height perception as shown by a decreased overall level of responding in Octave 5 compared to the training octave, Octave 4. Reducing the influence of one pitch ability on another is good science, but in practice completely eliminating the influence of pitch height perception on octave equivalence is probably impossible.

Octave Equivalence, Music and Language Training

Is it genes, environment, or experience that determine whether humans perceive octave equivalence? Our favorite answer to this question is "yes". Most sounds in the environment include harmonics and approximately half the harmonics heard in these sounds are at distances of an octave apart (Pierce, 1999). So, the ability to resolve out the octave in sound and sequences of sounds might be a useful skill. Octave equivalence affects both speech (Peter, Stoel-Gammon, & Kim, 2008) and music (Burns, 1999), and octaves are used in this way in all

145

cultures (Crickmore, 2003). Consider an example, sex recognition and sexual signals are basic skills common across species. In humans, the fundamental frequencies of male and female voices are roughly an octave apart (Titze, 2000) and thus, octave equivalence might be especially useful to humans comparing speech or song produced by a man to that produced by a woman. Untangling this knot of determinants may be impossible.

Allen (1967) found differences in octave equivalence in favor of participants with musical training. He tested with an equivalence-rating task that probably used trained musicians' extensive experience in making subjective octave judgments. In contrast, Kallman (1982) found evidence of octave equivalence in less trained individuals only when he reduced the influence of pitch height by testing over only a small range of pitches. Our most prominent example of the confluence of music ability and octave equivalence must be the single AP possessor in Experiment 2. Her ability to perceive octave equivalence in the generalization test exceeded 90%, so accurate that to make sense of the rest of the results, we needed to exclude her data. Pitch height perception can obscure the perception of octave equivalence but absolute pitch perception appears to amplify equivalence, though one might wish for more extensive confirming evidence. Also, even if correct, our finding about absolute pitch does not untangle training from the genes–the question as to which is more important to absolute pitch perception is still not settled.

In Experiments 2 and 3, more extensive music training hindered

acquisition of the note-range discrimination and reduced octave generalization. These effects may be the result of negative transfer from the skills acquired in extensive music training. Then, during the transfer phase of Experiment 3, more extensive music training made octave equivalence a positive factor in acquiring the same discrimination and a negative factor in acquiring the opposite discrimination in a higher octave. In other words, our results were determined by the complex interaction of music training with our task. We are reasonably confident in our results for correlations between acquisition and generalization of equivalence with music training, as they were obtained with a sample of 121 participants; even so the influence of music training here deserves further study.

All this said, the dependence of octave equivalence on training in music might be limited, as rhesus monkeys can be induced to show octave generalization across musical passages (Wright, Rivera, Hulse, Shyan, & Neiworth, 2000; Wright, 2007). As we have shown here, humans show octave generalization with minimal instructions, which included no reference to music or octaves, and in a task that presented notes individually and at random with respect to their height and chroma. We conclude that music training may sometimes enhance octave equivalence in humans but appears unnecessary to the basic perception of equivalence.

The relationship between music and language is receiving increased attention (e.g., Fitch, 2005; Patel, 2003; Masataka, 2009). Researchers studying pitch perception, but not octave equivalence, have found effects of tonal language

147

learning on relative and absolute pitch (Deutsch et al., 2009; Hove et al., 2010).

Given that relatively few of our participants initially spoke tonal languages and

the observed effect was not significant, one should not be surprised that we can

reach no conclusion about the influence of initial language on octave equivalence.

Research explicitly designed to distinguish among language groups would be

helpful for determining the relationship between language and octave

equivalence.

Conclusion

Cynx (1993), whose procedure we replicated in Experiment 1, decided

that his failure to observe octave generalization in starlings meant starlings lacked

pitch chroma perception. We propose an alternative interpretation: Cynx's (1993)

procedure is a measure of pitch height rather than chroma perception in both

songbirds and humans. This finding offers a useful lesson for comparative

psychologists: despite their seeming reasonableness, no sound conclusions flow

from comparisons between experimental evidence about one species and

historical, personal, or anecdotal evidence about another. With the results of

Experiments 2 and 3, we have solid evidence for octave equivalence in humans,

whether tutored in music or not. We are ready now to tackle the question Cynx

(1993) posed nearly 20 years ago: do songbirds perceive pitch chroma and

therefore, show octave equivalence? We also have more questions about human

octave equivalence. For example, does how we parse the octave during pitch

148

range discrimination affect the perception of octave equivalence? And we have learned that octave equivalence requires neither musical stimuli nor musical training, but instead is a common feature of human auditory perception.

$$y = 738.679 * 10^{-0.991x} \quad r = 0.962$$

○ Other Tones
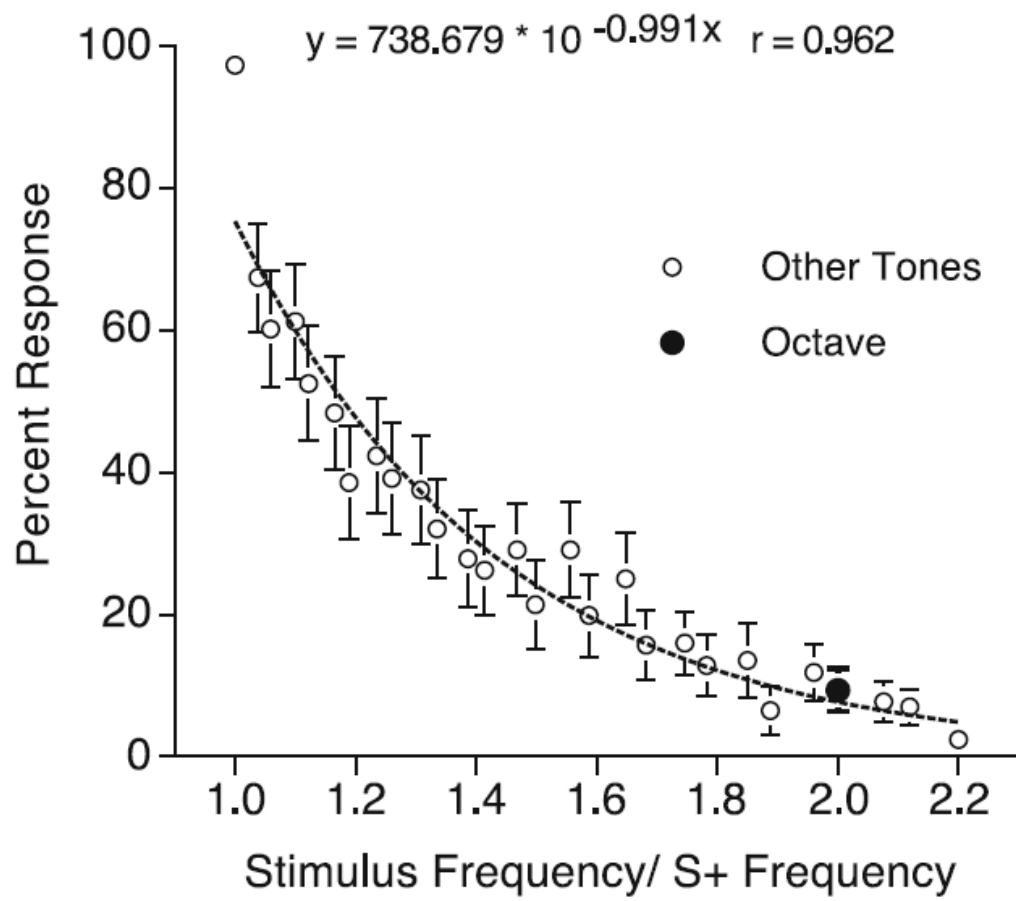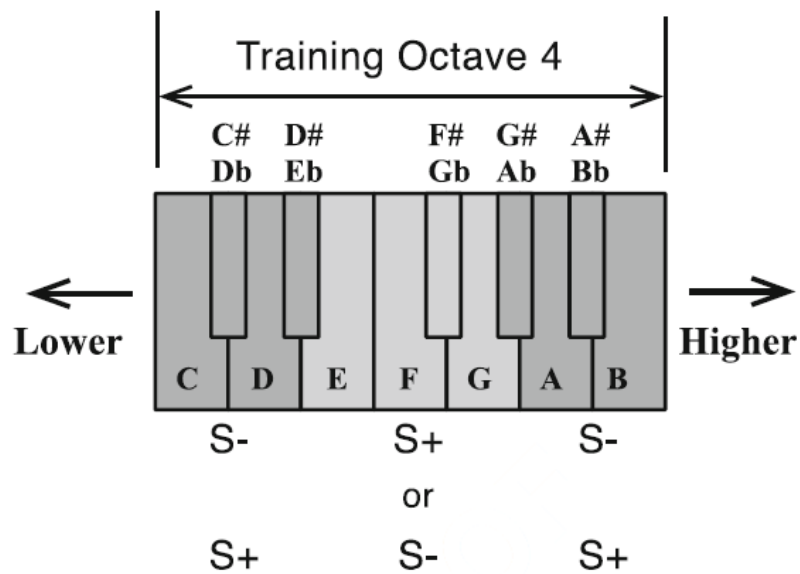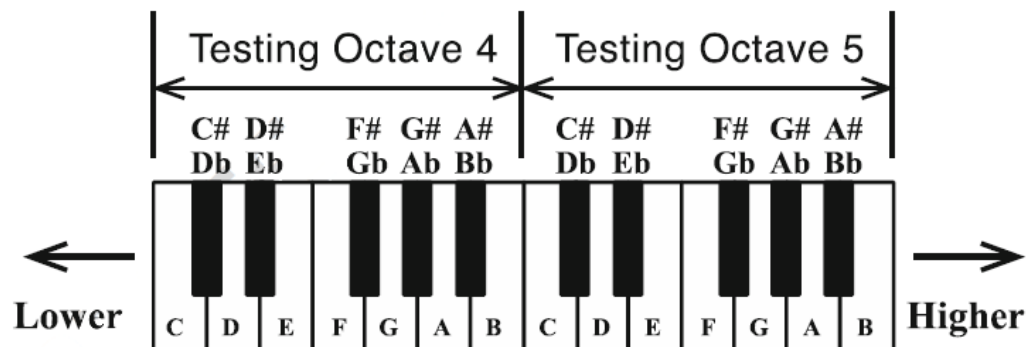
● Octave
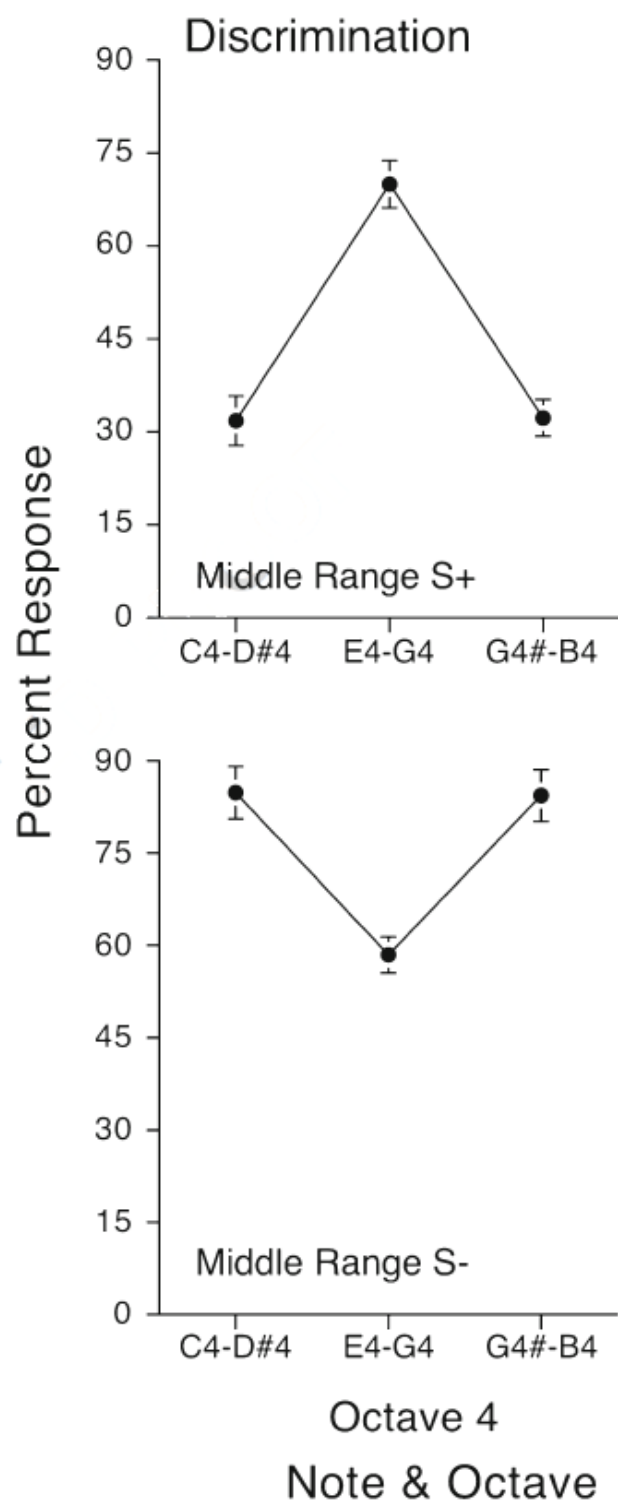
Figure 4-1.

Figure 4-2.

Figure 4-3.
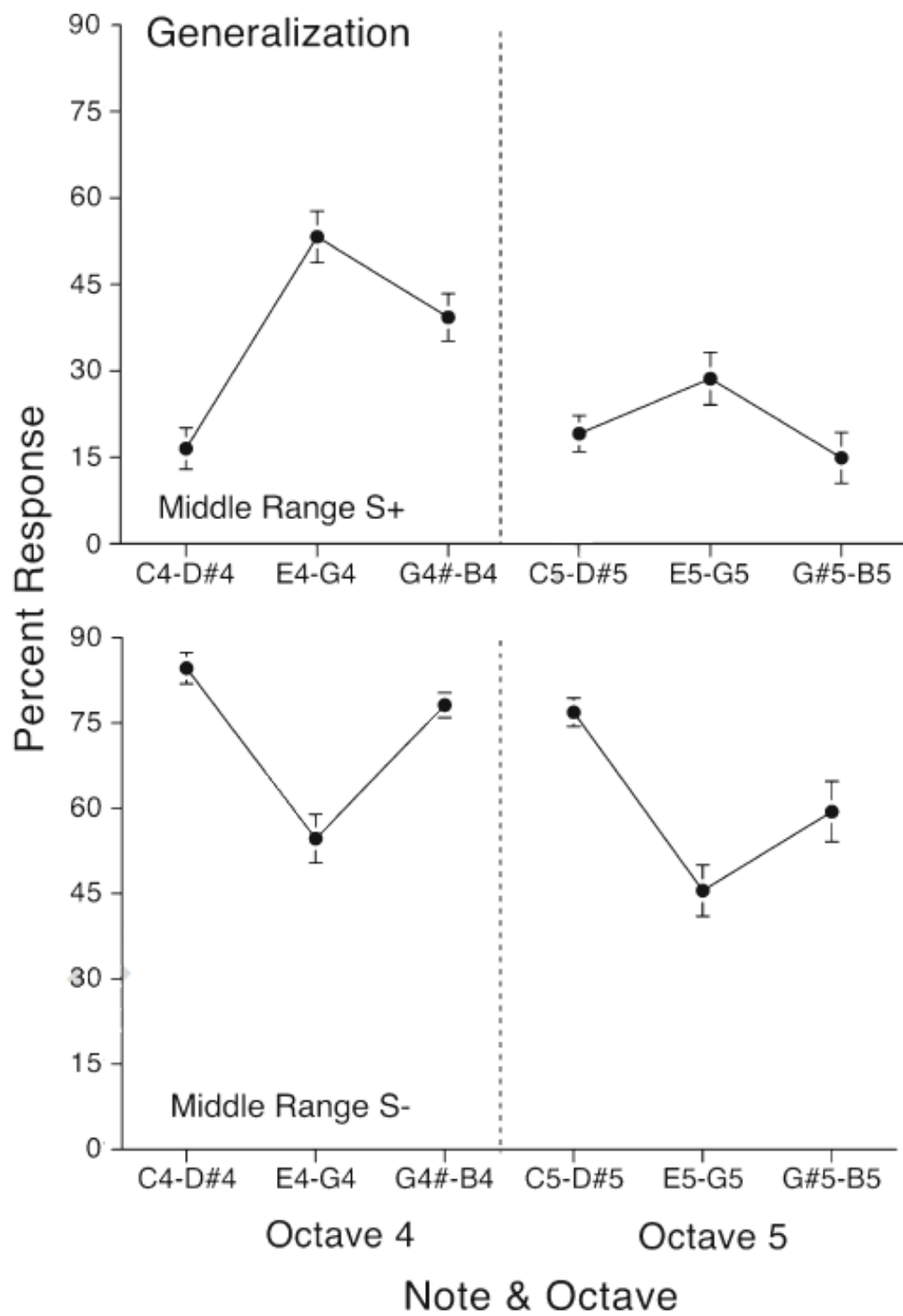
Figure 4-4.
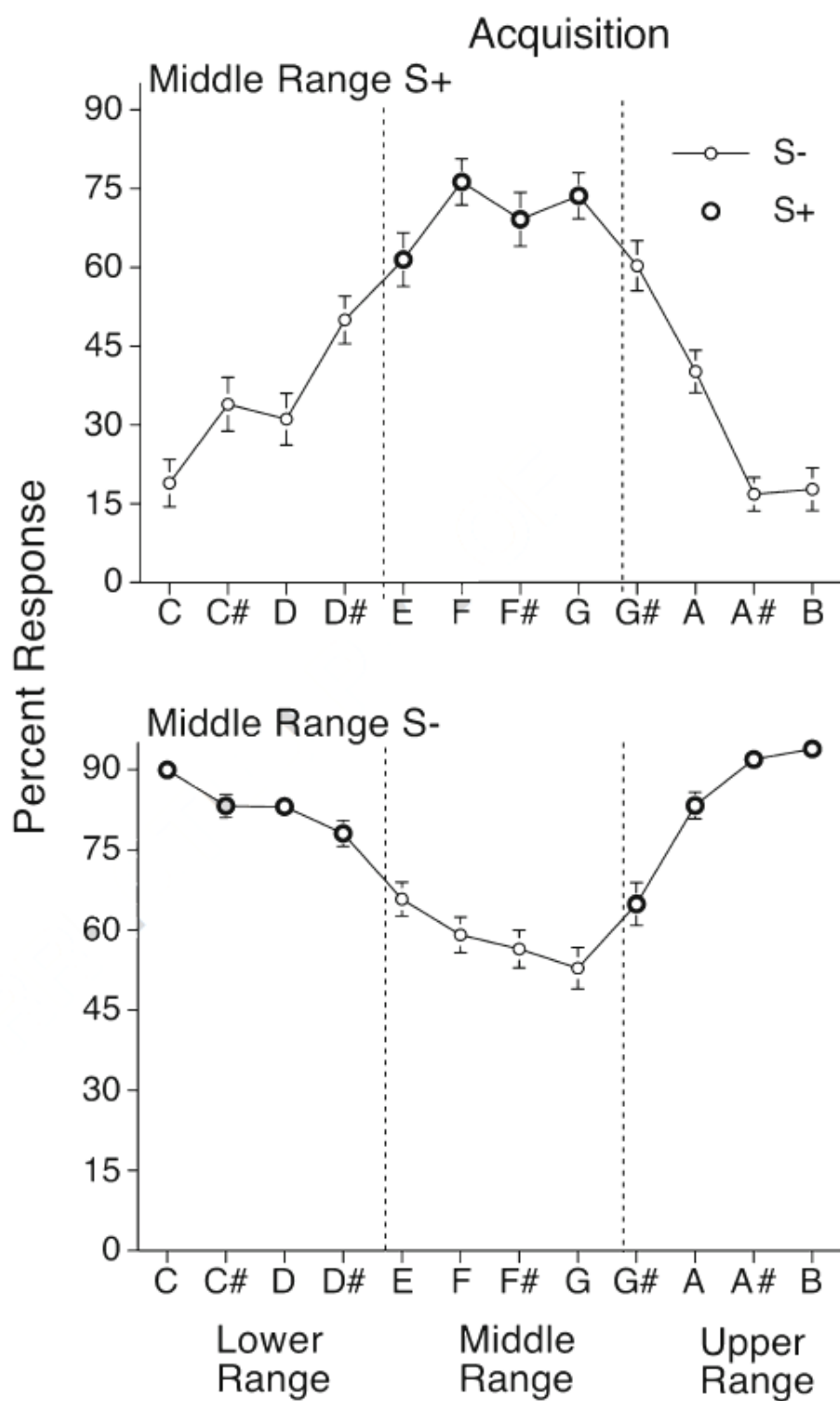
Figure 4-5.
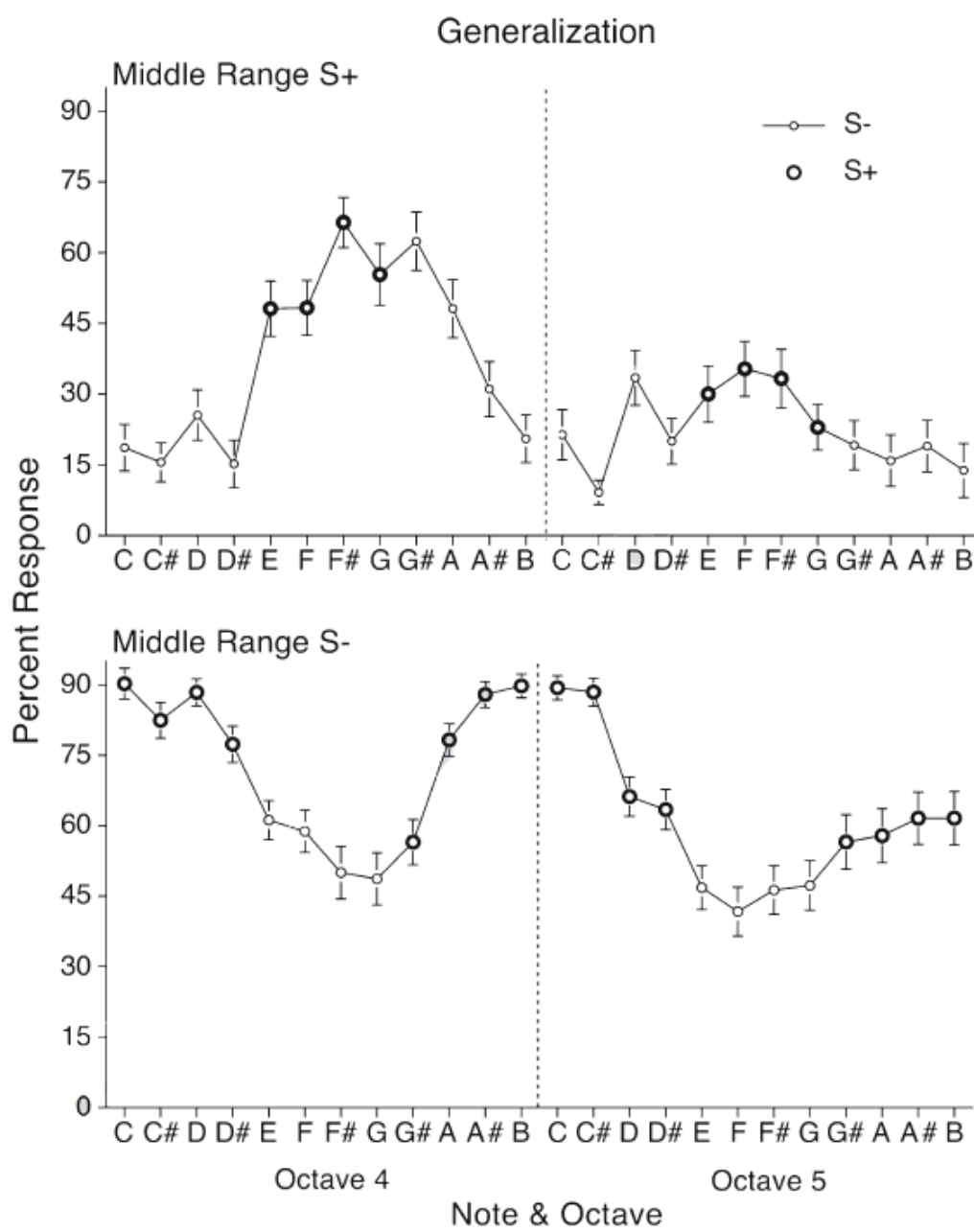
Figure 4-6.
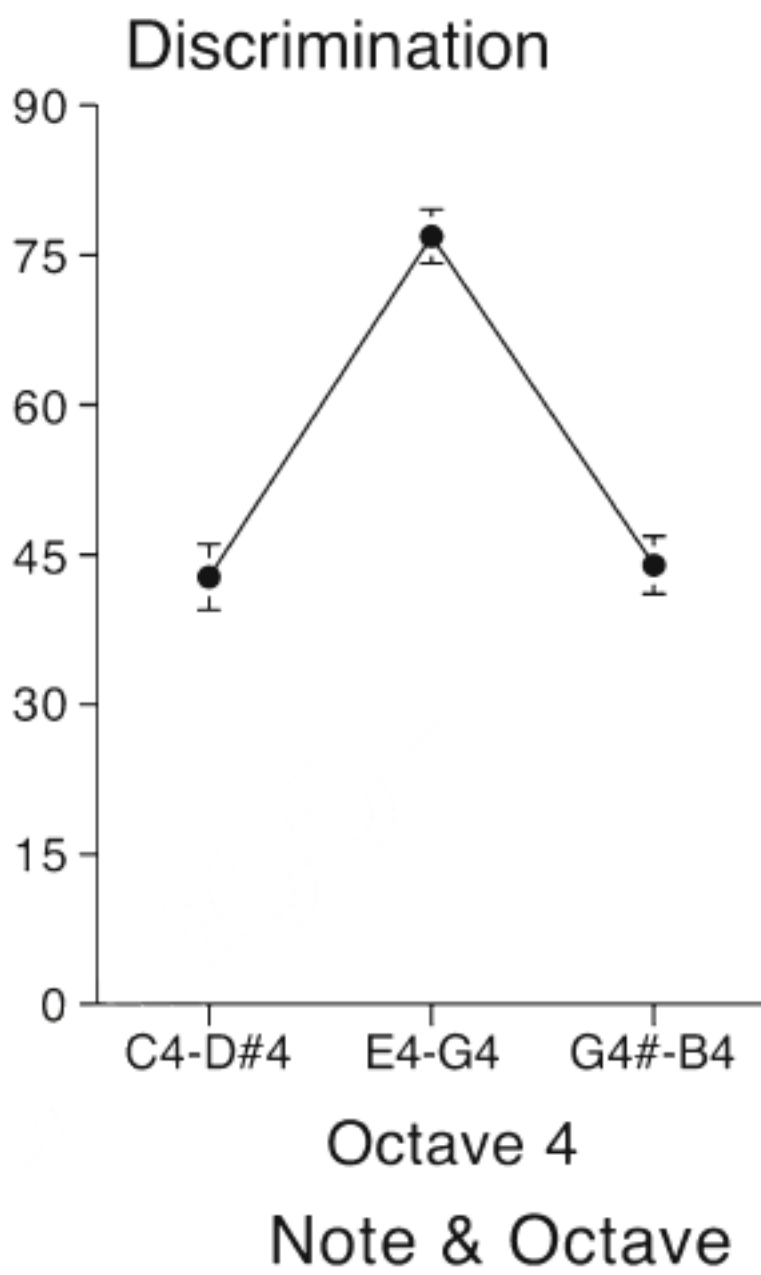
Figure 4-7.
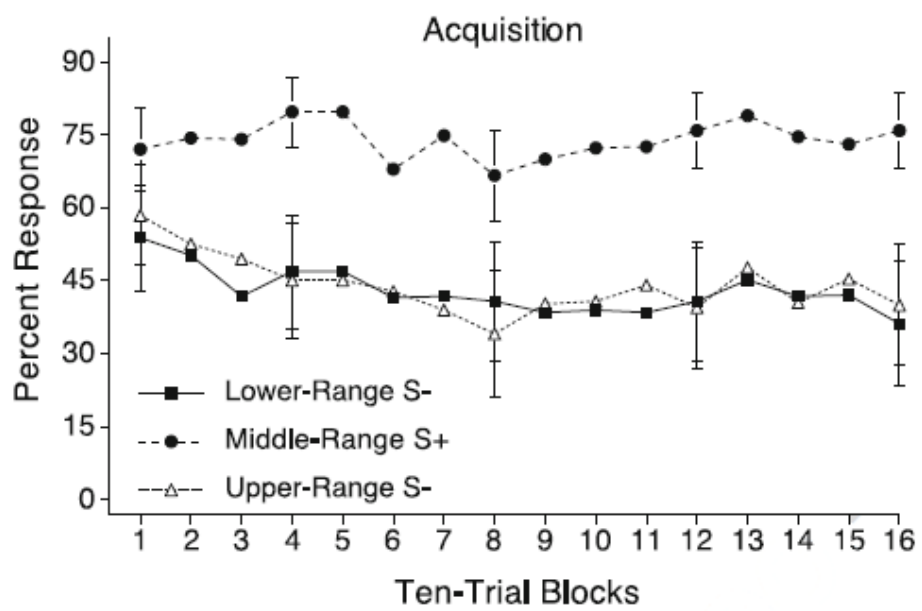
Figure 4-8.
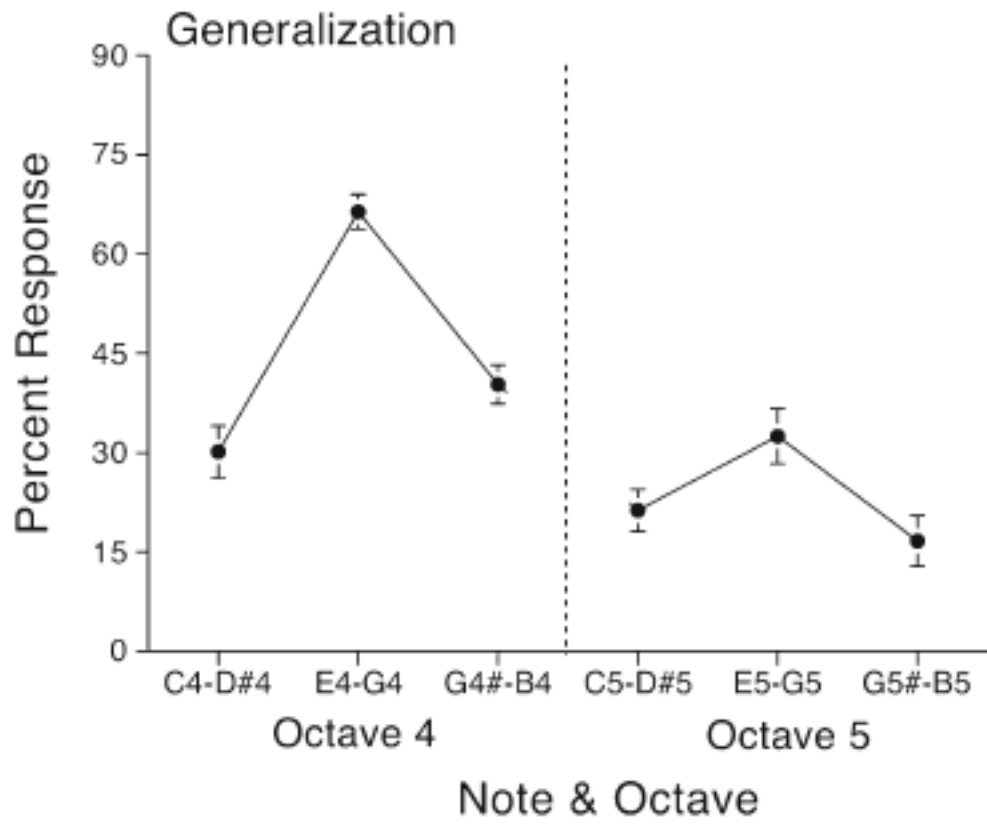
Figure 4-9.
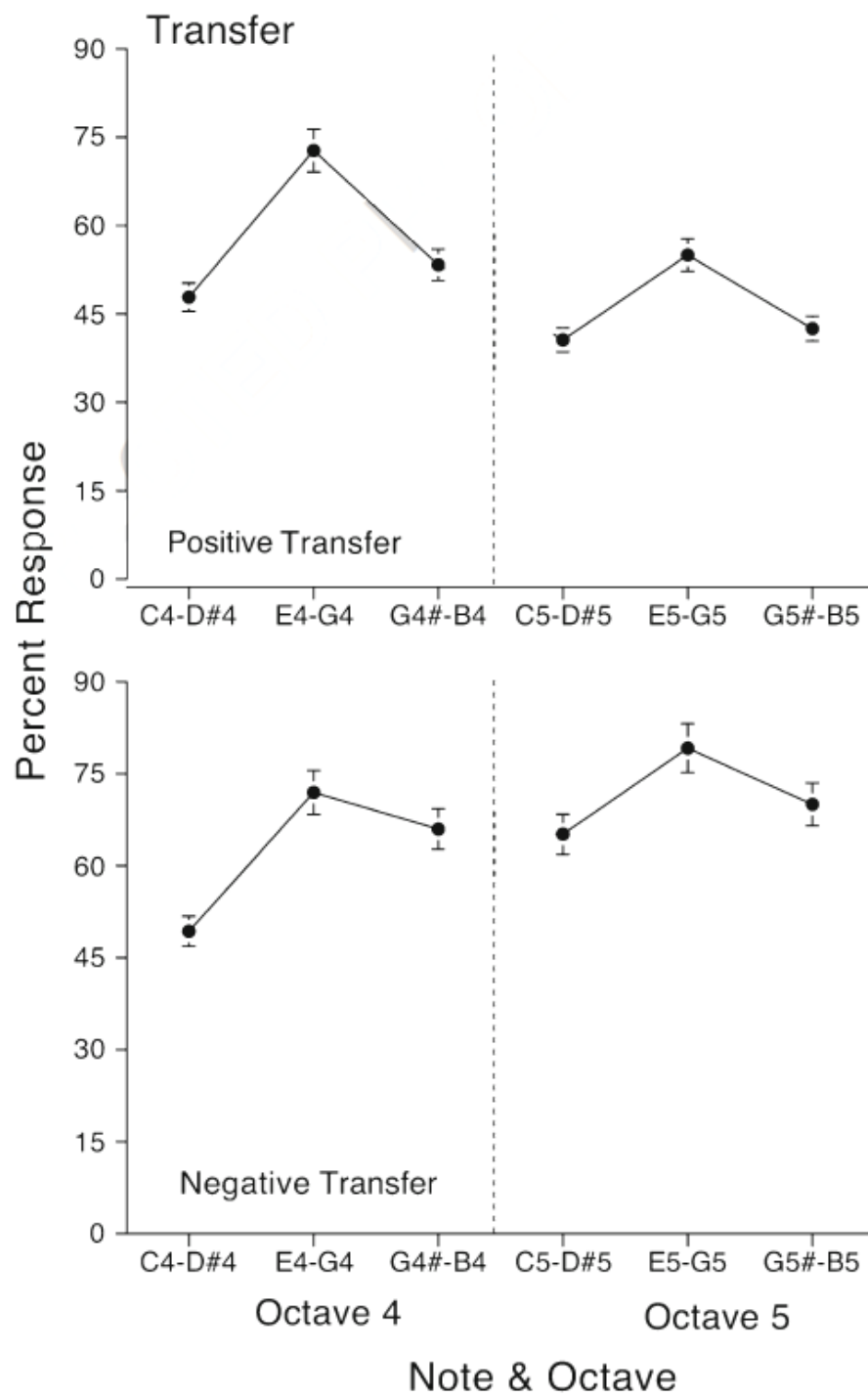
Figure 4-10.

References

Allen, D. (1967). Octave discriminability of musical and non-musical subjects. *Psychonomic Science, 7,* 421-422.

Athos, E. A., Levinson, B., Kistler, A., Zemansky, J., Bostrom, A., Freimer, N., & Gitschier, J.(2007). Dichotomy and perceptual distortions in absolute pitch ability. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 14795-14800.

Blackwell, H. R., & Schlosberg, H. (1943). Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology, 33*, 407-419.

Burns, E. M. (1999) Tuning. In (Ed) Deutsch, D. *The psychology of music.* San Diego, CA: Academic Press.

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

Cynx, J. (1993). Auditory frequency generalization and a failure to find octave generalization in a songbird, the European starling (*Sturnus vulgaris*). *Journal of Comparative Psychology,107*, 140-146.

Demany, L., & Armand, F. (1984). The perceptual reality of tone chroma in early infancy. *Journal of the Acoustical Society of America, 83,* 687-695.

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception &. Psychophysics, 11*, 411-412.

Deutsch, D., Dooley, D., Henthorn, T., & Head, B. (2009) Absolute pitch among
students in an American music conservatory: Association with tone
language fluency. *Journal of the Acoustical Society of America, 125*,
2398-2409.

Dowling, W. J., & Hollombe, A.W. (1977). The perception of melodies distorted
by splitting into several octaves: Effects of increasing proximity and
melodic contour. *Perception &. Psychophysics, 21*, 60-64.

Fitch, W. T. (2005). The evolution of music in comparative perspective. *Annals
of the New York Academy of Sciences, 1060*, 29-49.

Hove, M. J., Sutherland, M. E., & Krumhansl, C. L. (2010). Ethnicity effects in
relative pitch. *Psychonomic Bulletin & Review, 17*, 310-316.

Kallman, H. J. (1982). Octave equivalence as measured by similarity ratings.
Perception & Psychophysics, 32, 37-49.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of
tonal functions within a diatonic context. *Journal of Experimental
Psychology, 5*, 579-594.

Langner, G., & Ochse, M. (2006). The neural basis of pitch and harmony in the
auditory system. *Musicae Scientiae, 10,* 185-208.

MacDougall-Shackleton, S. A., & Hulse. S. H. (1996). Concurrent absolute and
relative pitch processing by European starlings (*Sturnus vulgaris*). *Journal
of Comparative Psychology, 110*, 139-146.

Masataka, N. (2009). The origins of language and the evolution of music: A

comparative perspective. *Physics of Life Reviews, 6*, 11-22.

McClain, E. G. (1978). *The Pythagorean Plato: Prelude to the Song Itself.* York
    Beach, ME: Nicolas-Hays.

Moore, B. C. J. (1989). *An introduction to the psychology of hearing (3rd ed.).*
    London: Academic Press.

Moore, G. (2006). Lecture 15: Aural Harmonics.
    www.physics.mcgill.ca/~guymoore/ph224/notes/lecture15.pdf.

Newman, E. B., Stevens, S. S., & Davis, H. (1937). Factors in the production of
    aural harmonics and combination tones. *Journal of the Acoustical Society
    of America, 9,* 107-118.

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

Peter, B., Stoel-Gammon, C., & Kim, D. (2008). Octave equivalence as an aspect
    of stimulus response similarity during nonword and sentence imitations in
    young children. *Speech Prosody 2008*, 731-734.

Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of
    musical pitch in tone language speakers. *Attention, Perception, &
    Psychophysics, 71*, 1385-1398.

Pierce, J. R. (1999). The nature of musical sound. In (Ed) Deutsch, D. *The
    psychology of music.* San Diego, CA: Academic Press.

Sergeant, D. (1983). The octave: percept or concept. *Psychology of Music, 11*, 2-
    18.

Shepard, R. N. (1982). Geometrical approximations to the structure of musical

      pitch. *Psychological Review, 89*, 305–333.

Spetch, M. L., & Friedman, A. (2006). Pigeons see correspondence between

      objects and pictures. *Psychological Science, 17*, 966-972.

Suits, B. H. (2012) The Physics of Music-Notes. Michigan Technological

      University. Retrieved February 11, 2012 from

      http://www.phy.mtu.edu/~suits/scales.html

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

      for Voice and Speech.

Ward, W. D., & Burns, E. M. (1982). Absolute pitch. In (Ed) Deutsch, D. *The*

      psychology of music. San Diego, CA: Academic Press.

Weisman, R. G., Balkwill, L. L., Hoeschele, M., Moscicki, M. K., Bloomfield, L.

      L., & Sturdy, C. B. (2010a). Absolute pitch in boreal chickadees and

      humans: exceptions that test a phylogenetic rule. *Learning and*

      Motivation, 41, 156-173.

Weisman, R. G., Hoeschele, M., Bloomfield, L. L., Mewhort, D., Sturdy, C. B.

      (2010b). Using network models of absolute pitch to compare frequency-

      range discriminations across avian species. *Behavioural Processes, 84*,

      421-427.

Weisman, R., Njegovan, M., Sturdy, C., Phillmore, L, Coyle, J., Mewhort, D.

      (1998). Frequency range discriminations: Special and general abilities in

      zebra finches. *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. *Behavioural Processes, 66,* 289-307.

Wright, A. A. (2007). An experimental analysis of memory processing. *Journal of the Experimental Analysis of Behavior, 88*, 405-433.

Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology: 805 General, 129*, 291-307.

Chapter 5 Chickadees fail human-standardized tests for pitch chroma discrimination, generalization, and transfer[4]

---

[4] A version of this chapter is currently under preparation to be submitted to *Animal Cognition.* Authors: Hoeschele, M., Weisman, R.G., Guillette, L.M., Hahn, A.H., & Sturdy, C.B.

Introduction

An octave is defined as the distance in frequency between two acoustic events that differ by double or half. The octave is the only pitch interval used in music by all human cultures and is therefore a fundamental component of music (Crickmore, 2003). Not only is the octave used across cultures, but notes separated by an octave are perceived to have a similar pitch quality, a quality known as "chroma". This cross-cultural perceptual similarity of notes separated by an octave is known as octave equivalence. Importantly, octave equivalence occurs across cultures even though the absolute frequencies of notes, their labels, and the number of notes within one octave vary across cultures (Cartelle & Kendall, 1999). Based around the logarithmic psychophysics of sound, the octave is thus a unique example of Weber's law at work in human music perception and a highly important referent in discussion of human auditory perception.

In fact, when assessing the absolute pitch of a sound, octave equivalence, or pitch chroma, is one of two main determinants. The other determinant is pitch height, which is simply a log-linear scale of frequency where nearby frequencies are perceived as more similar than distant frequencies. It is important to note that these two determinants of pitch perception are often in conflict. To a listener attending to pitch chroma alone, pitches with a frequency relationship of 1:2 are perceived as more similar than those that share a relationship of 2:3, whereas for a listener attending to pitch height the opposite would be true.

In humans, not only is the octave used in music (Burns, 1999), the octave

166

relationship is used when imitating speech (Peter, Stoel-Gammon, & Kim, 2008) and song (Levitin, 1994). That is, people will transpose a set of pitches by an octave to a more comfortable range when reproducing them in both speech and song. Because of these links between music and language, it has led some researchers to propose that many aspects of musical perception may have biological, rather than purely cultural, origins (e.g., Fitch, 2005; Hauser & McDermott, 2003; Masataka, 2009). In fact, octaves occur in many natural sounds that contain harmonics. Harmonics are integer multiples of the fundamental frequency of a sound. Since octaves are a doubling in frequency, the first harmonic of any sound, being two times the fundamental, is also the first octave from the fundamental (Pierce, 1999). This relationship makes it reasonable to suppose that animals may come to associate pitches with their octaves because of how commonly they occur together.

However, research on non-human animal octave perception has been sparse. In 1943, Blackswell and Schlosberg's work suggested that rats readily perceived octaves, but these studies were later criticized for not controlling for harmonics. Blackswell and Schlosberg trained rats with a 10 kHz tone, and tested them with a 5 kHz tone (1943). The first harmonic of the 5 kHz tone would have produced a 10 kHz harmonic in the test stimuli, equivalent to the octave of the 5 kHz tone. Burns (1999) notes that this study was never successfully replicated. There have been several studies that suggest dolphins may have octave equivalence; however, because dolphins are difficult to study in large groups,

167

these studies usually only have one subject per experimental group (e.g.,Ralston & Herman, 1995; Richards, Woltz, & Herman, 1984). Ralston and Herman (1995) found that a dolphin was able to transfer a perceptual discrimination to a novel octave. Richards, Woltz, and Herman (1984) found that a dolphin, which was trained to imitate vocalizations, would octave transpose the imitations into a preferred octave. Wright, Rivera, Hulse, Shyan, and Neiworth (2000) found that rhesus monkeys were more likely to classify two melodies as being the "same" if they were 1 or 2 octaves apart compared to being 0.5 or 1.5 octaves apart. As with the dolphin studies, however, Wright et al.'s (2000) study with rhesus monkeys only had two subjects, so it is difficult to know whether these results are generalizable. As far as avian species go, to date there is no evidence of octave equivalence in any species. The only study that we are aware of that looked at octave equivalence in birds is Cynx's (1993) work. Cynx found that starlings trained to respond to a 1000 Hz tone and not a 2200 Hz tone (or vice versa) would not respond to the octave of the S+ and concluded birds did not have octave equivalence. However, he did not control for the influence of pitch height (see Hoeschele, Weisman, & Sturdy, in press).

Although the importance of octaves has been known for quite some time because of its use in music, developing a perceptual task of octave perception even in humans has been difficult (e.g. see Allen, 1967; Kallman, 1982; Krumhansl & Shepard, 1979; Deutsch, 1972; Dowling & Hollombe, 1977 for conflicting evidence of octave perception). Previously, we developed, refined and

then replicated a test of human octave perception using operant conditioning that did not require explicit musical training (Hoeschele, Weisman & Sturdy, in press). We divided a Western music octave into three ranges and trained human participants with alternating contingency pattern (i.e., either the middle range was reinforced and the other two ranges were not, or vice versa), and then tested their response to the same set of notes in a neighbouring octave. We concluded that humans will naturally transfer their discrimination to a novel octave (Hoeschele, Weisman & Sturdy, in press), which suggests that humans are paying attention to pitch chroma in this task. This test provided the first evidence of octave perception in a simple task that could be used to test other species without the requirements of musical training and language, both of which had characterized other, previous tasks used with human participants.

In the current study, we applied our octave perception procedure with humans (Hoeschele, Weisman & Sturdy, in press) to black-capped chickadees using an operant go/nogo task. Black-capped chickadees reside year-round in Canada and in the northern two-thirds of the United States (Smith 1991). Over this wide range except for a few geographically isolated "islands", black-capped chickadees sing the same relative pitch based fee-bee song (Gammon &Baker 2004; Kroodsma et al. 1999; Weisman et al. 1990). There are good reasons to expect that the pitch perception abilities of birds and humans differ. Most important, birds are vastly more accurate than humans at using pitch height perception to group tones into rewarded and unrewarded sequential ranges of

pitches (Weisman et al. 2004). In the present experiment, we insured that predictions from both the octave and pitch-range hypotheses could be detected in testing for generalization and transfer. We chose to study chickadees for several reasons: 1) as a songbird, chickadees rely heavily on acoustic stimuli in the wild, 2) through our past work with chickadees we have a strong understanding of their pitch perception in other respects, 3) by studying such a distantly related species, we can determine whether octave perception is a fundamental component of auditory perception or whether it is species-specific.

There were three possible outcomes that we contemplated before running this experiment: 1) the birds could show octave generalization and respond following the same pattern in the training and testing octaves (i.e. ABA ABA), 2) because we trained with three ranges, the birds could show the opposite pattern by ignoring the octave and learning a general rule that every other range is reinforced (i.e., ABA BAB), or 3) the birds could respond nondifferentially to the higher octave because it was outside the range of training.

Method

Subjects

Twenty black-capped chickadees (ten male and ten female as identified by DNA analysis; Griffiths et al 1998) were tested between September 2010 and February 2012. These birds were captured at the Barrier Lake Field Station in

170

Kananaskis Country, Alberta, Canada (51◦02' N, 115◦03' W) or in and around

Edmonton, Alberta, Canada (Stony Plain, 53˚46N, 114˚01W; North Saskatchewan

River Valley, 53˚53N. 113˚53W; Mill Creek Ravine, 53˚52N, 113˚47W) between

December 2008-January 2011. By examining the color and shape of their outer

tail retrices (Pyle 1997), we judged the birds were at least one year of age at the

time of their capture.

Chickadees were maintained on a light/dark cycle that mimicked the

natural light cycle for Edmonton, Alberta, Canada. Each chickadee had prior

experience with relative pitch discriminations involving natural and synthetic

conspecific songs (Hoeschele et al. 2012) but was naïve to the current stimulus

set.

Apparatus

Prior to the experiment, birds were housed individually in Jupiter Parakeet

cages ($30 \times 40 \times 40$ cm (Rolf C. Hagen, Inc., Montreal, Canada) in colony rooms.

The birds had visual and auditory, but not physical, contact with one another.

Birds had free access to food (Mazuri Small Bird Maintenance Diet; Mazuri, St

Louis, MO, USA), water (vitamin supplemented on alternating days; Prime

vitamin supplement (Hagen, Inc., Montreal, Canada), grit and cuttlebone. Birds

were given three to five sunflower seeds daily. Birds also received one superworm

(*Zophobas morio*) three times a week and a mixture of greens and eggs twice a

week.

During the experiment, birds lived in modified Parakeet cages. Each cage was placed inside a ventilated, sound-attenuating chamber and contained three perches and dispensers for water and grit. The chambers were illuminated by a 9-W, full spectrum fluorescent bulb. An opening on the side of the cage (11 × 16 cm) gave each bird access to a motor-driven feeder (Njegovan et al., 1994). Infrared cells mounted on the feeder and the perch closest to the feeder (the request perch) monitored the position of the bird. A personal computer connected to a single-board computer (Palya & Walter, 2001) scheduled trials and recorded responses to stimuli. Stimuli were played from a CD through either a Cambridge A300 Integrated Amplifier (Cambridge Audio, London, England) or a NAD310 Integrated Amplifier (NAD Electronics, London, England) and then through a Fostex FE108 full-range speaker (Fostex Corp., Japan; frequency range 200–16000 Hz) located beside the feeder. See Sturdy and Weisman (2006) for a detailed description and illustration of the apparatus.

Stimuli

The training and test tones were sine waves synthesized at the frequencies of the 24 successive chromatic musical notes from C4 (262 Hz) to B5 (988 Hz). The exact frequency of each note can be found using the formula $Fx = 2(1/12) \times Fx\text{-}1$, where C4 = 262 Hz. One version of each tone was synthesized at 70 dB and a second version at 80 dB to provide controls for amplitude. This strategy made pitch a more salient determinant of the discrimination by eliminating the

confounding of pitch with loudness (Moore, 1989). The amplitude of each note

was standardized in the sound attenuating chamber at the approximate position of

the bird's head using a Brüel & Kjær Type 2239 (Nærum, Denmark) dB meter to

ensure that all frequencies were presented at the correct dB level (70 or 80 ± 0.3

dB, A weighting, slow response) and then were measured periodically during the

experiment with a Radio Shack Sound Level Meter (A weighting, slow response;

Radio Shack, Fort Worth, USA).

Procedure

Pretraining.

Once a bird had learned to use the request perch and feeder to obtain food,

the pretraining phase began. This phase was designed to promote a high, uniform

percent responding to all the notes before discrimination training. Birds received

food for responding on every trial (with future training and probe notes from

Octaves 4 and 5, respectively). A trial began when the bird landed on the request

perch and remained for 900-1100 ms. A note (one of the 24 notes already

described played at 70 or 80 dB: 48 notes in all) was selected randomly without

replacement and played once in each cycle through the tones. If the bird left the

perch before stimulus offset, the trial was treated as incomplete and resulted in a

30-s timeout with the house light off. If the bird left the perch and entered the

feeder within 1 s after stimulus completion, it received 1 s access to food then a

30-s intertrial interval with the house light on. If the bird remained on the perch

throughout the entire note and the 1 s following stimulus completion, it received a 60-s intertrial interval with the house light on, but could terminate the intertrial interval early if it left and subsequently returned to the perch. This increased the probability that the bird made a response to each note (i.e., on every trial). Birds remained in the pretraining phase until they had consistently high feeding rates (either feeding on $\geq$ 60% of the trials for six 500 trial blocks or feeding on $\geq$ 30% of the trials for 12 blocks) and roughly equal responding to future S+ and S- notes ($\leq$ 3% difference between responding to future S+ and S- stimuli for at least four 500 trial blocks).

Discrimination training.

The procedure was similar to pretraining, except responses (approaches to the feeder) to some of the notes in Octave 4 were no longer rewarded and all notes from Octave 5 were omitted. Specifically, responses to the middle four notes in Octave 4: E, F, F#, or G (the S+ notes), were rewarded as during pretraining. Responses to the lower notes: C, C#, D, or D# and the upper notes: G#, A, A#, or B (the S- notes) were not rewarded. Responses on S- trials now resulted in a 30-s intertrial interval with the houselights off. The criterion to complete discrimination training was a minimum of six 500-trial blocks with a discrimination ratio (DR) $\geq$ 0.80. The final two blocks above 0.80 DR had to be consecutive. See response measures for DR calculations. One female chickadee

was removed from the experiment after failing to reach this criterion in well over 100 blocks.

Discrimination training with reduced reward.

This training phase was identical to the discrimination training just described, except that responses to S+ tones were rewarded with food intermittently: with a probability of 0.85. On unrewarded S+ trials, the ITI was 30 s with the house light on, but without access to food. Intermittent reward was introduced to blur the distinction between rewarded trials and unrewarded test trials during subsequent generalization. Chickadees were trained until they completed two consecutive 500-trial blocks with a DR ≥ 0.80.

Generalization testing.

Trials were similar to discrimination training with reduced reinforcement, except that we added the 12-test notes from Octave 5 to each block of trials. Test notes were never reinforced. The notes from Octave 4 were presented 40 times per trial block (20 times at each dB level) for a total of 480 Octave 4 trials and 12 Octave 5 trials per block. The presentation order of notes was random without replacement.

Each chickadee heard half the test tones at 70 dB and the other half at 80 dB. Either odd ordered notes in ascending order of frequency (i.e., C5, D#5, F5 etc.) were played at 70 dB and the even ordered tones (i.e., C#5, E5, F#5 etc.)

were played at 80 dB, or the reverse. We counter-balanced such that half of the birds of each sex in each group received each version of the test. Each bird completed a minimum of three blocks of testing.

Transfer of training.

After the generalization test, birds completed a minimum of one block of discrimination training with a DR ≥ 0.80 before beginning transfer training. The birds were divided into two groups: the original-discrimination transfer group (n = 9, 4 females) and the opposite-discrimination transfer group (n = 10, 5 females). We conducted transfer with the 24 tones that comprise octaves 4 and 5, each selected randomly and played once at 70 and 80 dB in each cycle of training. In both transfer groups, responses to the tones in Octave 4 were rewarded following the same contingencies as during training; that is, responses to E, F, F#, or G were rewarded but responses to C, C#, D, or D#, the lower tone range and G#, A, A#, or B, the upper tone range were not rewarded. However, the transfer groups had different contingencies in Octave 5. The original-discrimination transfer group had the same contingencies in Octave 5 as in Octave 4. The opposite-discrimination transfer group had the opposite contingencies in Octave 5, i.e., responses to notes in the upper and lower note ranges were rewarded and responses to notes in the middle range were not.

Response Measures.

In many analyses we report the average percent of response either for an individual note or across each note range (e.g., the percent response for notes E, F, F#, and G averaged over a designated block of trials). To determine whether the birds successfully learned to discriminate the S+ note range from the two S- note ranges we calculated a discrimination ratio (DR) for each block of training. The DR is a standard measure that falls between 0 and 1, where 0 means all visits to the feeder followed S- notes, 0.5 represents chance with equal responding to S+ and S- notes, and 1 represents perfect discrimination of S+ from S- note ranges. DR is calculated by dividing the average percent response to S+ notes by average total percent responses to all notes. As in previous work, we excluded interrupted trials from the total before calculating percent response (see Guillette et al 2010 for details of the calculations).

Results

Discrimination.

We conducted an ANOVA comparing percent response across ranges in the final 500 trials and observed a significant effect of note range, see Figure 5-1, $F(2,36) = 187.75$, $p = .0001$, $\eta 2 = 0.912$. In planned comparisons, we determined that the chickadees responded significantly more to notes in the middle range than

to notes in the lower or upper ranges, ts(18) ≥ 14.40, ps < .0001. Every bird

responded more to notes in the middle than the lower or upper ranges.

Generalization.

In repeated measures ANOVA comparing percent response across note

ranges over generalization testing, we observed a significant effect of note range,

$F(5,90) = 43.78$, $p < .0001$, η2 = 0.701, as shown in Figure 5-2, across the note

ranges in Octaves 4 and 5. We then conducted separate sets of planned

comparisons in Octave 4 (the former training octave) and Octave 5 (the novel

generalization octave). In Octave 4, chickadees responded significantly more to

notes in the middle S+ range than to notes in either the lower or upper S- ranges,

ts(18) ≥ 17.92, ps < .0001, much as they did during discrimination training. In

Octave 5, chickadees responded less to notes in the middle range than to notes in

the upper or lower ranges, but these differences were not statistically significant,

ts(18) ≤ 1.26, ps ≥ .225. In Octave 4, the birds responded much as they did during

discrimination training, but in Octave 5, they responded similarly to notes across

ranges, and hence gave no evidence of generalization from Octave 4 based on

octave equivalence.

Transfer.

We conducted an ANOVA comparing percent response across ranges in

the final 500 trials and observed a significant effect of transfer group, $F(1,17) =$

12.24, $p = .0028$, η2 $= 0.419$, note range, $F(5,85) = 36.49$, $p < .0001$, η2 $= 0.520$,

and a significant interaction between group and range, $F(5,85) = 16.71$, $p <$

0.0001, η2 $= 0.238$, see Figure 5-3. In Octave 4, tests of simple effects found the

original-discrimination transfer group and opposite-discrimination transfer group

gave very similar percentages of responding to the three note ranges, $ts(17) \leq$

1.08, $ps \geq 0.295$, see Figure 5-3. Planned comparisons showed that in Octave 4,

both the original-discrimination transfer group, $ts(8) \geq 7.74$, $ps < .0001$, and the

opposite-discrimination transfer group, $ts(9) \geq 9.30$, $ps \leq 0.001$, responded

significantly more to notes in the middle S+ range than to notes in either the lower

or upper S- ranges much as they did during discrimination and generalization.

In Octave 5, the transfer octave, tests of simple effects showed that the

opposite-discrimination transfer group responded much more in each range than

the original-discrimination transfer group, $ts(17) \geq 2.37$, $ps \leq 0.029$, which was

expected, given that the opposite-discrimination transfer group was rewarded on

twice as many trials as the original-discrimination transfer group (we counter-

balanced this without any effect in the work with humans, see Hoeschele,

Weisman & Sturdy in press). Planned comparisons in Octave 5 showed that,

although the original-discrimination transfer group responded more to notes in the

middle range than notes in the lower or upper ranges, these differences were not

statistically significant, $ts(8) \leq 1.44$, $ps \geq .19$. That is, despite extensive training

in the original discrimination in Octave 4, the original-discrimination transfer

group failed to learn the identical discrimination in Octave 5 and thus failed to

show significant positive transfer of the discrimination from Octave 4. In contrast, planned comparisons in Octave 5 in the opposite-discrimination transfer group found significantly more responding to notes in the upper and lower ranges than the middle range, ts(9) ≥ 2.47, ps ≤ .035. That is, the opposite-discrimination transfer group learned the discrimination in Octave 5 and thus appears to have benefited from positive transfer from learning the opposite discrimination in Octave 4.

As further evidence that the opposite- but not the original-discrimination transfer group showed positive transfer from initial discrimination training, we present a correlational analysis of the match between the percent responses to individual notes in Octave 5 during the generalization test and during original- or opposite-discrimination transfer group training. Percent response note-by-note in Octave 5 is shown for two transfer groups in the upper and lower panels of Figure 5-4. Equivalent results for the generalization test, administered prior to transfer training, are shown for comparison in gray and without error bars to reduce confusion with results for the transfer tests in Figure 5-4. Note-by-note, the results for generalization and transfer in the opposite-discrimination transfer group were moderately and significantly correlated, r(17) = .55, p = .015, whereas the results for generalization and transfer in the original-discrimination transfer group were not significantly correlated, r(17) = -.13, p = 0.53. That is, generalization successfully predicted positive transfer to the opposite

discrimination but provided no evidence for transfer to the original discrimination.

Finally, we asked whether female chickadees differed from males in the accuracy of their discrimination, generalization to Octave 5, or transfer to a rewarded discrimination in Octave 5. As shown in Table 1, in none of these instances did females and males DRs differ significantly. Indeed, phase-by-phase in the experiment females and males had either identical or nearly identical DRs.

Discussion

Here we applied our operant test for octave equivalence, first verified in humans, in black-capped chickadees. We found that the chickadees did not show evidence of octave equivalence in generalization testing or transfer. This is in contrast to humans, who showed evidence of octave equivalence even when contingencies were reversed in the two octaves using the same operant paradigm (Hoeschele, Weisman, & Sturdy, in press).

If anything, the birds' responses in the testing octave trended towards being opposite to responses in the training octave during generalization. That is, the responses trended towards our alternate predicted outcome that the birds could show an opposite pattern in the two octaves by ignoring the octave and learning a general rule that every other range is reinforced (i.e., ABA BAB rather than ABA ABA). This is quite odd when one considers what effects pitch height alone

should generate on responding in the testing octave. Since we used three ranges of alternating contingency, the lowest and highest range within an octave had the same contingency. The testing octave, being one octave higher than the training octave, was closest in pitch height to the highest range of the training octave. If anything, responding in the testing octave, at least in the lowest range, should be similar to responding to the highest range of the training octave according to pitch height. However, birds had an easier time when this range had the opposite contingency. There are several possible explanations for this. One possible explanation is the difference in reinforcement between the two transfer groups. The reversed transfer group had twice the opportunity for reinforcement in the higher octave, since two ranges were reinforced instead of one. Because songbirds have much more precise absolute pitch perception than humans (Weisman et al. 2006; see Weisman et al. 2010 for exceptions), it may be that they simply learned the reversed discrimination more quickly because of the greater chance for reinforcement in this octave. By replicating this experiment with the opposite contingencies we could test this idea. However, this idea does not explain why responding during generalization testing, where the birds received no feedback, also trends towards a similar pattern. Another explanation for our results is that the birds learned an alternating rule. That is, the birds are learning the rule that after equal log-linear spaces in frequency, contingencies reverse. Because in the reversal group every other range was reinforced, this may have been a simpler pattern for the birds to learn than the unreversed pattern which has a large

182

intermediate range of the same reinforcement (i.e., the highest range of the lower octave and the lowest range of the higher octave are essentially one range because they have the same contingencies in an ABA ABA pattern). Further exploration could help us understand what the birds were relying on in this discrimination.

This study is the first to directly support the notion that chickadees, and perhaps songbirds, or even birds in general, do not perceive chroma, but rely solely on pitch height perception. Previous work with European starlings also showed that starlings did not perceive octaves (Cynx, 1993). However, we applied Cynx's procedure to humans and found that humans also failed the task that he used to test the birds (Hoeschele, Weisman, & Sturdy, in press). This outlines one of the important points of our methodology: it was important to first show that humans would show octave equivalence in the same task before it was possible to interpret the behaviour of the chickadees in the current study. We believe this is an important point for comparative work in general: without similar methodologies between species it is very difficult to interpret the data, and in particular, negative findings such as Cynx's (1993) failure to detect octave equivalence. We now have a protocol with which we can test any animal, using an operant conditioning paradigm, to respond to sinewave stimuli. Our goal is to test many species to understand the breadth of the octave equivalence phenomenon.

But what animals may be the most relevant to test in our paradigm? Because of the potential biological connection between music and language (Fitch, 2005; Hauser & McDermott, 2003; Masataka, 2009; Deutsch, 2006), it

may make the most sense to study animals that share other properties of complex auditory perception, or language-like abilities. As it turns out, black-capped chickadees, as members of the oscine songbirds, are a species with a rare language-like ability. These songbirds, along with parrots, hummingbirds, cetaceans, bats, and elephants are the few groups of animals along with humans that are thought to be vocal learners (Tyack, 2008). Vocal learners produce normal adult vocalizations only if they are exposed to a tutor when young (but see Fehér, Wang, Saar, Mitra & Tchernichovski, 2009 for an interesting departure). Because of this, similar neural architecture, and the many closely-related species among which comparisons are possible, songbirds are an excellent model for human language (see Doupe & Kuhl, 1999 for review). But without octave perception, a hallmark of human pitch perception in both language and music, there may be some limitations as to the similarity between the songbird and human vocal learning systems. It would be interesting to turn both to non-vocal learning and vocal-learning mammals to see whether octave perception is limited to humans, limited to mammalian vocal-learners, or limited to mammals as a whole or some other subset of mammals.

Octave perception might be especially relevant for humans compared to chickadees. For one thing, humans tend to vocalize in different octaves depending on their sex and development. Male and female voices are roughly an octave apart (Titze, 2000), and men and women will often sing together by singing an octave apart (Dowling & Harwood, 1986). Chickadees, on the other hand, do not

undergo major pitch height changes in vocal pitch as human males do (Baker, Baker & Gammon, 2003). Male and female chickadee vocalizations also occupy the same range of pitches (Ficken, Ficken & Witkin, 1978). As such, although chickadees do match pitches in some circumstances (Foote, Fitzsimmons, Mennill & Ratcliffe, 2008), it is not possible for them to octave generalize when singing to be able to match the pitch of another individual. Therefore it is possible that octave perception would be more likely to be found in species that have large ranges of vocalizations, or where male and female or young members of the species vocalize in different octave ranges. When choosing further species to study it may be relevant to consider whether octaves could possibly play a role in the natural vocalizations of that species. Bottlenose dolphins, for instance, can cover a very large range of frequencies that cover many octaves (Jensen, Beedholm, Wahlberg, Bejder, & Madsen, 2011). As a mammalian vocal learner, the dolphin is an ideal candidate to test for octave perception using our protocol. To date, there is some evidence that dolphins perceive octaves (Ralston & Herman, 1995; Richards, Woltz, & Herman, 1984), but by testing dolphins with our protocol we can directly compare this species to humans, now chickadees as well and, ideally, many other species in the future.

An alternate hypothesis to the idea that vocal range might influence the perception of octave equivalence, is that octave equivalence is a simple coping mechanism for species with inferior pitch height perception. Black-capped chickadees are highly adept absolute pitch height processors compared to humans

(Weisman, Njegovan, Williams, Cohen, & Sturdy, 2004). Because black-capped chickadees can very easily discriminate pitch height, the addition of pitch chroma perception would not improve their ability to discriminate pitches. Even the rare humans that are considered to have absolute pitch (estimated to be less than 1 in 10,000; Bachem, 1955), frequently make octave errors (Bachem, 1937; Weisman et al., 2010) which suggests that they actually have absolute pitch chroma rather than absolute pitch height. Unlike with absolute pitch, humans tend to be good at perceiving relative pitch relationships among notes (Weisman & Ratcliffe, 2004). For example, being able to recognize the same melody played in a different key (notes have different absolute pitches, but are the same pitch distance relative to one another on a log-linear scale). It is possible that octave equivalence may be a relative pitch mechanism to compensate for poor pitch height ability.

Our overarching goal in this research is to understand the fundamental components of complex auditory processing across species. For humans, it seems that octave perception is a fundamental component of how sounds are processed and produced. Without octave perception, music and language would both likely be quite different. Despite this, complex auditory processing does not appear to depend on octave perception, as black-capped chickadees show complex vocal properties such as open-ended call generation (Hailman & Ficken, 1996) and categorization (Bloomfield, Sturdy, Phillmore & Weisman, 2003), and highly sophisticated vocal markers such as learned associations between predators and threat (Avey, Hoeschele, Moscicki, Bloomfield & Sturdy, 2011). We plan to

186

continue studying the influence of octave equivalence on pitch judgments in

humans while simultaneously studying what factors influence the development of

octave equivalence by studying many other species with our paradigm.

Table 5-1. Mean ± standard deviation of male and female discrimination ratios in each octave of each phase of the experiment. Student t tests results are also provided showing that males and females were not statistically different in any comparison.

Table 5-1 Analysis of Female and Male Performance (Discrimination Ratios)

Training/Testing Phase	Females	Males	Statistical Test
Acquisition	.83±.01	.83±.01	$t(17)=0.19, p=0.85$
Generalization-Octave 4	.83±.01	.83±.02	$t(17)=-0.19, p=0.85$
Generalization-Octave 5	.41±.07	.48±.07	$t(17)=0.91, p=0.38$
Transfer-Octave 4	.82±.02	.82±03	$t(17)=0.22, p=0.83$
Transfer-Octave 5	.47±.06	.46±.08	$t(17)=0.15, p=0.88$

A   Three-Range Discrimination

Training Octave 4

Lower   Higher

S-   S+   S-
or
S+   S-   S+

B   Generalization

Testing Octave 4 | Testing Octave 5

Lower   Higher

Figure 5-1.

189
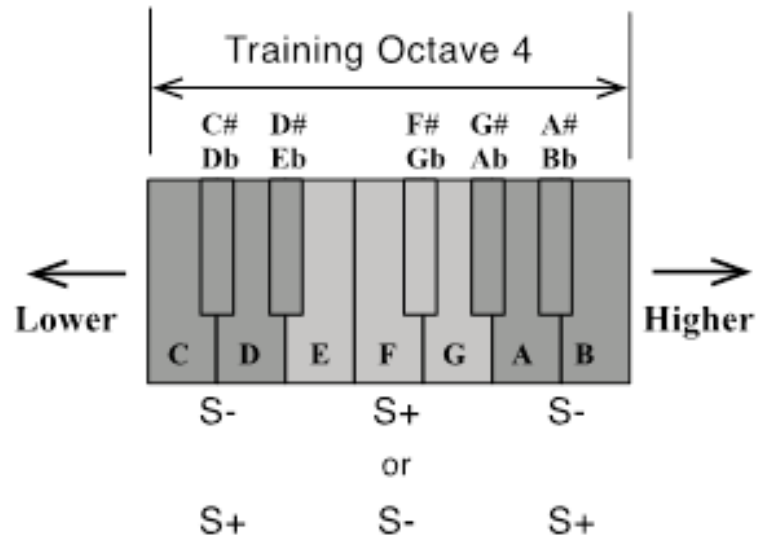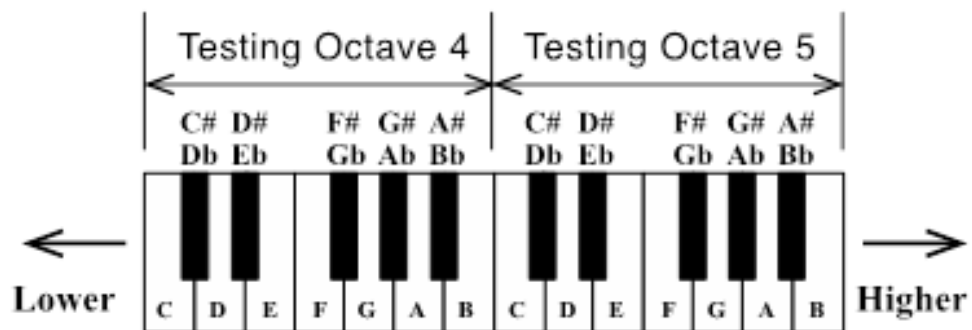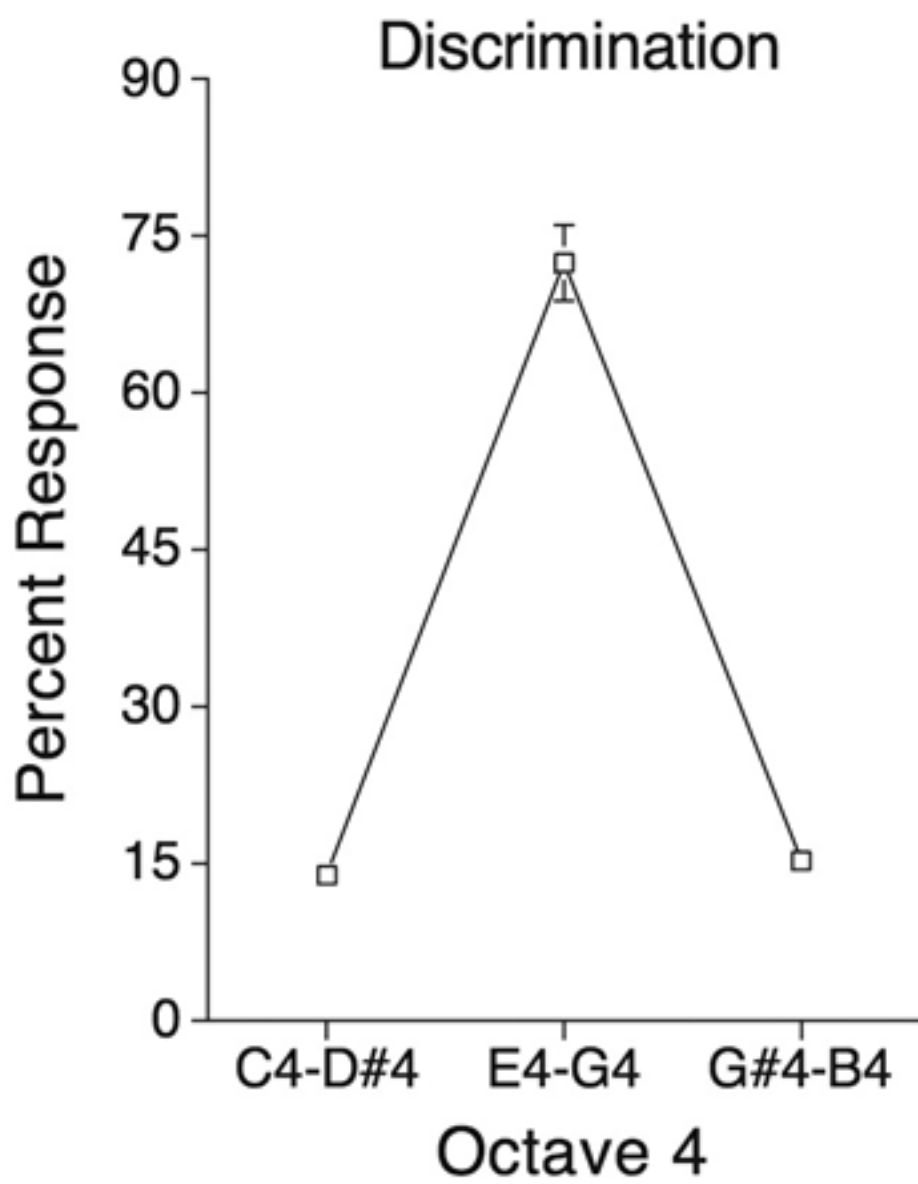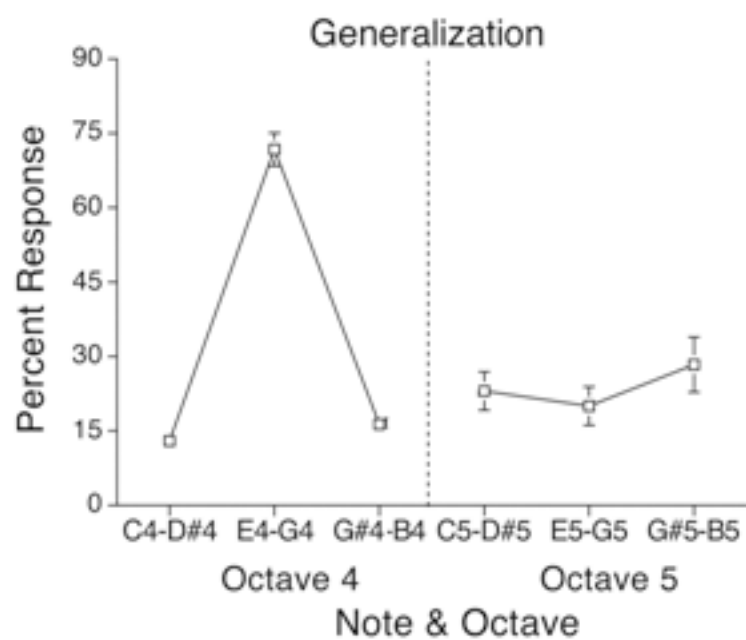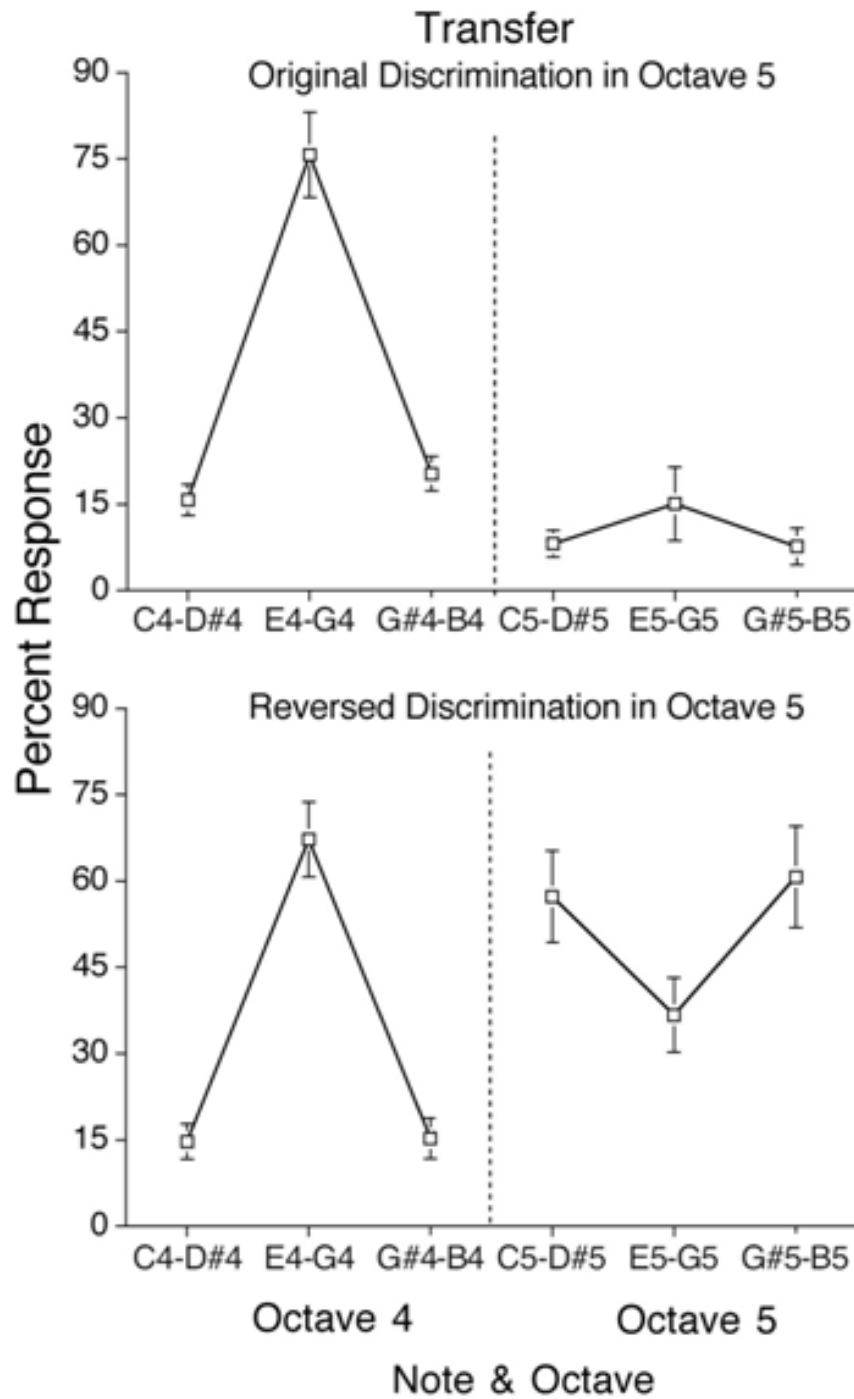
Figure 5-2.

Figure 5-3.

Figure 5-4.
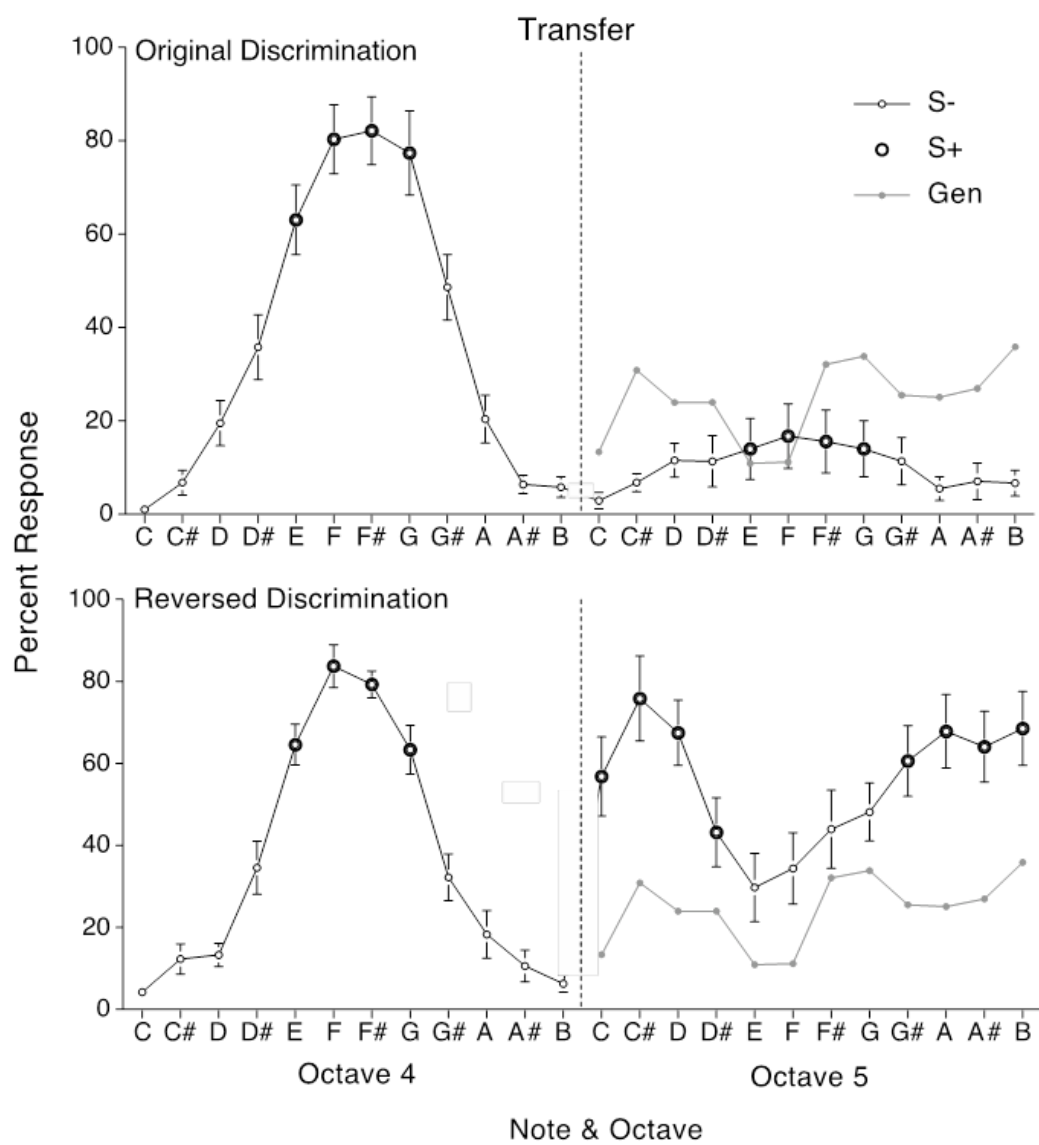
Figure 5-5.

References

Allen, D. (1967). Octave discriminability of musical and non-musical subjects. *Psychonomic Science, 7*, 421-422.

Avey, M. T., Hoeschele, M., Moscicki, M. K., Bloomfield, L. L., & Sturdy, C. B. (2011). Neural correlates of threat perception: Neural equivalence of conspecific and heterospecific mobbing calls is learned. *PLoS One, 6*(8), e23844 1-7.

Bachem, A. (1955). Absolute pitch. *Journal of the Acoustical Society of America, 27*(6), 1180-1185.

Baker, M. C., Baker, M. S. A., & Gammon, D. E. (2003). Vocal ontogeny of nestling and fledgling black-capped chickadees *Poecile atricapilla* in natural populations. *Bioacoustics, 13,* 265-296.

Blackwell, H. R., & Schlosberg, H. (1943). Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology, 33*, 407-419.

Bloomfield, L. L., Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (2003). Open-ended categorization of chick-a-dee calls by black-capped chickadees *(Poecile atricapilla)*. *Journal of Comparative Psychology, 117*(3), 290–301.

Cartelle, E. C., & Kendall, R. A. (1999). Comparative music perception and cognition. In D. Deutsch (Ed.), *The Psychology of Music 2nd Ed* (pp. 215-264). San Diego, California: Academic Press.

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

Cynx, J. (1993). Auditory frequency generalization and a failure to find octave generalization in a songbird, the European starling (*Sturnus vulgaris*). *Journal of Comparative Psychology,107*, 140-146.

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception &. Psychophysics, 11*, 411-412.

Deutsch, D. (2006). The enigma of absolute pitch. *Acoustics Today, 2,* 11-19.

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 Inc.

Dowling, W. J., & Hollombe, A.W. (1977). The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception &. Psychophysics, 21*, 60-64.

Fehér, O., Wang, H., Saar, S., Mitra, P. P., & Tchernichovski, O. (2009). De novo establishment of wild-type song culture in the zebra finch. *Nature, 459*(7246), 564-568.

Ficken, M. S., Ficken, R. W., & Witkin, S. R., (1978). Vocal repertoire of the black-capped chickadee. *The Auk, 95*, 34-48.

Fitch, W.T., (2005). The evolution of music in comparative perspective. *Annals of*

the New York Academy of Sciences, 1060, 29-49.

Foote, J. R., Fitzsimmons, L. P., Mennill, D. J., & Ratcliffe, L. M. (2008). Male
chickadees match neighbors interactively at dawn: support for the social
dynamics hypothesis. *Behavioral Ecology, 19*(6), 1192–1199.

Gammon, D. E., Baker, M. C. (2004). Song repertoire evolution and acoustic
divergence in a population of black-capped chickadees, *Poecile
atricapillus*. *Animal Behavior, 68*, 903-913.

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.

Hailman, J. P. and Ficken, M. S. (1996). Comparative analysis of vocal
repertoires, with reference to chickadees. In: Kroodsma, D. E. and Miller,
E. H., (Eds). *Ecology and Evolution of Acoustic Communication in Birds*
(pp. 136–159). Ithaca, New York: Comstock Publishing Associates of
Cambridge University Press.

Hauser, M.D. McDermott, J. (2003). The evolution of the music faculty: a
comparative perspective. *Nature Neuroscience, 6*(7), 663-668.

Hoeschele, M., Guillette, L.M., & Sturdy, C.B. (2012). Biological relevance of
acoustic signal affects discrimination performance in a songbird. *Animal
Cognition, 15*(4), 677-688.

Hoeschele, M., Weisman, R.G., & Sturdy, C.B. (in press). Pitch chroma

discrimination, generalization and transfer tests of octave equivalence in humans. *Attention, Perception, & Psychophysics*, MS# PP-ORIG-12-057.R2.

Jensen, F. H., Beedholm, K., Wahlberg, M., Bejder, L., & Madsen, P. T. (2012). Estimated communication range and energetic cost of bottlenose dolphin whistles in a tropical habitat. *Journal of the Acoustical Society of America, 131*(1), 582.

Kallman, H. J. (1982). Octave equivalence as measured by similarity ratings. *Perception & Psychophysics, 32*, 37-49.

Kroodsma, D. E., Byers, B. E., Halin, S. L., Hill, C., Minis, D., Bolsinger, J. R., Dawson, J., Donelan, E., Farrington, J., Gill, F. B., Houlihan, P., Innes, D., Keller, G., MacAulay, L., Marantz, C. A., Ortiz, J., Stoddard, P. K., Wilda, K. (1999). Geographic variation in black-capped chickadee songs and singing behavior. *The Auk, 116*, 387-402.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology, 5*, 579-594.

Levitin, D. J. (1994). Absolute memory for musical pitch: evidence from the production of learned melodies. *Perception & Psychophysics, 56*(4), 414–423.

Masataka, N. (2009). The origins of language and the evolution of music: A comparative perspective. *Physics of Life Reviews, 6*, 11-22.

197

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.

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.

Peter, B., Stoel-Gammon, C., & Kim, D. (2008). Octave equivalence as an aspect
of stimulus response similarity during nonword and sentence imitations in
young children. *Speech Prosody 2008*, 731-734.

Pierce, J. R. (1999). The nature of musical sound. In (Ed) Deutsch, D. *The
psychology of music*. San Diego, CA: Academic Press.

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

Ralston, J. V., & Herman, L. M. (1995). Perception and generalization of
frequency contours by a bottlenose dolphin (*Tursiops truncatus*). *Journal
of Comparative Psychology, 109*(3), 268–277.

Richards, D. G., Wolz, J. P., & Herman, L. M. (1975). Vocal mimicry of
computer-generated sounds and vocal labeling of objects by a bottlenosed
dolphin, *Tursiops truncates. Journal of Comparative Psychology, 98*(1),
10–28.

Smith, S. M. (1991). *The Black-Capped Chickadee: Behavioral Ecology and
Natural History*. Ithaca, New York: Cornell University Press.

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

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

Tyack, P. L. (2008). Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *Journal of Comparative Psychology, 122*(3), 319-331.

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

Weisman, R. G., Njegovan, M. G., Williams, M., Cohen, J., & Sturdy, C. B. (2006). The comparative psychology of absolute pitch. In E. A. Wasserman & T. Zentall (Eds.), *Comparative Cognition: Experimental Explorations of Animal Intelligence* (pp. 71-86) Oxford: Oxford University Press.

Weisman, R. G., & Ratcliffe, L. M. (2004). Relative pitch and the song of black-capped chickadees. *American Scientist, 92*(6), 532-539.

Weisman, R. G., Njegovan, M. G., Williams, M., Cohen, J., & Sturdy, C. B. (2006). The comparative psychology of absolute pitch. In E. A. Wasserman & T. Zentall (Eds.), *Comparative Cognition: Experimental Explorations of Animal Intelligence* (pp. 71-86). Oxford: Oxford

University Press.

Weisman, R .G., Balkwill, L.-L., Hoeschele, M., Moscicki, M. K., Bloomfield, L.
   L., & Sturdy, C. B. (2010). Frequency-range discriminations in boreal
   chickadees and humans: Exceptions that test a phylogenetic rule. *Learning
   and Motivation, 41*, 156–173.

Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiworth, J. J. (2000).
   Music perception and octave generalization in rhesus monkeys. *Journal of
   Experimental Psychology: 805 General, 129*, 291-307.

Chapter 6 General discussion

Summary of data chapters

Human music perception is a thoroughly studied field, but rarely addresses the biological underpinnings of this perception. In this thesis my goal was to take the knowledge we already have for music perception, and design protocols to be able to assess how this perception functions across species in order to gain insight about the biological origins. The species I focused on were humans and black-capped chickadees. In chapter 2, I found that the biologically-relevant relative pitch interval in the black-capped chickadee *fee-bee* song influences the way this species discriminates pitches. In chapter 3, using a similar go/no-go operant paradigm to comparatively study humans and black-capped chickadees, I found that these species discriminate musical chords in a similar manner. They also are able to transfer this relative pitch discrimination to novel absolute pitches. In chapter 4, I designed an operant go/no-go task in humans that demonstrated the natural tendency of humans to perceive pitch chroma. In chapter 5, using a similar protocol to chapter 4, I found that black-capped chickadees do not appear to use pitch chroma when discriminating absolute pitch.

Biological relevance

Chapter 2 was the only study that was not comparative and focused solely on black-capped chickadees. The reason for this was the unique song of the black-capped chickadee, the *fee-bee* song, which contains a fixed relative pitch interval that can be performed at different absolute pitches (Horn et al. 1992; Weisman et

202

al. 1990). While it is known that humans develop relative pitch ability at a young age (Miyazaki, 1995), chickadees and songbirds as a whole are not known for their relative pitch. Quite the contrary, these species tend to rely more on absolute pitch (Weisman et al. 2006; see Weisman et al. 2010 for exceptions). The absolute pitch of stimuli can even sometimes hinder the ability of these animals to learn or transfer simple rules. For example starlings, cowbirds and mockingbirds can learn the relative pitch rule of whether a series of tones are ascending or descending, but are unable to transfer this discrimination to a novel pitch range (Hulse & Cynx, 1985). This reliance on absolute versus relative may be specific to pitch. Black-capped chickadees in particular were recently shown to be able to learn to discriminate sets of acoustic stimuli based on the relational concept of "same" versus "different". These birds also transferred this discrimination to a number of novel stimuli except stimuli that were from a novel pitch range (Hoeschele, Cook, Guillette, Hahn, & Sturdy, in press). However, at least one other study with a songbird species, in addition to chapter 2, have recently shown that if biologically relevant stimuli are being discriminated, relative pitch is used more readily. Specifically, Bregman, Patel, and Gentner, (2012) showed that starlings can recognize melodies that have been pitch shifted, if the melodies are the songs of other starlings. They do not recognize pitch shifted piano melodies. These results suggest that the preference to use absolute pitch may be true only for sounds outside of conspecific vocalizations or perhaps vocalizations in general. It would be interesting to study how songbirds respond to the relative pitch of similar

heterospecific vocalizations. With black-capped chickadees this could be achieved by studying the song of the mountain chickadee, which is fairly similar in structure to the song of the black-capped chickadee (Lohr, 2008).

That relative pitch is used more readily when it comes to biologically-relevant stimuli has implications for the interpretation of the rest of the studies in this thesis. Both the synthesized chord stimuli in chapter 3 and the sinewave stimuli used to test for chroma perception in chapter 5 are unnatural sounds. It is possible that if I had used pitch-shifted vocalizations in the later studies that the results would have been quite different. Potentially, chickadees would show chroma perception if the *fee-bee* song had been used as a stimulus instead of sinewave tones. However, there are other reasons to doubt that songbirds have chroma perception, some of which are discussed below (see chroma perception section).

I left one interesting finding unexplained in chapter 2: the chickadees not only had an easier time discriminating the *fee-bee* song interval from other intervals, they also had a comparatively easy time discriminating the largest interval I used in the study (both in the natural shifted *fee-bee* song acquisition, and in averaged final performance). It's important to note that the group trained with the largest intervals were not trained with the *fee-bee* song interval at all, so direct comparisons could not be made. Overall this large interval group learned the discrimination the slowest compared to the other two groups. Despite this, this group managed to discriminate the larger interval within their discrimination more

quickly from the S+ interval than the smaller interval. An explanation that I considered for this that goes beyond the scope of chapter 2 is consonance and dissonance.

Consonance and dissonance are usually used to describe the subjective experience of music. A sound that is more consonant is a sound that is more pleasing, stable, and resolved, and a sound that is more dissonant is more displeasing, unstable, and unresolved (Rasch & Plomp, 1999). Although there is some cultural variation about what intervals are pleasing and displeasing across cultures, part of what makes a sound consonant and dissonant in music is universal (Cook & Fujisawa, 2006; Rasch & Plomp, 1999). In fact, consonance and dissonance partly have to do with the physical structure of sound. The simpler the ratio of the two notes in an interval, the more likely they are to be rated "consonant". The simplest pitch ratios (1:1, 1:2, 2:3, 3:4) were thought to have already been identified by Pythagoras (Plomp & Levelt, 1965) as the "perfect" intervals because of their high consonance. Other scholars, such as Galileo, Mersenne, Leibniz, and Euler noted that notes with simple ratios follow similar pulsing pattern (Plomp & Levelt, 196; Pierce, 1999). For example, Figure 6-1 shows the fundamental frequency of a note and its octave, with a ratio of 2:1, compared to two sinewaves that have a complex ratio. When in phase, a note and its octave (ratio of 1:2; upper panel) will make zero crossings occur together, that is, the relationship between the two tones stays constant over time. In contrast, when two notes are separated by a small ratio (lower panel), they drift further and

further apart which creates a pulsing sound that musicians use to tune instruments. Reducing the pulsing sound reduces the different between two notes and brings about a stable sound as the notes approach a ratio of 1:1. In the 19th century, the harmonics of notes were also considered by von Helmholtz and Wundt (reviewed in Plomp & Levelt, 1965). While a rough beating sound occurs with sinewaves only with small differences in frequency, non-sinewave sounds with harmonics that have a complex ratio with one another will have harmonics that are close in frequency and also create this rough beating sound. These physical properties of consonance and dissonance are often referred to as sensory consonance (Plomp & Levelt, 1965).

It turns out that the interval with the simplest ratio in chapter 2 was the largest interval. This interval approximated a perfect 4th in Western music and is one of the most consonant intervals with a ratio of 3:4 (one of the intervals defined as perfect, see Plomp & Levelt, 1965). In terms of Plomp and Levelt's (1965) theoretical model of consonance and dissonance, this interval is by far the most consonant of all the intervals used in this study on a consonance/dissonance continuum (see Figure 6-2). Thus, it may be that the birds are detecting this physical consonance and dissonance and are using it to discriminate the intervals. Chapter 3 begins to address the question of whether black-capped chickadees attend to consonance and dissonance in complex auditory stimuli.

Because of the potential influence of consonance and dissonance in Chapter 2's task, it may be interesting, in hindsight, to compare how humans

respond to this task as well. Although the *fee-bee* song is not biologically relevant for humans, other aspects of pitch intervals may be, and by comparing their response we might be able to get a clearer picture of what general factors (i.e., not species-specific or biologically relevant factors) could be contributing to the response patterns in the chickadees. If consonance and dissonance are affecting the way chickadees solve the task, this should be seen in humans as well, but if the biological relevance of the *fee-bee* song controls responding, this should not be seen in humans.

Relative pitch

Chapter 3 showed that humans, black-capped chickadees, and pigeons (pigeons based on comparisons with prior data) performed similarly when discriminating musical chords. The important issue is whether the similar response patterns are due to similar underlying mechanisms. Theoretically, similar response patterns could be achieved with very different mechanisms. However, there are probably some underlying mechanisms that are shared across species. One possibility is that all species were attending to the frequency bandwidth of pitch range of the signal. Frequency bandwidth could explain why manipulations of the 5th were easier to detect than manipulations of the 3rd. The 3rd manipulations being manipulations of the middle note, they do not change the overall frequency range covered by the stimuli, whereas 5th manipulations expanded or reduced the frequency range of the fundamental frequencies. On the other hand, because these were harmonically complex stimuli that were generated

with a synthesized French horn timbre, the frequency bandwidth was fairly similar across all stimuli and may not have been that easy to detect.

Based on the findings of chapter 2, it seems that the *fee-bee* song interval may be important in general when chickadees are determining the relative pitch of an interval. Related to this, the sus4 (suspended 4th) chord in chapter 3 contained the *fee-bee* song interval and was generally discriminated more readily by the chickadees than the humans. It would be interesting to test whether chickadees trained with 2 note intervals presented in chapter 2 that either matched or did not match the interval from their *fee-bee* song would transfer this discrimination to chords that contained the same intervals. In fact, it would be interesting to see whether the chickadees could transfer the relative pitch interval from notes separated by time to notes that are played simultaneously at all.

In fact, notes presented separately may be interpreted quite differently than notes that are presented simultaneously because of important physical properties and how these relate to vocalizations. Nonlinearities in vocal production are essentially unpredictable interactions in the vocal apparatus that usually occur in high stress situations when the vocal apparatus has been overdriven (Blumstein, Davitian & Kaye, 2010) Sounds with nonlinearities are considered "harsh" because of their noisy structure and could be considered a kind of dissonant stimulus. These harsh vocalizations are thought to occur in natural vocalizations when an animal is in distress (Blumstein, Davitian, & Kaye, 2010) and may be used to attract attention to the producer (Fitch, Neubauer, &

Herzel, 2002). This chaotic interaction of sounds occurring in a single vocalization might make notes played simultaneously with complex ratios that also have rough-sounding interactions (sensory dissonance) especially striking.

That dissonant intervals where the notes are played simultaneously may be more striking than those separated by time is not to mention the memory component that is necessary when evaluating stimuli that are not played simultaneously. It would be interested to begin investigation on this matter in humans, who already as infants can recognize melodies of separated notes by their pitch relationships (i.e., they can recognize a melody that has been transposed to a novel absolute pitch, Trehub, 2003). Human infants also produce nonlinearities when crying, so it is possible that nonlinearities are also biologically relevant in humans, although adult humans rarely produce nonlinearities (Fitch, Neubauer, & Herzel, 2002). There is also evidence to suggest that film scores also use nonlinearities to manipulate the emotional valence of a situation (Blumstein, Davitian, & Kaye, 2010). Studying to what extent humans can transfer relative pitch information from individual notes to chords and what factors might play a role in this ability (e.g., how consonant the intervals are) may shed light on what role memory and nonlinear vocalizations might have in perceiving pitch intervals.

Another thing to consider is whether consonance and dissonance was a large determinant in the results of chapter 3 for all species or not. As suggested in the discussion of that chapter 3, models have been proposed to explain cross-

cultural consonance and dissonance in humans that go beyond physical consonance and dissonance (e.g., Cook & Fujisawa, 2006). To summarize a point from Cook and Fujisawa's (2006) model, because three note chords are more complex than two note intervals, a general grouping mechanism might account for how easily these chords are processed. If the middle note of a chord is the same ratio distance away from the highest and lowest note, it is less clear whether that note should be grouped with the higher or lower note than if the middle note is closer to one of the notes. This creates what is called "tension" (Cook & Fujisawa, 2006). Both humans and chickadees may have been paying attention to the "tension" of the chords as they discriminated the augmented chord easily for the major chord despite its relatively similar physical or sensory consonance.

It was also an advantage for this study to have a pigeon comparison. This allowed me to begin to consider phylogenetic aspects of relative pitch perception. Pigeons are more closely related to chickadees being another avian species, but chickadees and humans are both vocal learners, unlike pigeons (Tyack, 2008). If relative pitch perception is influenced most strongly by phylogenetic relatedness, we would expect the responses of chickadees to be more similar to pigeons. However, if relative pitch perception is influenced most strongly by vocal learning, we would expect the response of chickadees to be more similar to humans. Surprisingly, the responses of pigeons and humans were more similar to each other than either species was compared to the chickadees. This suggests that there may be a third factor that humans and pigeons share that chickadees do not.

One possible factor could be the absolute pitch of each species' vocalizations. Pigeons and humans have vocalizations that are much lower in pitch than chickadees (pigeon: Baptista & Abs, 1983; human: Titze, 2000; chickadee: Ficken, Ficken, & Witkin, 1978). As a result, these species may have been attending primarily to the fundamental frequencies. Chickadees have much higher-pitched vocalizations and may have been attending to higher spectral frequencies. In fact, the fundamental frequencies of the chords, which were between 262-466 Hz, are below the optimal hearing range of chickadees (see Henry & Lucas, 2010 for the hearing range of closely-related Carolina chickadees and Langemann, Gauger, & Klump, 1998 for the hearing range of the closely-related great tit), whereas they are within both the human (Fay, 1988) and pigeon (Goerdel-Leich & Schwartzkopff, 1984) optimal hearing ranges. Potentially, if we tested chickadees with the same chords presented at a higher absolute pitch their responses would be more similar to those of humans and pigeons.

Chapter 3 also only had human participants that were currently living in North America and therefore exposed to Western music. Testing participants from a region where another musical system predominates could lead to insights as to how much of a role culture played in the response patterns of our participants. This is important when comparing the results of humans to other species, because the results may not be completely representative of the species as a whole. In addition, cross-cultural study could also lend further support to models such as Cook and Fujisawa's (2006) model that try to find simple patterns and

explanations in cross-cultural similarities of pitch perception. If we fully grasped the role culture plays in chord perception, it may be easier to determine what features of chord processing could have biological origins.

Chroma perception

Chapters 4 and 5 assessed whether humans and chickadees, respectively, possessed chroma perception. The findings suggested that humans, but not chickadees, detect pitch chroma. But can this result be generalized to all songbirds? Or even all avian species? This is a difficult conclusion to make without further study. However, it is plausible that chroma perception would not be present in other species that, like chickadees, have vocalizations that do not make use of the octave relationship. In humans, male and female voices are roughly an octave apart (Titze, 2000), and the octave relationship is used when singing together (Dowling & Harwood 1986). Further comparative studies could make it clearer what role vocal range and sexual dimorphism in pitch production might have. Another possibility is that mammals have chroma perception, but avian species do not. There is some evidence that rhesus monkeys (Wright, Rivera, Hulse, Shyan & Neiworth, 2000) and dolphins (Ralston & Herman, 1995; Richards, Woltz, & Herman, 1984) have chroma perception. However, these studies all had very small sample sizes and it may be difficult to replicate these results. I plan to continue to use my protocol to test more species to attempt to determine what factors influence the ability to perceive chroma.

While the range of notes presented were far below those used in chickadee

vocalizations, it does not seem likely that a higher range of notes would be easier for the chickadees to identify. All the chickadees except one learned the discrimination to a high accuracy. Thus, although in humans, the high pitches near the end of the range of a piano (8th octave) are very difficult for people even with absolute pitch to identify (Baharloo, Johnston, Service, Gitschier, & Freimer, 1998), chickadees do not seem to be limited to their preferred range. Regardless, by testing the chickadees with higher pitched notes that come from their own vocal range I could address this potential criticism.

Conclusions

My PhD work has lead me to not only understand a lot more about black-capped chickadee pitch perception, but also human pitch perception and the biological underpinnings of pitch perception. I compared these two species directly by using highly similar operant training and testing protocols; something which tends to be rare in biomusicology.

Using an operant paradigm, rather than having humans rate similarity of chords and tones, meant that humans and chickadees could be directly compared. It is not always the case that verbal responses are the same as intuitive operant responses. For example, Ludvig and Spetch (2011) found that humans will be risk averse for gains and risk seeking for losses if they are asked to choose between verbally described probabilities, but will be risk seeking for gains and risk averse for losses in an experience-based operant task with the exact same probabilities.

However, my results in chapter 3 matched fairly well with what is known about human perception of chords. Prior human data where humans rated the similarity of the chords in chapter 3 resulted in a similar pattern of results (Brooks & Cook, 2010). This suggests that human relative pitch perception may be more consciously accessible than risky decision making. That is, the same mechanism may be used to both rate chords and discriminate them in an operant paradigm.

Conversely, although the perception of octaves has been discussed for a long time in human history (Plomp & Levelt, 1965; Pierce, 1999), empirical data showing the perception of octaves has been difficult (reviewed in chapter 4). In chapter 4, human participants showed that they did perceive chroma across octaves. Given the difficulty Kallman (1982) had getting human participants, especially nonmusicians, to rate notes separated by an octave as being similar, I do not believe it would be possible to obtain similar results if we had used a rating system instead of an operant paradigm. It would have been informative to have asked participants what strategy they were using after they completed the experiment. However, anecdotal support of this notion is that many participants asked at the end of the experiment what the rule was they were supposed to use in the experiment. Many were surprised that the notes had been divided into 3 ranges of alternating contingency. If participants did not even understand what rule they had been using to solve the task, it is unlikely that they consciously octave generalized their response pattern in octave 4 to octave 5. Alternatively, in the chord task participants generally seemed more certain about what it was they were

listening for.

I expect that the studies presented in this thesis will have a significant impact on the field of biomusicology for several reasons. One reason is that the field is at this time very small and relatively unexplored. There has been especially little research on chroma perception in nonhuman animals. Another reason is that systematic human-nonhuman cross-species comparisons using similar tasks are rare, if not absent, in most biomusicology pursuits. Often a direct human comparison is not available, and assumptions are made about human perception based on other studies with very different, language-based, methodologies. Both the chord and octave protocols I present in this thesis have not only been tested in both human and non-human animals, but they have the potential to be applied to any other animal species that can be trained in a Go/NoGo task.
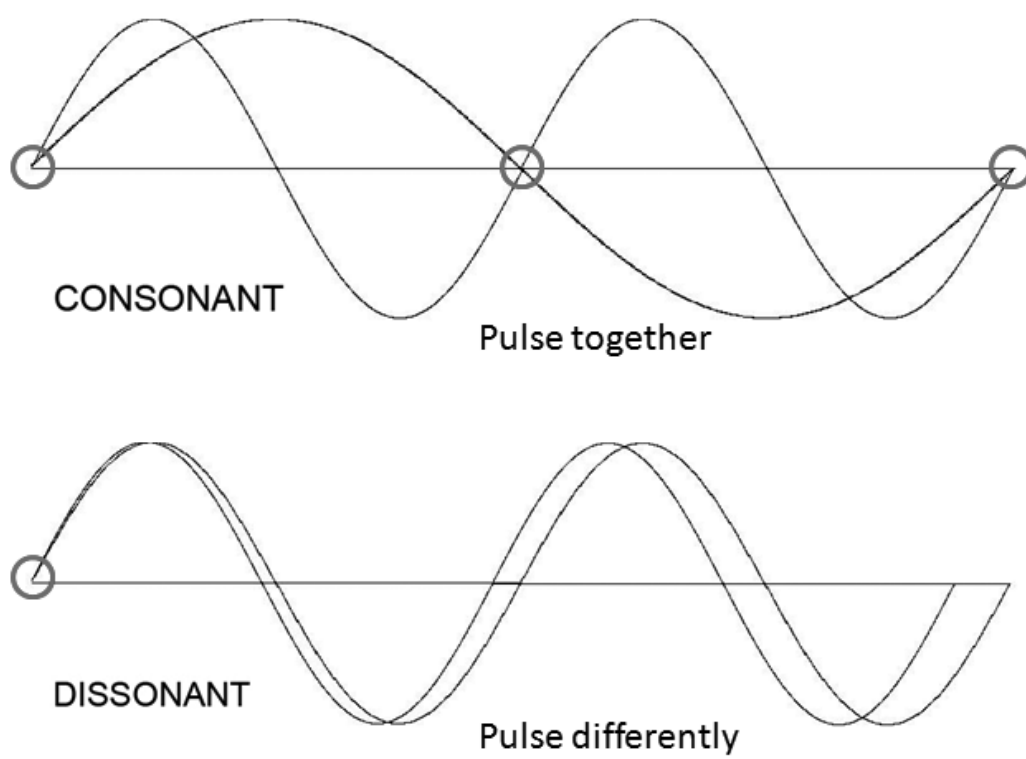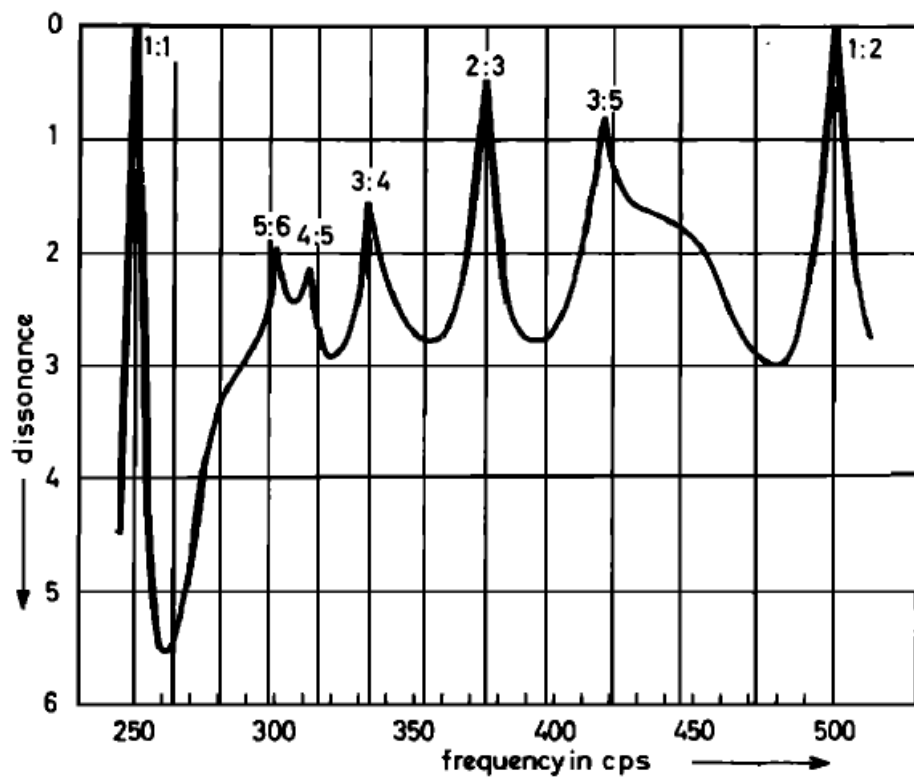
Figure 6-1.

Figure 6-2.

217

References

Baharloo, S., Johnston, P. A., Service, S. K., Gitschier, J., & Freimer, N. B. (1998). Absolute Pitch: An Approach for Identification of Genetic and Nongenetic Components. *American Journal of Human Genetics, 64*, 224-231.

Baptista, L. F., & Abs, M. (1983). Vocalizations. In: M. Abs, (Ed.) *Behavior and physiology of the pigeon* (pp. 309-325). New York, NY: Academic Press.

Blumstein, D. T., Davitian, R., & Kaye, P. D. (2010). Do film soundtracks contain nonlinear analogues to influence emotion? *Biology letters, 6*(6), 751–4.

Bregman, M. R., Patel, A. D., Gentner, T. Q. (2012). Stimulus-dependent flexibility in non-human auditory pitch processing. *Cognition 122*, 51-60.

Brooks, D. I., & Cook, R. G. (2010). Chord Discrimination by Pigeons. *Music Perception 27*(3), 183-196.

Cook, N. D., & Fujisawa, T. X. (2006). The Psychophysics of Harmony Perception: Harmony is a Three-Tone Phenomenon. *Empirical Musicology Review, 1*(2), 106-126.

Cynx, J. (1993). Auditory frequency generalization and a failure to find octave generalization in a songbird, the European starling (*Sturnus vulgaris*). *Journal of Comparative Psychology,107*, 140-146.

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

Fay, R.R. (1988) *Hearing in vertebrates: A psychophysics Databook*. Winnetka, Il: Hill-Fay Associates.

Ficken, M. S., Ficken, R. W., & Witkin, S. R., (1978). Vocal repertoire of the black-capped chickadee. *The Auk, 95*, 34-48.

Fitch, W. T., Neubauer, J., & Herzel, H. (2002). Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour, 63,* 407-418.

Goerdel-Leich, A., & Schwartzkopff, J. (1984). The auditory threshold of the pigeon (*Columba livia*) by heart-rate conditioning. *Naturwissenschaften, 71,* 98-99.

Henry, K. S., & Lucas, J. R. (2010). Auditory sensitivity and the frequency selectivity of auditory filters in the Carolina chickadee, *Poecile carolinensis*. *Animal Behaviour, 80*(3), 497–507.

Hoeschele, M., Cook, R. G., Guillette, L. M., Hahn, A. H., & Sturdy, C. B. (in press). Auditory same/different concept learning and generalization in black-capped chickadees (*Poecile atricapillus*). *PLoS ONE,* MS# PONE-D-12-17777R1.

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.

Hulse, S. H., & Cynx, J. (1985). Relative pitch perception is constrained by absolute pitch in songbirds (*Mimus, Molothrus*, and *Sturnus*). *Journal of*

Comparative Psychology, 99, 176-196.

Kallman, H. J. (1982). Octave equivalence as measured by similarity ratings. *Perception & Psychophysics, 32*, 37-49.

Langemann, U., Gauger, B., Klump, G. M. (1998). Auditory sensitivity in the great tit: Perception of signals in the presence and absence of noise. *Animal Behavior, 56*, 763-769.

Lohr, B. (2008). Pitch-related cues in the songs of sympatric mountain and black-capped chickadees. *Behavioural processes, 77*(2), 156–165.

Ludvig, E. A, & Spetch, M. L. (2011). Of black swans and tossed coins: is the description-experience gap in risky choice limited to rare events? *PLoS ONE, 6*(6), e20262.

Miyazaki, K. (1995). Perception of relative pitch with different references: some absolute-pitch listeners can't tell musical interval names. *Perception & Psychophysics, 57*(7), 962–70.

Pierce, J. R. (1999). The nature of musical sound. In (Ed) Deutsch, D. *The psychology of music.* San Diego, CA: Academic Press.

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

Ralston, J. V., & Herman, L. M. (1995). Perception and generalization of frequency contours by a bottlenose dolphin (*Tursiops truncatus*). *Journal of Comparative Psychology, 109*(3), 268–277.

Rasch, R., & Plomp, R. (1999). The Perception of Musical Tones. In D. Deutsch

(Ed.), The Psychology of Music 2nd Ed (pp. 215-264). San Diego,
California: Academic Press.

Richards, D. G., Wolz, J. P., & Herman, L. M. (1975). Vocal mimicry of
computer-generated sounds and vocal labeling of objects by a bottlenosed
dolphin, *Tursiops truncates. Journal of Comparative Psychology, 98*(1),
10–28.

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

Trehub, S. E. (2003). The developmental origins of musicality. *Nature
Neuroscience, 6*(7), 669–73.

Tyack, P. L. (2008). Convergence of calls as animals form social bonds, active
compensation for noisy communication channels, and the evolution of
vocal learning in mammals. *Journal of Comparative Psychology, 122*(3),
319-331.

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

Weisman, R. G., Njegovan, M. G., Williams, M., Cohen, J., & Sturdy, C. B.
(2006). The comparative psychology of absolute pitch. In E. A.
Wasserman & T. Zentall (Eds.), *Comparative Cognition: Experimental
Explorations of Animal Intelligence* (pp. 71-86). Oxford: Oxford
University Press.

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.

Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology: 805 General, 129*, 291-307.