The Influence of Sensory and Cognitive Consonance/Dissonance on Musical Signal Processing

Susan E. Rogers
Department of Psychology
McGill University, Montréal
June, 2010

A thesis submitted to McGill University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Experimental Psychology.

Table of Contents

Abstract

This thesis investigates possible origins of the distinction between consonant and dissonant auditory events and how persons with and without formal musical training judge the distinction. Two studies comprising six experiments used behavioral methods to explore perceptual and cognitive differences between musicians and nonmusicians. The first three experiments concern the qualitative assessment of auditory roughness — a primary component of sensory dissonance. The remaining three experiments concern short-term memory for musical intervals as distinguished by their properties of consonance and dissonance. An original contribution of this thesis is to quantify several differences that musical training confers upon both bottom-up (sensory-driven) and top-down (knowledge-driven) processing of musical sounds. These studies show that knowledge of a tonal hierarchy in a given culture cannot be reliably dissociated from the evaluation of a musical sound's features. Moreover, they show that robust, accurate auditory short-term memory exceeds the duration previously reported in the literature. These findings are relevant to theories of music perception and cognition, auditory short-term memory, and the psychophysical scaling of auditory event properties.

Résumé

Dans cette thèse nous étudions les origines possibles de la distinction entre événements auditifs consonants et dissonants, ainsi que la façon dont cette distinction est révélée dans le traitement auditif par les personnes ayant une formation musicale ou pas. Deux études comprenant six expériences ont employé des méthodes comportementales pour explorer les différences perceptives et cognitives entre musiciens et non musiciens. Les trois premières expériences concernent l'évaluation qualitative de la rugosité auditives — un composant élémentaire de la dissonance sensorielle. Les autres trois expériences concernent les différences de la mémoire à court terme entre les intervalles musicaux consonants et dissonants. Une contribution originale de cette thèse est de quantifier plusieurs différences que la formation musicale confèrent sur les traitements ascendants (conduits par les sensations) et descendants (conduits par les connaissances) des sons musicaux. Ces études montrent que les connaissances sur la hiérarchie tonale dans une culture donnée ne peuvent pas être fiablement dissociées de l'évaluation des attributs d'un son musical et que la durée de la mémoire auditive à court terme, qui est robuste et précise, excède celle rapportée précèdemment dans la littérature.

Acknowledgements

This thesis was written under the guidance of Dr. Daniel J. Levitin, who supervised the work reported in Chapter 3, and Dr. Stephen McAdams, who supervised the work in Chapters 2 and 4. I wish to thank Daniel Levitin for inviting me to study at McGill and allowing me to pursue a research question of my own. His perspectives on cognitive processing and memory mechanisms provided the inspiration for this work. I acknowledge his graciousness in permitting me to seek supervision in psychoacoustics outside of his laboratory. I am deeply grateful to Stephen McAdams for his mentorship, care, faith, patience, and attention to detail. His oversight of every aspect of this work was crucial to my development as a scientist and I am proud to be a member of his team. Dr. Evan Balaban provided the original question that led to this avenue of exploration. Professor Emeritus Al Bregman helped me to sharpen my focus, laying the foundation for future work.

Bennett Smith provided technical assistance by programming the experimental paradigms and instructing me on all things software-related. Karle-Philip Zamor provided additional instruction and assistance with computer and audio software and I am very appreciative. I also wish to thank Giovanna Locascio for her time and advice.

Undergraduate research assistants Sabrina Lytton and Mattson Ogg helped collect the data reported in Chapter 2. This work is indebted to the following graduate student and post doctoral colleagues for piloting the experiments and offering insightful comments: Anjali Bhatara, Bruno Gingras, Bruno Giordano, Bianca Levy, Georgios Marentakis, Nils Peters, Eve-Marie Quintin, Finn Upham, and Michel Vallières.

My entire scholastic experience was made possible through the success I enjoyed in partnership with Barenaked Ladies and so I thank Jim Creeggan, Kevin Hearn, Steven Page, Ed Robertson, and Tyler Stewart. I am grateful to have the support and encouragement of my friend and dean Stephen Croes and my colleagues and students at Berklee College of Music.

Finally, I wish to express an endless gratitude for the love and inspiration that I found in my beloved Boston Terrier Gina, that I have kept in Tommy Jordan, and that has been renewed by Matthew McArthur.

Preface

Manuscript-based Thesis

The present work is submitted in the form of a manuscript-based thesis in accordance with McGill University's Graduate Thesis Guidelines for dissertations. A manuscript-based thesis consists of three papers formatted for submission to a peer-reviewed journal. The guidelines specify that these papers must form a cohesive work with a unifying theme representing a single program of research. Chapters must be organized in a logical progression from one to the next and connecting texts must be provided as segues between chapters. In accordance with the guidelines, the present thesis consists of three chapters of original research in journal-submission form. Chapter 2 is in preparation for submission to the *Journal of the Acoustical Society of America.* Chapters 3 and 4 are a two-part submission for the *Journal of Experimental Psychology: Learning, Memory, and Cognition.* An introductory chapter is included with a literature review and a discussion of the rationale and objectives of the research. A concluding chapter summarizes the work and describes future directions.

In accordance with the Guidelines, a description of the contributions from each chapter's co-authors, including myself, is submitted below.

Contributions of Authors

<u>Chapter 1</u>: Introduction and Overview of Literature
Author: Susan E. Rogers

I am the sole author and am responsible for all content; Drs. McAdams and Levitin read drafts and made suggestions.

<u>Chapter 2</u>: Roughness Evaluations of Just- and Micro-tuned Dyads from Expert and Nonexpert Listeners.
Authors: Susan E. Rogers and Stephen McAdams
Contributions:

- First author — Rogers: I conducted the literature review, prepared the stimuli, tested the participants (with the cooperation of two undergraduate research assistants), analyzed all data, researched and implemented the audio analyzers, prepared all figures and tables, wrote the manuscript, and presented this work a conference.
- Co-author — McAdams (thesis co-advisor): provided the need for the subjective data and the original research question, summarized and advised the statistical analysis. Earlier work by Dr. McAdams relates to the theoretical issues discussed in this paper. Dr. McAdams gave counsel and direction at all stages and feedback on drafts of the manuscript. Dr. Levitin read drafts and made suggestions.

<u>Chapter 3</u>: Short-term Memory for Consonant and Dissonant Pure-Tone Dyads
Authors: Susan E. Rogers, Daniel J. Levitin, and Stephen McAdams
Contributions:

- First author — Rogers: I conducted the literature review, conceived of and designed the experiment, prepared the stimuli, tested the participants, analyzed the data, prepared all figures and tables, wrote the manuscript, and incorporated the contributions from co-authors, and presented this work at conferences and an invited lecture. I was the first author on summaries of this work published in *Canadian Acoustics* (2007) and the *Journal of the Audio Engineering Society* (2007).

- Co-author — Levitin (thesis co-advisor): gave counsel and direction at all stages and feedback on drafts of the manuscript. Earlier work by Dr. Levitin on memory for musical pieces provided inspiration for this study.
- Co-author — McAdams (thesis co-advisor): directed the statistical analysis, gave counsel and direction at all stages and feedback on drafts of the manuscript.

Chapter 4: Short-term Memory for Consonant and Dissonant Complex-Tone Dyads — Just- and Microtuned.
Authors: Susan E. Rogers, Stephen McAdams, and Daniel J. Levitin
Contributions:

- First author — Rogers: I conducted the literature review, prepared the stimuli, modified the experimental paradigm from the previous investigation, tested the participants, analyzed the data, prepared all figures and tables, wrote the manuscript, incorporated the contributions from co-authors, and presented this work at conferences.
- Co-author — McAdams (thesis co-advisor): suggested the idea to test memory for microtuned dyads and described their construction, gave advice at all stages of the study and provided feedback on drafts of the manuscript.
- Co-author — Levitin (thesis co-advisor): gave counsel and feedback on drafts of the manuscript.

Chapter 5: Summary and Conclusions
Author: Susan E. Rogers

I am the sole author and am responsible for all content; Drs. McAdams and Levitin read drafts and made suggestions.

CHAPTER 1

Introduction

Auditory events have both a physical form (an acoustic waveform) and a meaningful function (conveying information about the environment). Music-making requires the manipulation of auditory form and function to achieve an emotional end. Humans choose chords and musical instrument timbres to affect an (intended) function in composition and performance. Objective properties of frequency, amplitude, phase, and temporal delay are balanced against subjective properties such as whether a sound is euphonic or consonant versus suspenseful or dissonant. *Consonance* versus *dissonance* is a continuum that can be discussed in two ways: according to the sensation in the auditory periphery induced from the interaction of multiple tones and according to the tones' music-theoretical representation in a cognitive schema. It is the dual nature of consonance and dissonance — sensory and cognitive — that is the focus of this work.

The nature of consonance and dissonance has presented opportunities for interdisciplinary study for generations of scholars. Music-theorists (Cazden, 1945, 1980; Rameau, 1722/1971; Sethares, 1998; Tenney, 1988) and scientists (Helmholtz, 1885/1954; Kameoka & Kuriyagawa, 1969a, 1969b; Plomp & Levelt, 1965; Schellenberg & Trehub, 1994a, 1994b, 1996; Seashore, 1918; Stumpf, 1898; Terhardt, 1974a, 1974b, 1984; Tramo, Cariani, Delgutte, & Braida, 2003; Wild, 2002) have investigated it within cultural, philosophical, mathematical, perceptual, cognitive, and neurophysiological frameworks. Early work focused on linking the sensation of sound to its acoustical parameters and to the mechanics of the mammalian auditory system (e.g., Greenwood, 1961; Helmholtz, 1885/1954; Kameoka & Kuriyagawa, 1969a, 1969b; Plomp, 1967; Plomp & Levelt, 1965; Plomp & Steeneken, 1967). Music theorists and scientists adopted each other's ideas and findings. Explorations in the late twentieth century narrowed the focus and established that the relationship between what psychoacousticians called "sensory (or tonal) consonance" and what music theorists called "musical consonance" was not perfectly parallel (Terhardt, 1984; Tramo et al., 2003).

Technological advances in the late twentieth century have made new methods available to study individual differences in consonance/dissonance (C/D) perception. Neuroimaging and brain functional mapping tools provide data on the ways in which musical training, attention, expectancies, exposure, and implicit and explicit learning processes shape the perception of musical sounds. Yet there remain many outstanding questions. How and when in the signal processing chain are the physical features of a sound transduced into the psychological percept of a musical interval? What is the duration of the temporal window during which a chord's acoustical features and tonal function cannot be dissociated? How does familiarity with a given tonal music system affect whether a chord is perceived as consonant or dissonant? How is the C/D distinction reflected in other cognitive processes, such as memory?

This thesis contributes two perspectives to the body of knowledge on the C/D distinction, and thereby to theories of auditory perception and cognition. The first perspective concerns the perception of form by collecting judgments of auditory roughness (a critical component of dissonance). The work described in Chapter 2 advances the psychoacoustician's understanding of C/D by segregating musical experts from nonexperts, controlling sources of signal distortion and error that contaminated earlier findings, analyzing results with new statistical methods, and accounting for familiarity with musical tonal systems.

The second perspective concerns the idea of "natural intervals" — the belief that consonance endows a signal with innate cognitive processing advantages over dissonance (Burns & Ward, 1982; Schellenberg & Trainor, 1996; Schellenberg & Trehub, 1996). The experiments described in Chapters 3 and 4 presented musical intervals to listeners in a short-term memory paradigm to learn whether some musical intervals were mentally more robust than others. These chapters comprise an integrated series of three experiments that are the first of their type in C/D research. Musicians and nonmusicians performed a novel/familiar memory task using intervals that varied along the two axes of sensory and cognitive consonance and dissonance. The observed differences in memory strength and fragility provided clues to the nature of the original signal processing. The findings inform theories of nonlinguistic auditory memory as well as the centuries-old discussion on the origins of the C/D distinction.

Overview of the Literature

A brief history of consonance and dissonance.

The twin phenomena of consonance and dissonance have intrigued the scientist/philosopher since Pythagoras introduced it to the Greeks in the 6th century B.C.E. (Cazden, 1958, 1980; Tenney, 1988). Gottfried Leibniz, the 17th century co-inventor of infinitesimal calculus, linked consonance to the Beautiful and believed that humans unconsciously calculate the frequency ratios that describe musical intervals. According to Bugg's (1963) interpretation of Leibniz, the soul performs the calculations (albeit oblivious to the math) and deems only the octave and the perfect 5th to be truly consonant. Leonhard Euler, an 18th century mathematician who advanced geometry and calculus, suggested that simple-ratio intervals appeal to the human need for order and coherence and thus cause the corresponding sensation of agreeableness (Burdette & Butler, 2002). In the 19th century, philosopher Arthur Schopenhauer believed that the harmony necessary for perfection in music was a copy of our animal nature and "nature-without-knowledge" (1819/1969, p. 154). Helmholtz (1873/1995) agreed, remarking, "the mind sees in [harmony and disharmony] an image of its own perpetually streaming thoughts and moods."

Helmholtz (1885/1954) formalized the observations of Pythagoras by linking C/D to the physical properties of sounds. Periodic musical tones and speech sounds have partial tones that correspond to the harmonic series, i.e., overtones related to the fundamental frequency f_0 by integer multiples nf_0. The integer ratio describing a dyad identifies the number of its coincidental (or nearly so) partials. Small-integer ratio dyads such as octaves (1:2) and perfect 5ths (2:3) have few or no narrowly separated, noncoincidental partials. Large-integer ratio dyads such as minor 7ths (9:16) and Major 2nds (8:9) feature many noncoincidental partials, argued to be the source of their relative dissonance (p. 194). Helmholtz believed that dissonance could be predetermined given that it was the property of the absolute frequency differences between tones. His writings assumed that all astute listeners judged dissonance the same way.

Early twentieth century researchers assembled a corpus of data on the evaluated C/D of musical chords (summarized in Chapter 2). Seminal work on the subjective assessment of consonance versus dissonance was conducted during this period by Plomp and Levelt (1965) and Kameoka and Kuriyagawa (1969a, 1969b). Their research advanced the topic through systematic examinations of listeners' responses to pure-tone and complex-tone dyads across a wide range of fundamental frequencies. They extended the work of Helmholtz (1885/1954) by describing C/D as a product of the *relative*, not absolute, frequency difference between two tones. Their findings greatly advanced understanding of the association between acoustical phenomena and the physical behaviors and constraints of the human hearing mechanism. Nevertheless, methodological issues remained for interpreting C/D. Participants in these studies were pre-selected for their abilities to differentiate consonance from dissonance *as defined by the researchers*. By so doing, both sets of data may have inadvertently excluded participants representative of the normal population. In addition, the adjectives used to describe "consonance" and "dissonance" in Dutch (Plomp & Levelt, 1965) and Japanese (Kameoka & Kuriyagawa, 1969a) may not have described exactly the same phenomena. The fact that the understanding of particular adjectives and some training were necessary prior to C/D assessments revealed a need for more precise definitions of the terms.

Terhardt (1984) helped resolve ambiguities by codifying the terms used in the C/D discussion, based on findings of Kameoka and Kuriyagawa (1969a, 1969b) and Terhardt and Stoll (1981). Terhardt argued that *sensory C/D* referred to the presence of one or more potentially "annoying" factors: roughness, loudness[1], sharpness (the loudness weighted spectral center computed on a physiological dimension — the Bark scale — related to the mapping of frequency into the auditory system), and tonalness (how well tones fuse into a single percept or provide a strong pitch cue). Unlike sensory C/D, *musical C/D* was confined to musical tones. It referred to an interval's sensory C/D plus its degree of *harmony.* Terhardt (1984, 2000) defined harmony as tonal affinity plus the ease with which the fundamental note or root pitch may be extracted from a chord.

Models of C/D based on the harmonic series and the contribution from partial roughnesses dominated the early literature (Ayers, Aeschbach, & Walker, 1980; Butler & Daston, 1968; Geary, 1980; Greenwood, 1991; Guernsey, 1928; Helmholtz, 1954/1885; Kameoka & Kuriyagawa, 1969a, 1969b; Malmberg, 1918; Plomp & Levelt, 1965; Plomp & Steeneken, 1968; Regnault, Bigand, & Besson, 2001; Sethares, 1993, 1998; Terhardt, 1974a, 1974b; Van de Geer, Levelt, & Plomp, 1962). Unfortunately, methodological inconsistencies made cross-experimental comparisons difficult or in some cases impossible. Each decade's researchers used the technologies available at the time, but nevertheless signal path distortions and lack of control due to unreliable modes of signal generation and/or reproduction contributed a nontrivial amount of error. It is germane to this thesis that most early work collected data from a homogeneous sample and in many cases ignored or failed to report participants' levels of musical training. (Chapter 2 lists exceptions where data from two groups of participants — musicians and nonmusicians — were collected and analyzed separately.)

"Natural" musical intervals.

So called "natural intervals" (Burns & Ward, 1982) are those defined by small-integer frequency-ratio relationships. The idea that the human brain has adapted to favor some musical intervals or otherwise regard them as innately easier to process is an important concept for the work of this thesis. Explanations for the link between consonance and small-integer frequency-ratio relationships have taken psychoacoustical and neurobiological approaches.

Evidence for the existence of natural intervals evolved from the work on the cognition of tonality — the affinity of tones. The influence of frequency-ratio size to C/D perception was shown to extend beyond the physical correlates in the cochlea. The relative C/D of horizontal or *melodic intervals* — tones played sequentially — depends upon factors that include the frequency-ratio relationship between the tones (Krumhansl & Kessler, 1982). Maps of the relative C/D of melodic intervals have provided evidence for internalized tonal schemata that influence the perception of musical events, even when those events are presented outside of a melodic context.

Another explanation for "natural intervals" relates to the human propensity for speech acquisition. Human utterances are the most salient naturally occurring periodic sounds in our personal and collective environments. As in consonant dyads, harmonic energy in speech sounds is distributed at simple frequency ratios like 1:2 and 2:3 (Ross, Choi, & Purves, 2007; Schwartz, Howe, & Purves, 2003). The frequency of occurrence of small-integer ratio acoustical energy distributions in speech sounds is argued to quickly train (Schwartz et al., 2003; Terhardt, 1974b) or predispose (Schellenberg & Trehub, 1996; Trainor, Tsang, & Cheung, 2002) the human auditory system to regard simple, small-integer ratio intervals as more "natural" and thus easier to process than more complex ratio intervals.

Subjective assessments of the C/D of musical intervals have yet to be explored in a standardized, "culture-independent" way (Patel, 2008, p. 90). Researchers have focused their attention recently on the neuroelectrical and neurovascular sources of the C/D distinction (reviewed in Chapter 2) with the aim of uncovering universal principles underlying the phenomena. Theories of C/D rooted in neurological processes note that closely spaced partials causing certain mechanical interactions in the cochlea lead to qualitatively distinct representation in auditory neural coding (Tramo et al., 2003). Sounds perceived as rough or sensory dissonant give rise to neural firing patterns in the auditory nerve and brainstem that are readily distinguished from firing patterns caused by smoother, sensory consonant sounds (Fishman et al., 2001; McKinney, Tramo, & Delgutte, 2001). The all-order auditory nerve firing pattern corresponding to evaluated consonance has also been found to correlate positively with the perceived salience of a sound's pitch (Bidelman & Krishnan, 2009; Cariani, 2004). These findings explain consonance preference (as defined by an interval's integer-ratio complexity) as a product of innate auditory system processing constraints that favor small-integer ratio musical intervals.

Should "natural musical intervals" be regarded by the brain as categorically distinct, there is no a priori reason to believe that the distinction would appear in a nonmusical cognitive task such as short-term memory. The work presented in this thesis hopes to contribute to Peretz and Zatorre's (2005) call to "determine at what stage in auditory processing … the computation of the perceptual attribute of dissonance is critical to the perceptual organization of music."

Auditory short-term memory.

The literature on auditory memory is concerned primarily with aural language stimuli. Literature on auditory memory that excludes mnemonic pathways through lexical or visual associations is sparse, chiefly for practical reasons. It is a safe assumption, for example, that memory for the melody associated with "Happy Birthday to You" is recalled along with the words and sights that usually accompany hearing it. So-called "genuine auditory memory" for a stimulus or a task excludes nonauditory forms of coding such as visual or linguistic associations (Crowder, 1993). Typically only those rare individuals with absolute pitch (AP) perception — the ability to immediately label or produce a specific pitch chroma in the absence of an external reference — have the option of encoding a single tone's active neural trace by an attribute other than its pitch (Levitin & Rogers, 2005). By immediately and accurately identifying its pitch chroma, AP possessors can encode the signal with a label or its visual equivalent on a musical staff, (presumably) increasing the chances of its later retrieval. Most humans lack this ability and thus are capable of exhibiting genuine auditory memory free from the confounds of verbal labels for both familiar and unfamiliar musical intervals presented in isolation, outside of a melodic or tonal context.

The conscious perception of an auditory stimulus is the by-product of its initial representation (Crowder, 1993; Näätänen & Winkler, 1999). Differential memory retention can provide clues to the underlying differences in mental organization caused by stimulus type. If sensory and/or cognitive C/D encoding recruits anatomically distinct pathways, differential memory may mirror the distinction. This would not be due to separate memory stores necessarily but due to the fact that the perceptual events were initially encoded or otherwise processed differently (Crowder, 1993).

If memory for one set of auditory stimuli is more accurate than for another, the characteristics of the set should reflect categorical distinctions between them, innate or otherwise. Thus differential memory for consonance and dissonance could reflect a hierarchal categorization scheme that automatically places consonant (or dissonant) intervals in a less accessible cognitive position. It could also indicate differential rates of forgetting (Tierney & Pisoni, 2004; Wickelgren, 1977), driven by either Gaussian or deterministic auditory feature decay (Gold, Murray, Sekuler, Bennett, & Sekuler, 2005). Where no discrepancy is found, this suggests that although the brain recognizes a physical distinction between consonant and dissonant dyads (Blood, Zatorre, Bermudez, & Evans, 1999; Brattico et al., 2009; Fishman et al., 2001; Foss, Altschuler, & James, 2007; McKinney et al., 2001; Minati et al., 2009; Passynkova, Neubauer, & Scheich, 2007; Regnault, Bigand, & Besson, 2001; Tramo et al., 2003), it regards these events as cognitively equivalent.

Short-term memory (STM) is cognitively easy. Unlike working memory, STM does not require mental operations such as the application of a rule or the transformation of items (Engle, Tuholski, Laughlin, & Conway, 1999). Its neural representation is fragile in contrast to representations in long-term storage because STM is quickly degraded by time and interference from new incoming items (Cowan, Saults, & Nugent, 1997; Crowder, 1993; Keller, Cowan, & Saults, 1995; Näätänen & Winkler, 1999; Winkler et al., 2002). Accurate STM reflects a level of processing that ranges from conscious knowing (i.e., remembering or recollecting) to unconscious perceptual fluency — a processing level more information-driven than simply guessing (Jacoby, 1983; Wagner & Gabrieli, 1998). In instances where perceptual fluency is the only option for processing, i.e., when the participant has no conceptual knowledge of the stimulus's meaning or function, the similarity of successive stimuli has a strong effect on STM recognition accuracy (Stewart & Brown, 2004).

The experiments reported in Chapters 3 and 4 of this thesis modified a novel/familiar experimental protocol from Cowan, Saults, and Nugent (1997) that tested STM for single pitches. Tasks of this type require a listener to compare the features of a new sound in STM against the features of recently stored sounds (Nosofsky & Zaki, 2003). A correct answer on a familiar trial results if some property of the stimulus exceeds a criterion threshold for a correct match. For novel trials the stimulus properties have to fall below the criterion value (Johns & Mewhort, 2002; Stewart & Brown, 2005). This kind of processing makes a novel/familiar recognition task useful for determining categorization schemes because correct rejections of novel stimuli indicate that psychological lines have been drawn around stimulus sets (Johns & Mewhort, 2002; Nosofsky & Zaki, 2002). Novel/familiar recognition taps implicit memory for an object and tests the participant's ability to decide whether or not a trace was left by a recently encountered object or event (Petrides & Milner, 1982).

When guessing is the only strategy that can be used, the rate of guessing is revealed by the proportion of false alarms. Analysis methods developed from Signal Detection Theory provide the researcher with information on the participant's "decision axis" — an internal standard of evidence for or against an alternative (Macmillan & Creelman, 2005; Wickens, 2002, p. 150). The descriptive statistic *d'* reflects the magnitude of the participant's decision ability and thus the "strength of evidence" (Macmillan & Creelman, 2005; Pastore, Crawley, Berens, & Skelly, 2003).

How long does an uncategorized sound (i.e., apropos of nothing or significant in no larger context) remain in STM? In cases where alternate coding strategies (e.g., rehearsing, visualizing, labeling) are ruled out by the stimulus or task, STM for a single pitch will fade in less than 30 s. Winkler et al. (2002) showed that memories for single pitches were available after 30 s of silence, but only when the pitches were encoded in the context of a regular, repetitive sequence (a pitch train). These researchers concluded that *acoustic regularity* causes a single pitch to be encoded as a permanent record in long-term memory (LTM). For comparison, they also conducted a simple two-tone task — one in which there was no stimulus regularity. In the absence of regularity, only one of their participants was able to retain a single pitch in STM after 30 s of silence.

Other studies have failed to demonstrate persistent STM for single pitches beyond 30 s, although it must be noted that they did not extend their retention periods beyond that time (Cowan et al., 1997; Dewar, Cuddy, & Mewhort, 1977; Kærnbach & Schlemmer, 2008; Keller et al., 1995; Massaro, 1970; Mondor & Morin, 2004).

One possibility is that seeing performance drop to near chance at moderate retention periods (as did the work of this thesis for certain classes of dyads) discouraged researchers from exploring beyond 30 s.

Auditory roughness in sensory dissonance.

The definition of auditory roughness describes a degree of signal modulation in the range of 15-300 Hz (Zwicker & Fastl, 1991) that listeners typically report as "unpleasant" or "annoying" (Terhardt, 1974b). Like pitch and loudness, it is a subjective property, represented throughout the auditory system from the cochlea to cortical areas (De Baene, Vandierendonck, Leman, Widmann, & Tervaniemi, 2004; Fishman, Reser, Arezzo, & Steinschneider, 2000; Greenwood, 1961b; Plomp & Levelt, 1965). Its perception contributes to the sensory dissonance of musical sounds and it is linked to the feeling of musical tension (Pressnitzer, McAdams, Winsberg, & Fineberg, 2000). Evaluating auditory roughness requires listeners to detect, attend to, and label the perception, and that can be difficult for some listeners, in some circumstances. Researchers report inconsistent roughness assessment in the absence of thoughtful experimental design (Kreiman, Gerratt, & Berke, 1994; Prünster, Fellner, Graf, & Mathelitsch, 2004; Rabinov & Kreiman, 1995).

Helmholtz (1885/1954) wrote that for musical sounds, roughness and slower fluctuations (termed *beating*) could readily be heard, but "(t)hose who listen to music make themselves deaf to these noises by purposely withdrawing attention from them" (p. 67). Assuming that this is true, a musical interval's degree of roughness, imbued as it is with tonal (Krumhansl, 1991) and emotional (Balkwill & Thompson, 1999; Pressnitzer, et al., 2000) associations in a given musical culture, could be expected to elicit a range of evaluative responses from listeners in a psychophysical scaling task. The quality and quantity of a listener's musical experiences should mediate his or her sensitivity to roughness components and subsequently, to sensory and cognitive dissonance. Early influential studies of evaluated sensory C/D may have underappreciated the role of individual differences in interval quality judgments (Kameoka & Kuriyagawa, 1969b; Plomp & Levelt, 1965; Van de Geer, Levelt, & Plomp, 1962). Accounting for these differences refines the understanding of sensory C/D processing.

Psychophysical scaling of musical sounds.

"Object constancy is fundamental to perception and attribute scaling is not fundamental" (Lockhead, 2004, p. 267). Lockhead offered this theoretical viewpoint to argue that humans did not evolve for the purpose of abstracting single elements from an object, and thus, "there is no a priori reason to expect people to be good . . . sound meters" (p. 267). Serving as a meter to measure a single element, he argued, would disrupt the listener's goal of identifying the object associated with the element.

The argument that perceivers find it naturally difficult to attend to isolated elements is supported in psychophysical scaling tasks involving *related elements that change* in a moving object (Lockhead, 2004; Zmigrod & Hommel, 2009), as is the case for sounds produced by musical instruments and vocal chords. Indeed it is attention to the unfolding changes across elements comprising frequency spectra and temporal envelope that permit the listener to identify a sound's source (Dowling &

Harwood, 1986). Given that he is likely to attend to the relations among sonic elements, psychophysical scaling of elements in musical sounds are predicted to occur in the context of the perceiver's knowledge of sounds having similar relations, rather than *absolutely* in terms of elements in the experimental set (Lockhead, 2004; Ward, 1987). Perceived elements of a dyad, such as roughness cues, are therefore confounded with implicit knowledge of the interval's role and frequency of occurrence in the listener's musical culture. This implicit musical knowledge is linked to the fact that the harmonic relationship between a dyad's two tones mirrors its distribution in Western tonal musical compositions (Krumhansl, 1990, 1991). (For example, perfect consonant intervals such as octaves and perfect 5ths are more prevalent in music than dissonant intervals such as minor 2nds and tritones; Cambouropoulos, 1996; Cazden, 1945; Dowling & Harwood, 1986.) Uncertainty over what to expect within the context of a psychophysical scaling experiment, i.e., unfamiliarity with the items being judged, diminishes the perceiver's capacity to imagine where his or her judgments reside on the "true" scale of all possible items, leading to less-reliable ratings (Lockhead, 2004; Ward, 1987). Thus listeners relatively unfamiliar with assessing musical intervals in the absence of a musical, tonal context could be expected to show less agreement and poorer rating consistency than those listeners experienced in regarding intervals as items in a known or familiar set.

Rationale and Research Objectives

The work presented in these chapters makes a unique contribution to fundamental topics in psychoacoustics and auditory memory in part by accounting for the recently known processing advantage conveyed by musical expertise in the perception and cognition of musical intervals. Each of the studies reported here used strict experimental protocols, rigorously controlled and calibrated audio recording and reproduction tools, and methods of statistical analysis not used in previous studies of these types. Each experiment was conducted using three unique stimulus sets to control for ecological validity and exposure to Western tonal musical materials. In addition, each engaged a large number of participants to strengthen the power of the findings.

These studies advance knowledge of music perception and cognition by showing the extent to which musical expertise moderates the dual auditory processing streams of sensory form (bottom-up) and conceptual knowledge (top-down). The findings contribute to theories of nonlinguistic auditory memory and signal processing and assist in the development of new audio tools that better reflect the range of human perceptual abilities.

Three manuscript-style chapters form the body of this thesis. The objectives of each chapter are summarized as follows:

Chapter 2 reports on the expert and nonexpert assessment of auditory roughness — a primary component of sensory dissonance. Musical expertise has gone unreported in most of the behavioral data on evaluated sensory C/D, yet recent neurophysiological reports show that expert listeners (those with years of formal musical training) process auditory signals differently than nonexperts. This three-part experiment segregated the two populations and adopted a more controlled design protocol than previously used in evaluations of sensory dissonance, eliminating or reducing sources of error that confounded earlier studies of this type. The work controlled for exposure to musical intervals by including microtuned dyads — mistuned from the familiar Western standard by a quartertone — that are only rarely found in Western music. Ratings were compared both within and across participant groups. The application of statistical tests new to sensory C/D work provided a clearer insight into the variance and stability of internal standards found in the psychophysical scaling of auditory roughness. Ratings were also compared to objective ratings from two auditory roughness analyzers and two sensory C/D models in the literature to learn the extent to which musical expertise was assumed by their designs.

Chapter 3 explores a cognitive basis for the distinction between the sensory and cognitive properties of musical intervals outside of a musical, tonal context. Musicians and nonmusicians listened to sequentially-presented pure-tone dyads in a STM recognition paradigm. Dyads spanned a range of sensory and cognitive C/D so that differential memory, if observed, could provide evidence for or against the argument for "natural musical intervals." Each dyad was presented twice, separated by a varying number of intervening stimuli. Participants responded by indicating whether the dyad was novel or had been recently heard. Mapping the time course of STM for musical intervals provides information on auditory feature availability and processing differences for dyads according to classification and type between musicians and nonmusicians.

Chapter 4 expands the study of STM for pure-tone dyads by exploring memory for complex-tone dyads. The work aimed to discern how relationships among harmonic partials contribute to dyad robustness against decay over time and interference from incoming sounds. In two studies, listeners of Western tonal music performed the novel/familiar recognition memory task reported in Chapter 3. Stimuli featured either commonly known just-tuned dyads or unfamiliar microtuned dyads (mistuned from common musical intervals by a quarter tone). Microtuned intervals, rare in the Western tonal system, provided a control for different levels of musical experience between expert and nonexpert listeners. The use of these dyads also provided a necessary constraint on STM processing by reducing or eliminating its reliance on categorized exemplars from long-term storage to successfully perform the task.

Chapter 5 reviews and integrates information presented in the previous four chapters and develops the conclusions drawn from this research. New proposals for future work are introduced and discussed in terms of their potential contributions to areas of psychology.

The research reported herein addresses the perceptual and cognitive distinctions between consonance and dissonance with the aim of advancing understanding of how auditory signals are processed and how individual differences affect their interpretation.

[1]Terhardt's later writing (2000) omitted loudness as a component of sensory dissonance, although Kameoka and Kuriyagawa (1969a,b) provided evidence for the association.

CHAPTER 2

Roughness Ratings for Just- and Micro-Tuned Dyads from Expert and Nonexpert Listeners

Susan E. Rogers and Stephen McAdams

Author affiliations:
Susan E. Rogers [a]
Department of Psychology and Center for Interdisciplinary Research in Music, Media, and Technology, McGill University

Stephen McAdams
Schulich School of Music and Center for Interdisciplinary Research in Music, Media, and Technology, McGill University

[a)] Department of Psychology
McGill University
1205 Dr. Penfield Avenue, 8th floor
Montreal, Quebec, CA H3A 1B1
Electronic mail: susan.rogers@mail.mcgill.ca

Abstract

To explore the extent to which musical experts and nonexperts agreed, listeners rated pure- and complex-tone dyads (two simultaneous pitches) for auditory roughness — a primary component of sensory dissonance. The variability of internal roughness standards and the influence of musical training on roughness evaluation were compared along with objective ratings from two auditory roughness analyzers. Stimulus sets included dyads in traditional Western, just-tuned frequency-ratio relationships as well as microtuned dyads — mistuned from the familiar Western standard by a quartertone. Within interval classes, roughness ratings for just-tuned dyads show higher rater consistency than ratings for microtuned dyads, suggesting that knowledge of Western tonal music influences perceptual judgments. Inter-rater reliability (agreement among group members) was poorer for complex-tone dyads than for pure-tone dyads, suggesting that there is much variance among listeners in their capacity to isolate roughness components present in harmonic partials. Pure-tone dyads in frequency ratio relationships associated with musical dissonance received higher roughness ratings than those in musical consonance relationships from musical experts, despite the absence of signal elements responsible for the sensation. Complex-tone, just-tuned dyad ratings by experts correlated more closely with a theoretical model of Western consonance than did those of nonexperts (Hutchinson & Knopoff, 1978). Roughness ratings from audio analyzers correlated better with just-tuned than with micro-tuned dyad ratings. Accounting for sources of listener variability in roughness perception assists in the development of audio analyzers, music perception simulators, and experimental protocols, and aids in the interpretation of sensory dissonance findings.

keywords: auditory roughness, auditory perception, sensory dissonance, sensory consonance, microtuning

I. INTRODUCTION

"The ability to judge the quality of two-clang as in consonance is now the most general test of sensory capacity for musical intellect" (Seashore, 1918). Seashore regarded some individuals to be more sensitive than others in assessing the qualities of musical sounds and believed this sensitivity was innate. Since his time, the effect of musical training has been invisible in much of the 20th century data on the sensory (physiological) dissonance of dyads — two simultaneous pitches. Seminal research and writings on the perception of sensory dissonance has for the most part omitted the musical expertise of the listener as a covariate (e.g., Ayres, Aeschbach, and Walker, 1980; Butler and Daston, 1968; DeWitt and Crowder, 1987; Guirao and Garavilla, 1976; Kameoka and Kuriyagawa, 1969a, 1969b; Plomp, 1967; Plomp and Levelt, 1965; Plomp and Steeneken, 1967; Schellenberg and Trainor, 1996; Terhardt, 1974a; Viemeister and Fantini, 1987). (Exceptions include Geary, 1980; Guernsey, 1928; Malmberg, 1928; Pressnitzer, McAdams, Winsberg, and Fineberg, 2000; Van de Geer, Levelt, and Plomp, 1962; and Vos, 1986.) The prevailing assumption has been that outside of a musical, tonal context, listeners could attend strictly to the physical form of a musical interval. Because the sensory *consonance/dissonance* (hereafter abbreviated C/D) distinction originates in the auditory periphery, any meaning implied in the relationship of a dyad's frequency components could be effectively ignored (Terhardt, 1974a).

This decade's neurophysiological findings have overturned that assumption by demonstrating that in passive listening tests using isolated dyads or chords, adults with musical training often process musical intervals in different brain regions, at different processing speeds, and with greater acuity than nonmusicians (Brattico *et al.* 2009; Foss, Altschuler, and James, 2007; Minati *et al.* 2009; Passynkova, Neubauer, and Scheich, 2007; Regnault, Bigand, and Besson, 2001; Schön, Regnault, Ystad, and Besson, 2005). Therefore, the need exists for a new perceptual assessment of the sensory dissonance of dyads, acknowledging the relative contribution from diverse capacities for auditory discrimination. This study asks how well expert and nonexpert listeners agree in their judgment of auditory roughness — a primary component of sensory dissonance — to explore the variance of internal roughness standards and the extent to which musical training influences sensory dissonance perception.

A. Auditory roughness and sensory dissonance

The term 'roughness' is used by speech pathologists when describing a hoarse, raspy vocal quality (Kreiman, Gerratt, and Berke, 1994) and by acousticians when describing a degree of signal modulation in noise or in complex tones (Daniel and Weber, 1997; Hoeldrich, 1999). In its simplest form, a sensation of auditory roughness can result when a tone or noise is amplitude- or frequency-modulated at rates ranging from about 15 to 300 cycles per second (Zwicker and Fastl, 1990). As the modulation rate increases to the point where the human auditory system can no longer resolve the changes, modulation depth is reduced along with the roughness sensation (Bacon and Viemeister, 1985). Fluctuations slower than 15 Hz are termed *beating* (where two tones are perceived as one tone with audible loudness fluctuations), and very slow fluctuations below 4 Hz are not perceived as rough (Zwicker and Fastl, 1990).

With regards to musical intervals, psychoacoustic researchers label 'roughness' as a particular sound quality contributing to *sensory dissonance* — a measure of a chord's harshness or annoyance that is the opposite of *sensory consonance* — a measure of its tonal affinity or euphoniousness. Roughness is frequently discussed as synonymous with 'unpleasantness,' although the strength of this association warrants further investigation. At least one study found roughness to be only moderately unpleasant as compared to the qualities of 'sharpness' and 'tenseness' (Van de Geer *et al.* 1962). Since the early 17th century, music theorists have linked consonance to the absence of roughness, and perceptual data have supported this idea (Kameoka and Kuriyagawa, 1969a; Plomp and Levelt, 1965; Plomp and Steeneken, 1968; Van de Geer *et al.* 1962; see Tenney, 1988, for an historical review).

Roughness can be difficult for listeners to isolate, as sound quality assessments using mechanical (Prünster, Fellner, Graf, and Mathelitsch, 2004) and voice (Kreiman *et al.* 1994; Rabinov and Kreiman, 1995) signals show. Establishing the roughness of musical signals is exceptionally challenging because the quality is subsumed under the broader perception of sensory dissonance — a multidimensional sensation (Terhardt, 1974b, 1984). Along with roughness, three other dimensions have been linked to the sensory dissonance of musical intervals: loudness, sharpness (a piercing quality — loudness weighted mean frequency on a physiological frequency scale), and toneness (a quality of periodicity — the opposite of noise — sometimes referred to as "tonality", risking confusion with musicologists' use of the term for a particular musical system; Terhardt, 1984). Of these, however, roughness is presumably the primary dissonance factor through its effectiveness at increasing a musical sound's perceptual tension (Pressnitzer *et al.* 2000) and its frequent association with musical unpleasantness (Blood, Zatorre, Bermudez, and Evans, 1999; Brattico *et al.* 2009; Helmholtz, 1885/1954; Terhardt and Stoll, 1981).

In music and speech signals comprised of harmonics, roughness is introduced in the auditory periphery by the physical interaction of two or more fundamental frequencies, lower order harmonics, and/or subharmonics that fall within a single critical bandwidth or auditory filter (Bergan and Titze, 2001; Greenwood, 1991; Terhardt, 1974a; Zwicker and Fastl, 1990). Maximum roughness occurs when two spectral components are separated in frequency by approximately 50% to 10% of the critical bandwidth, depending on the mean frequency of the components (the percentage decreases as the mean frequency increases; Greenwood, 1991, 1999). This nonlinear property of the human auditory system has intrigued mathematicians and music theorists for centuries. Long before empirical evidence existed to support it, theorists observed that a given musical interval could be more or less rough (i.e., more or less dissonant) depending on the frequency of its lowest tone (Rameau, 1722/1971, pp. 119-123).

The presence of harmonically related frequencies that can lead to perceived roughness in musical intervals may be calculated mathematically (Wild, 2002). In most cases, when the two fundamental frequencies of a complex-tone dyad form a small-integer ratio (e.g.,, 1:2 or octave, 2:3 or perfect 5th), the resultant sound has few or no harmonic partials co-occurring within a single critical band. Such an interval is likely to be judged as consonant (Ayres *et al.* 1980; Butler and Daston, 1968;

Guernsey, 1928; Kameoka and Kuriyagawa, 1969a, 1969b; Malmberg, 1918; Plomp and Levelt, 1965; Plomp and Steeneken, 1967; Schellenberg and Trainor, 1996; Tufts, Molis, and Leek, 2005; Van de Geer *et al.* 1962). (As noted above, exceptions can include small-integer ratio dyads with very low root notes, e.g., below C3, approximately 131 Hz.) A large-integer ratio dyad (e.g., 8:15 or Major 2nd, 9:16 or minor 7th) has narrowly-spaced partials that fall within a single critical band and thus has spectral components that generate a sensation of roughness and a concomitant judgment of dissonance.

If sensory dissonance could be plotted simply as a function of the degree of frequency interaction, listeners' ratings and objective acoustical analyses would agree. For very narrowly spaced pure-tone dyads (two combined sine waves), the sensory dissonance plot derived from listener assessments is considered a reliable indicator of critical bandwidth (Greenwood, 1991). Beyond a critical bandwidth, listeners' sensory dissonance ratings for pure-tone dyads can reflect prevailing cultural biases towards regarding large-integer ratio dyads as dissonant, even in the absence of physical components liable for the sensation (Terhardt, 1984; Tramo, Cariani, Delgutte, and Braida, 2003; see also Chapter 3, Table I). The conclusion that the dissonance phenomenon was more than just the absence of roughness prompted research into the neurophysiology of harmonically related pitches to learn how listeners extract dissonance from musical signals, and how musical expertise mediates this process.

B. Musical expertise and processing differences

The bottom-up, perceptual attributes of musical signals are associated with meaning and emotion (Balkwill and Thompson, 1999; Bigand and Tillman, 2005; Pressnitzer *et al.* 2000), and even nonhuman animals demonstrate altered brain chemistry from exposure to musical signals (Panskepp and Bernatsky, 2002). Studies exploring the neurovascular and neuroelectrical bases of the music/meaning association are relatively recent. In efforts to disable the top-down influence of musical knowledge and expectancies, researchers have presented isolated chords to listeners, presuming that music-theoretic or *cognitive C/D* could at least be somewhat segregated from *sensory C/D* outside of a musical, tonal context (Foss *et al.* 2007; Itoh, Suwazono, and Nakada, 2003; Minati *et al.* 2009; Passynkova, *et al.* 2007; Passynkova, Sander, and Scheich, 2005). These studies have provided some contradictory data on the neural correlates of the C/D distinction but intriguing processing differences between musicians and nonmusicians have been more consistent. Several studies are worth summarizing here to illustrate the dissociation in neural activation patterns between consonance and dissonance and between musicians and nonmusicians. Although the evidence from neuroimaging studies investigating sensory C/D is not entirely convergent, activation networks in three regions are frequently implicated in consonant-versus-dissonant processing.

Functional magnetic resonance imaging (fMRI) has revealed greater neural activation for dissonant over consonant chords in the left inferior frontal gyri (IFG) of musicians (Foss *et al.* 2007; Tillman, Janata, and Bharucha, 2003). Similar differential C/D activation was observed in the right IFG in nonmusicians (Foss *et al.*

2007). Activation in the IFG was shown to be sensitive to manipulations of both the music-theoretic and sensory properties of chords (Tillman *et al.* 2003).

The opposite pattern has also been reported in an fMRI study — greater activation for consonant (defined as 'pleasant') over dissonant (defined as 'unpleasant') chords — but the distinction was found in the left IFG of nonmusicians (Koelsch, Fritz, v. Cramon, Müller, and Friederici, 2006). A third, more recent study supported the Foss *et al.* (2007) and Tillman *et al.* (2003) findings of left-dominant processing in musicians and right-dominant processing in nonmusicians, but the valence of the difference agreed with Koelsch *et al*. — greater activity was seen for consonant chords (Minati *et al.* 2009). The diversity of musical materials used probably accounts for much of the variability.

The left medial frontal gyrus (MFG) of musicians and nonmusicians demonstrated stronger activity for dissonant over consonant chords in isolation (Foss *et al.* 2007). Differential activation between melodies in major and minor keys compared to a sensory dissonant, nontonal sequence was found in the left MFG of nonmusicians (Green *et al.* 2008). Beyond the frontal areas, dissonant chords generated greater activation than consonant chords in the left superior temporal gyri (STG) of musicians (Foss *et al.* 2007; Tillman *et al.* 2003), but this region did not show differential consonant-versus-dissonant processing in nonmusicians.

The sensory C/D distinction must be mapped within a very narrow time window in order to avoid the influence of top-down processing. Schön *et al.* (2005) tracked the time course of chord processing in order to determine when the consonant-versus-dissonant distinction emerged. Piano chords in isolation were presented to musicians and nonmusicians and brain activity recorded via long-latency event-related brain potentials (ERP). Differential neuroelectrical activation to consonant-versus-dissonant chords was observed in musicians and nonmusicians, however musicians processed C/D differences faster and with greater acuity than nonmusicians, as indexed by the early N1-P2 complex elicited 100-200 ms post stimulus. A rating task was included to allow comparison of neural activity to listeners' subjective assessments of the pleasantness of the stimuli. Musicians' differential C/D activity showed stronger correlations with *pleasantness* ratings than with the chords' music-theoretic C/D, supporting earlier findings that musicians engage in a subjective assessment of chords more rapidly than nonmusicians do (Zatorre and Halpern, 1979). Nonmusicians' ERPs also reflected sensory consonant-versus-dissonant processing differences, as shown in the N2 activity elicited 200 ms post stimulus. These differences, however, did not appear in the accompanying rating task. The researchers proposed that perhaps, "nonmusicians perceive the differences in the acoustical properties of sounds but do not rely on the perceptual analysis (for rating) because they lack confidence in their perceptual judgment."

In the same study, the N420 — a late negative component considered indicative of cognitive categorization — showed strong differential C/D activity in musicians but was weaker for nonmusicians. The amplitude of this late component was largest for imperfect consonances — the musical intervals midway between consonance and dissonance and the most difficult to categorize. This observation supported the authors' conclusion that musical expertise hones the neural responses to chord types presented in isolation and mediates their categorization (Schön *et al.* 2005).

Early and late differential processing of dyads can be elicited simply by differences in the frequency ratio between two simultaneous pure tones, also shown in an ERP study (Itoh *et al.* 2003). These researchers concluded that, "cortical responses to pure-tone dyads were affected not only by sensory roughness but also by other features of the stimuli concerning pitch-pitch relationships." The finding supported the hypothesis that the evaluation of pure-tone dyads is under the influence of knowledge-based processing. Unfortunately, the study involved only musicians. The inclusion of nonmusicians would have provided a valuable comparison because pure-tone dyads are infrequently used in neurophysiological C/D studies.

These results call attention to the usefulness of ERP in chord processing studies for separating sensory-driven from knowledge-driven (and possibly biased) processing. Test conditions can create expectancies, allowing participants to use probability to anticipate aspects of an upcoming stimulus, biasing their responses (Ward, 1987). The rapid-response N1-P2 complex is elicited under passive testing conditions, reflecting preattentive processing that is immune to (observer) probability. By contrast, long latency components occurring 250-600 ms post stimulus are elicited only while participants actively attend to the stimuli (Parasuraman and Beatty, 1980; Parasuraman, Richer, and Beatty, 1982). Differential rough-versus-nonrough chord processing has been measured in the early P2 component under passive listening conditions while participants read and ignored the stimuli (Alain, Arnott, and Picton, 2001). This study did not use musical intervals or screen for musical expertise, but did show that auditory roughness perception is preattentively encoded, a conclusion that has been supported elsewhere (De Baene, Vandierendonck, Leman, Widmann, and Tervaniemi, 2004).

Is the musician's heightened neural activation to the C/D distinction caused by enhanced perceptual sensitivity or by greater familiarity with intervals? Would Western musicians have a higher capacity for C/D discrimination than nonmusicians for chords outside of the Western tonal tradition? Brattico *et al.* (2009) used magnetoencephalography (MEG) to address this question. The change-related mismatch negativity (MMNm) response was measured using four categories of chords: major, minor, microtuned (mistuned from the traditional Western standard), and sensory dissonant. Processing differences were measured bilaterally in the primary and secondary auditory cortices of the temporal lobes at approximately 200 ms post stimulus. Musicians showed faster and stronger responses than nonmusicians to differences between all chord types and were the only group to elicit a difference between major and minor chords. For both groups the automatic response to sensory dissonance was greater and faster than for microtuned musical chords. The earliest P1m component did not differ between groups.

Taken together, these results indicate that initial, bottom-up, sensory-based chord processing is similar for musicians and nonmusicians. Musical expertise, however, rapidly enables top-down processing to assist in the categorization and assessment of both familiar and unfamiliar chords.

C. Subjective rankings: Sources of error and variability

The current experiment uses a psychophysical scaling task to update the data from sensory C/D ratings while attending to sources of rating variability. In contrast to neurophysical measures, behavioral measures from scaling judgments are prone to greater inter- and intra-subject variability. Four broad sources of error have been implicated in this type of task: long-term memory (LTM), short-term memory (STM), residual stimuli, and background stimuli (Ward, 1987).

Participants using a scale to make category judgments are expected to ignore any internal, absolute stimulus-response mappings in favor of new, relative maps based solely upon the experimental content, but this does not always happen. Long-term memory for stimulus-response mappings made hours or days earlier affect the responses made to the current stimuli under assessment. Participants asked to rate a stimulus set a day after providing an initial set of ratings showed bias in the direction of the previous day's mapping (Ward, 1987). The effect of rating the first stimulus set "taught" participants what to expect from the second set. Participants' STMs for stimuli also influence judgments by creating expectancies for the to-be-presented stimulus. The sequential dependency of a response to previous responses is independent of the changes in judgment induced by LTM and learning (Ward, 1987). Long- and short-term memory (STM) processes thus influence both the absolute and relative judgments that co-occur in category formation tasks (Lockhead, 2004). Presenting stimuli in randomized order reduces the cumulative effect of sequential dependency. Allowing participants to replay each stimulus as needed before making a decision reduces the STM trace for previous stimuli by increasing the inter-trial interval.

The internal representation of a stimulus is also biased by general experience with the specific sensory continuum being scaled. The psychological boundaries or internal scale endpoints may not be the same even within participant groups (Lockhead, 2004; Ward, 1987). In the case of roughness evaluation, string players such as violinists or cellists are likely to have experienced a greater variety of musical roughness as induced by their instruments than players of fretted instruments. Helson (1964) labeled these life experiences *residual stimuli* — referring to what a person knows of the stimulus type. These internal standards may or may not be used to judge the magnitude of a stimulus element and the experimenter will have difficulty knowing exactly how to account for this (Lockhead, 2004). Gathering information from participants on their musical culture, training, and listening habits adds valuable insight to the data and reduces error by pooling latent variables. Lastly, the enduring characteristics of an individual's sensory system or response characteristic influences scaling judgments and is termed *background stimuli* or internal noise, but this component plays a minor role (Ward, 1987).

How concerned should the experimenter be with the four sources of error listed above? Studies of voice quality assessments have concluded that most of the variability in rating data is actually due to task design and not listener unreliability, and can therefore be controlled (Kreiman, Gerratt, and Ito, 2007). Kreiman *et al.* (2007) identified four factors making the largest contribution to inter-rater variability: stability of listeners' internal standards, scale resolution, difficulty isolating the attribute, and attribute magnitude. The stability of the internal standard can be improved by providing listeners with an external comparison stimulus (Gerratt, Kreiman, Antoñanzas-Barroso, and Berke, 1993). (However, if the method of paired comparisons is used the experimenter must carefully design the paradigm to avoid inducing response biases; Link, 2005.) Surprisingly, inter-rater correlations are shown to substantially improve when a continuous, high-resolution scale is substituted for a discrete, low-resolution scale (Kreiman *et al.* 2007).

Intra-rater consistency depends on the listener's ability to isolate the property under test. Speech pathologists found it difficult to isolate and assess vocal roughness without including the contribution from a second quality — *breathiness* (Kreiman *et al.* 1994). Expert listeners could focus their attention on breathiness, but differed considerably in their capacity to focus attention on vocal roughness per se. Providing examples of roughness improved listener agreement (Kreiman *et al.* 2007). In addition, listeners' past experiences concentrating on any type of specific auditory signal helped them to isolate auditory attributes. When voice assessment novices rated the roughness of vowel sounds, ratings by those with musical training were more consistent than those who had little or no training (Bergan and Titze, 2001).

Lastly, the magnitude of the attribute is shown to affect listener agreement (Lockhead, 2004). Voice assessment ratings show greater agreement near the endpoints of the scale where items are more alike and more variability near the midpoint (Kreiman *et al.* 2007). Providing stimuli with properties having a broad range of noticeable differences allows the experimenter to account for this tendency. The present work aims to improve experimental control and thus supplement the behavioral data on dyad sensory dissonance by attending to these sources of inter- and intra-rater variability,

Raters of voice roughness were shown to be as reliable as objective roughness measures from auditory analyzers (Rabinov and Kreiman, 1995). The current work takes a similar approach by comparing listener roughness ratings with those provided by two software-based auditory roughness analyzers. In addition, our listeners' pure-tone dyad ratings were compared to sensory dissonance ratings interpolated from Plomp and Levelt's (1965, Fig. 9) plot of the pure-tone dyad consonance band. Likewise, our complex-tone, just-tuned dyad ratings were compared against ratings predicted by a theoretical model of the acoustic dissonance of complex-tone, Western dyads by Hutchinson and Knopoff (1978), to learn the extent to which musical training was assumed in their model.

II. EXPERIMENT 1: PURE-TONE, JUST-TUNED DYADS

A. Method

1. Participants

Participants (*N* = 30; 14 men and 16 women; 18 - 51 years; *M* = 23, *SD* = 5.9) were recruited from a classified ad and the Schulich School of Music at McGill University. Three (two musicians, one nonmusician) were volunteers who served without pay; 27 recruits were paid $5 for their time. Fifteen participants (musician group) had seven or more years of formal music training (*M* = 13, *SD* = 5.0); the remaining 15 participants (nonmusician group) had 2.5 or fewer years of training (*M* = 1, *SD* = 0.9). None had absolute pitch perception by self-report and all reported normal hearing. Musical training and music listening habits were assessed using a modified version of the Queen's University Music Questionnaire (Cuddy, Balkwill, Peretz, and Holden, 2005).

2. Apparatus and Stimuli

Participants were seated in a soundproof booth (IAC, Manchester, U.K.) at a Macintosh G5 PowerPC computer (Apple Computer, Cupertino, CA). Dyads were delivered from the Macintosh's digital output to a Grace m904 (Grace Design, Boulder, CO) digital interface, converted to analog and presented to listeners diotically through Sennheiser HD280 Pro 64 Ω headphones (Sennheiser, Wennebostel, Germany).

The software package Signal (Engineering Design, Berkeley, CA) was used to create 72 pure-tone (PT) dyads by summing in cosine phase a lower frequency sine wave (*f*1) with a higher frequency sine wave (*f*2). Dyads were amplitude normalized so that each stimulus had a sound pressure level of 57 $\pm$ 0.75 dBA SPL at the headphone as measured with a Brüel & Kjær 2203 sound level meter and Type 4153 Artificial Ear headphone coupler (Brüel & Kjær, Naerum, Denmark). (A level below 60 dB SPL was selected as optimal for ensuring sensitivity to stimulus differences while avoiding acoustic distortion products aural harmonics and combination tones; Clack and Bess, 1969; Gaskill and Brown, 1990; Plomp, 1965.) Each dyad was 500 ms in duration, including 10-ms raised cosine onset and offset ramps.

The 72 dyads formed the 12 musical intervals of the Western chromatic scale: minor 2nd (m2), Major 2nd (M2), minor 3rd (m3), Major 3rd (M3), perfect 4th (p4), tritone (tri), perfect 5th (p5), minor 6th (m6), Major 6th (M6), minor 7th (m7), Major 7th (M7), and octave (oct). Each interval's lower frequency (*f*1) was assigned by a random number table (without replacement) and spanned two octaves from C3 (130.8 Hz) to B4 (493.8 Hz). The upper frequencies (*f*2) ranged from D#3 (155.6 Hz) to A5 (880 Hz). Interval spacing corresponded to the just-tuned scale to conform to the protocol used in earlier work linking psychoacoustics and musical intervals.

Six unique dyads — three per octave span — were created at each pitch chroma. For example pitch chroma C was represented by p5, m6, and p4 at the root note of C3 (131 Hz) and M2, p4, and m7 at the root note of C4 (262 Hz). The reciprocity of the design allowed for six unique dyads for every musical interval. For example, the six m2s had root notes of C#3, D3, E3, A4, G4, and B4. (See Table A.I for a complete description.)

3. Procedure

Participants were tested individually and were instructed to rate each dyad for how "rough" or "smooth" it sounded, using their own judgment. Roughness was defined as analogous to an object's texture, as if it were a sensation upon the skin. A short, eight-trial practice session familiarized participants with the task and the equipment. The practice dyads were a subset of the dyads presented in the main experiment, chosen to represent probable extremes of rough and smooth. The rough exemplars were two m2s, a M2, and a low m3; the smooth exemplars were an octave, M7, M6, and high p5. Participants were told that these eight dyads represented rough and smooth exemplars but were not told which was which. It was emphasized that roughness should not be considered the same as musically inharmonious, nor should it be confused with pitch height or how "sharp" a tone sounded. Participants were asked to ignore all other dimensions of the sound, especially musical quality or personal preference, in order to focus on roughness as a distinct sensory dimension. Practice results were not analyzed.

Participants rated dyads by moving a continuously adjustable, linear, high-resolution on-screen slider implemented via a Max/MSP program (Cycling '74, San Francisco, CA) to a position that reflected the degree of roughness. The slider was a visual analog type — the leftmost end was labeled "smooth" and the rightmost end was labeled "rough." They were asked to use the full width of the slider, reserving the middle area only for dyads that seemed to be neither especially rough nor smooth. An on-screen "Hear again" button could be used by a participant to audition each dyad as many times as needed to make a decision. Slider positions were recorded with a resolution of 4 decimal places, corresponding to 10,000 discrete positions between the values of 0 for smoothest and 1 for roughest. Once a rating was entered, the next trial began. Earlier dyads could not be re-auditioned.

The 72 dyads in the experiment were presented in randomized order for each participant. Upon completion of the task, participants filled out a short questionnaire regarding their musical experience.

B. Results

To determine if and how roughness ratings differed across interval classes (dyads sharing the same frequency ratio relationship), box plots were created to show the median values for each interval and the upper and lower quartiles. Figure 1 shows that roughness ratings were high for typically narrow intervals (m2 and M2) and low for wider (M7 and octave), as expected from earlier findings for PT dyads (Plomp and Levelt, 1965). Differences between the two groups were apparent in musicians' tendencies to assign slightly lower roughness ratings to small-integer ratio intervals (M3, p4, p5, octave), compared to nonmusicians' ratings. Musicians, but not nonmusicians, rated the large-integer ratio tritones as rougher than neighboring small-integer ratio p4s and p5s. Twelve independent-sample t-tests were performed to evaluate the hypotheses that participant groups would rate interval classes differently. In each case the independent variable was group membership and the dependent variable was participants' average roughness ratings for each class's six dyads. The mean differences between musicians' and nonmusicians' PT roughness ratings were not significant under a Bonferroni corrected alpha for multiple tests of 0.004 (0.05/12 = 0.004). The p4 interval class came the closest to significance [t (28) = –2.13, p = 0.04], with musicians assessing these dyads as smoother than did the nonmusicians.

Different root pitches were represented within interval classes, meaning that dyads at a given interval were relatively narrow or wide with respect to critical bandwidth. Figure 2 depicts the PT dyads' mean roughness ratings as a function of the frequency difference between tones as indexed by the equivalent rectangular bandwidth (ERB). (The ERB-rate scale allows frequency to be converted into a unit of measure reflecting the critical bandwidth of a frequency's functional auditory filter. The number increases as the place of a frequency's maximum activation on the basilar membrane gets more distant from the oval window; Moore, 2003). Values on the ERB-rate scale were calculated for the frequencies of each dyad using the formula of Glasberg and Moore (1990) and the difference was taken (ERBn $f2$ – ERBn $f1$). As expected, dyads with tones separated by less than a critical band (ERBn difference < 1.0) were judged as rougher than dyads exceeding a critical bandwidth. Both groups judged m7s and M7s to be rougher on average than the octaves, despite the similar bandwidths for these wide dyads. This trend was more pronounced in the musician group. Musicians also showed a greater tendency to assess dyads within some classes as similarly rough, e.g., p4s, p5s, and tritones. In contrast, nonmusicians' averaged ratings for dyads in these classes were more sensitive to bandwidth differences.

Insert Figure 1 About Here

Insert Figure 2 About Here

"For perceptual ratings to be meaningful, listeners must use scales consistently" (Kreiman, Gerratt, Kempster, Erman, and Berke, 1993). Coefficient alpha (Cronbach, 1951) was chosen as an estimator of *intra-rater reliability* to examine rater consistency and to learn the extent to which frequency ratio relationship (a within-class constant) was a main factor in roughness ratings. High values of alpha indicate that the rater regarded all items in a set as having high commonality and low uniqueness (Cortina, 1993; Green, Lissitz, and Mulaik, 1977; Schmitt, 1996). Negative alphas reflect inconsistent codings and/or multidimensionality (i.e., listeners were confused over what constituted roughness, or had difficulty isolating the property). The analyses included intercorrelation matrices that measured the extent to which roughness ratings in an interval class reflected a single, unifying dimension, providing confidence intervals for the coefficient alpha (Cortina, 1993; Iacobucci and Duhachek, 2003; Schmitt, 1996).

Presuming the common factor among items in an interval class to be frequency ratio relationship, classes with some narrow and some wide dyads with respect to critical bandwidth (e.g., m3, M3, p4) were expected to show greater variance in ratings and correspondingly lower values of coefficient alpha than classes with exclusively narrow or wide (e.g., octave, m2) intervals. This hypothesis was supported more strongly by nonmusicians (see Table I). Nonmusicians' coefficient alpha values ranged from 0.32 to 0.89 ($M_{\mathrm{NMus}} = 0.66$, $SD_{\mathrm{NMus}} = 0.16$), with significantly low variance (high commonality among items) shown for the octave ($p \leq 0.05$). Nonmusicians showed the highest variance for dyads in p5, m3, M3, and m6 classes. Musicians showed a higher tendency than nonmusicians to regard each interval class as having a single unifying factor shared among the dyads, with coefficient alphas ranging from 0.55 to 0.89 ($M_{\mathrm{Mus}} = 0.76$, $SD_{\mathrm{Mus}} = 0.09$). Significantly low variance ($p \leq 0.05$) was seen for the octave, m7, and m2. An omnibus paired samples *t*-test (2-tailed) compared the two groups' mean coefficient alphas at each interval class. The test indicated a tendency for musicians to show greater intra-rater reliability than nonmusicians when judging PT dyad classes, although this fell just short of significance at the 5% level [$t(11) = 2.04$, $p = 0.07$].

Insert Table I About Here

"For ratings to be clinically useful, inter-rater agreement must be high" (Kreiman *et al.* 1993). The intraclass correlation coefficient (ICC) was used to test *inter-rater reliability* (Shrout and Fleiss, 1979). Because it measures coherence in a group of raters, the ICC is highest when all judges assign similar scores to an item. Negative ICC values indicate that differences among raters exceeded the naturally occurring variance of the items (McGraw and Wong, 1996). A two-way (subjects: 15 levels; dyads: 6 levels), random effects model was used considering that both dyads and raters represented a random sample of all possible dyads and respondents. The model looked for correlated ratings among judges rather than exact matches, therefore a *consistency* model was chosen over the *absolute* value type. Because we had no a priori assumptions regarding the amount of variability we would find within each group, we used listeners' individual ratings as the unit of analysis, and therefore a *single-measures* model was chosen over an *average-measures* model. (The other analysis option would have compared the averaged ratings from the musician group for each dyad with those from the nonmusicians, but it would have ignored the individual variance that we were interested in quantifying. Note that when the unit of analysis is individual ratings, the ICC is generally lower than found with an average-measures model; Shrout and Fleiss, 1979.) In the convention of McGraw and Wong (1996), this model is Type ICC(C,1). The ICC results are reported in Table II.

Nonmusicians showed somewhat higher inter-rater reliability for judging PT dyads than did musicians when ratings were analyzed by interval class, as indicated by a 2-tailed *t*-test [$t(11) = –2.06$, $p = 0.06$] and shown in Table II. Nonmusicians' ICCs ranged from –0.05 to 0.29 ($M_{\text{NMus}} = 0.14$, $SD_{\text{NMus}} = 0.11$). Their ratings in several interval classes (M3, p4, tri, p5, m6, M6, and m7) were in significantly high agreement ($p \leq 0.05$), as determined by *F* tests on the population value (ρ) of the 95% confidence intervals (Shrout and Fleiss, 1979). Musicians reported a broader range of ICC values than did nonmusicians, ranging from –0.06 to 0.37 ($M_{\text{Mus}} = 0.07$, $SD_{\text{Mus}} = 0.11$). Musicians' ratings showed significantly strong agreement in the interval classes of m3, p4, and m6. Even though the mean roughness ratings for all six PT dyads in some classes may have been similar, listeners did not agree as to which was the roughest or smoothest dyad of the six. This was indicated by negative ICCs for m2 (both groups) and M2 and M3 (musicians only). The lack of agreement for the m2 is presumably linked to the dual nature of roughness in this class. Some m2s were narrow enough to cause beating — a sensation distinct from roughness that might have been considered smooth by some raters.

Insert Table II About Here

III. EXPERIMENT 2: COMPLEX-TONE, JUST-TUNED DYADS

The second experiment examined the effect of harmonic partials on roughness ratings and followed the same protocol using just-tuned, complex-tone (CT) dyads.

A. Method

1. Participants

Participants who served in Experiment 1 also rated dyads in Experiment 2 and were paid an additional $5. The order of participation was random; roughly half of the participants performed Experiment 2 before Experiment 1.

2. Apparatus and Stimuli

Testing was conducted under the same conditions described in Experiment 1.

The McGill University Master Samples (MUMS) provided a sample of an alto saxophone playing the note D#4 (MUMS ID #16, volume 3, sample 16-03). The sample was digitally transferred to the audio processor ProTools (Digidesign, Daly City, CA) and digitally pitch-shifted higher or lower as needed using the ProTools plug-in “Pitch & Time - Algorithm B” to create 33 upper notes and 24 lower notes. The technique maintained the relative amplitudes of the harmonics upon transposition. Seventy-two CT dyads were created by combining in phase a lower frequency note ($f1$) with an upper frequency note ($f2$), each matched in amplitude. The lower frequencies ($f1$) of each musical interval were assigned according to a random number table ranging from C3 (130.8 Hz) to B4 (493.8 Hz). The upper frequencies ($f2$) ranged from D#3 (155.6 Hz) to A#5 (932.3 Hz). As in Experiment 1, dyads were presented to listeners at 57 ± 0.75 dBA SPL at the same headphone as measured with the same equipment. Each dyad was 500 ms in duration, including 10-ms raised cosine onset and offset ramps.

As before the 72 CT dyads formed the 12 musical intervals of the Western chromatic scale. Interval spacing corresponded to the just-tuned scale. The scheme of six unique intervals at each pitch chroma was used, but the interval assignments at each root note were not the same as in Experiment 1. (See Table A.II for a complete description of each dyad.)

3. Procedure

The same procedure and instructions used in Experiment 1 were followed here. The eight dyads used in the practice session were: two octaves, a M7, a m6, two M2s, a m2, and a low m3.

B. Results

Analysis of CT dyads proceeded as in Experiment 1. Box plots in Fig. 3 show a greater disparity between musicians’ and nonmusicians’ ratings of CT dyads, compared to the PT dyad results. More so than seen with PT dyads, musicians' averaged CT ratings reflected the full range of frequency ratio relationships. In contrast, nonmusicians’ averaged ratings are shown clustered near the middle of the scale.

Independent-sample *t*-tests showed some significant differences between musicians’ and nonmusicians’ ratings in the 12 classes of intervals. Musicians’ ratings of octaves, p5s, and p4s were significantly lower (i.e., smoother) than nonmusicians' ratings of these classes under the Bonferroni corrected alpha of $p \leq 0.004$ [octave: $t(28) = -3.25$, $p = 0.003$; p5: $t(28) = -3.09$, $p = 0.004$; p4: $t(28) = -3.59$, $p = 0.001$]. Other differences approaching significance between musicians and nonmusicians included the M3s [$t(28) = -2.64$, $p = 0.01$], rated smoother by musicians, and the m7s [$t(28) = 2.16$, $p = 0.04$] and M7s [$t(28) = 2.21$, $p = 0.03$], rated rougher by musicians.

Insert Figure 3 About Here

Insert Figure 4 About Here

Figure 4 shows the average roughness ratings of CT dyads as a function of the ERB-rate scale difference between the two tones of each dyad. As expected, musicians' roughness ratings reflected the roughness contribution from narrowly spaced harmonics by showing sharp differences between their ratings for large- and small-integer ratio CT dyads. In contrast, nonmusicians' assessments of dyad roughness appeared to be more a function of ERB rate scale differences than to the presence of harmonics. Octaves — dyads with only coincident partials — were rated as only marginally smoother than tritones and m7s — dyads with considerable roughness components. Nonmusicians rated tritones as slightly smoother than the less complex, but fundamentally wider, m7 and M7 dyads. It is not clear why musicians regarded m7s and M7s as rougher than tritones unless the separation between the fundamental frequencies was as important a criterion as the roughness induced by harmonic partials.

Compared with the PT dyad result, coefficient alphas measuring intra-rater reliability were expected to be higher for CT dyads. It was hypothesized that the shared harmonic relationship among dyads in each interval class would stabilize roughness ratings, but this was not supported. As shown in Table III, intra-rater reliability for CT dyad ratings was similar to that seen for PT dyads. *T* tests compared the 12 mean coefficient alphas for both stimulus types, and found no significant difference [$t(11)_{\text{Mus}} = 1.35$, $p = 0.21$; $t(11)_{\text{NMus}} = 0.26$, $p = 0.80$]. CT dyad reliability was statistically equivalent for both groups [$t(11) = 1.25$, $p = 0.24$]. Musicians' coefficient alphas ranged from 0.53 to 0.86 ($M_{\text{Mus}} = 0.72$, $SD_{\text{Mus}} = 0.11$). Nonmusicians' alphas ranged from 0.41 to 0.85 ($M_{\text{NMus}} = 0.67$, $SD_{\text{NMus}} = 0.14$). As seen in the PT dyad results, ratings at the extremes of the scale showed less variance (i.e., higher alphas) than ratings for those in the middle. Once again, the m2 showed significantly high ($p \leq 0.05$) intra-rater reliability, for both participant groups in this case. Nonmusicians also showed high reliability for octaves, as shown in Table III.

Insert Table III About Here

Perhaps due to the greater ecological validity of CT compared with PT dyads, individual differences appeared to have a stronger affect on assessments of these (presumably) more familiar-sounding stimuli, but only for nonmusicians. Nonmusicians showed significantly greater group disagreement for assessing CT than for PT dyads, when the mean ICCs were compared [$t(11)_{\text{NMus}} = 4.05$, $p < 0.01$]. Musicians' group coherence did not differ between PT and CT dyads [$t(11)_{\text{Mus}} = 1.60$, $p = 0.14$]. Inter-rater reliability for CT dyads was not significantly different between the two groups [$t(11) = 1.06$, $p = 0.31$]. Compared with the PT dyads, a higher number of negative ICCs were reported for CT dyads (see Table IV). Musicians' ICCs ranged from –0.05 to 0.12 ($M_{\text{Mus}} = 0.02$, $SD_{\text{Mus}} = 0.06$). Nonmusicians' ICCs ranged from –0.04 to 0.04 ($M_{\text{NMus}} = 0.00$, $SD_{\text{NMus}} = 0.03$). Musicians were in significantly high agreement ($p \leq 0.05$) in rating M2s and m3s, but nonmusicians did not reach agreement for any interval class, as indicated by *F* tests. These results suggest that roughness components contributed by harmonic partials were an unreliable predictor of how participants, especially nonmusicians, would behave as a group in assigning roughness ratings.

Insert Table IV About Here

IV. EXPERIMENT 3: COMPLEX-TONE, MICROTUNED DYADS

Experiment 3 aimed to account for individual differences by limiting the influence of exposure to Western tonal music in dyad roughness assessment. We asked musicians and nonmusicians to rate the roughness of complex tone dyads in microtuned (MT) frequency ratio relationships that are not often encountered in Western music.

A. Method

1. Participants

Participants (N = 30; 12 men and 18 women; 18 - 55 years; M = 28, SD = 9.4) were recruited from a classified ad and the Schulich School of Music at McGill University. One of the participants (a nonmusician) served in Experiments 1 and 2, but not on the same testing days. Six participants were volunteers and served without pay; 24 recruits were paid $5 for their time. Fifteen participants (musician group) had five or more years of formal music training (M = 16, SD = 8.1); the remaining 15 participants (nonmusician group) had two or fewer years of training (M = 0.8, SD = 0.8). None had absolute pitch perception by self-report and all reported normal hearing. Musical experience was assessed with the modified Queen's University Music Questionnaire (Cuddy *et al.* 2005) used in Experiments 1 and 2. Persons accustomed to listening to music that included microtuned intervals, such as Indian and Arabic music, were excluded from the study.

2. Apparatus and Stimuli

Testing was conducted under the same conditions described in Experiment 1.

The same sample of an alto saxophone playing the note D#4 used in Experiment 2 was used to create 72 MT dyads. Analog-to-digital conversion, pitch-shifting, and the combining of notes were accomplished using the same equipment and procedure. The lower notes were the same as in Experiment 2. Intervals were assigned to the lower notes by random assignment. The upper frequencies ($f2$) ranged from D3 plus a quartertone (151.1 Hz) to A5 plus a quartertone (905.8 Hz). As in Experiments 1 and 2, dyads were 500 ms in duration, including 10-ms raised cosine onset and offset ramps. Dyads were presented to listeners at 57 $\pm$ 0.75 dBA SPL at the same headphone as measured with the same equipment.

The 72 MT dyads were composed of altered versions of 12 musical intervals of the Western chromatic scale: unison (uni+), minor 2nd (m2+), Major 2nd (M2+), minor 3rd (m3+), Major 3rd (M3+), perfect 4th (p4+), tritone (tt+), perfect 5th (p5+), minor 6th (m6+), Major 6th (M6+), minor 7th (m7+), and Major 7th (M7+), where the upper notes of the intervals were augmented by a single quartertone (about a 3% increase in frequency).

To accurately describe the stimulus set, the frequency-ratio relationships were calculated by multiplying $2^{n/24}$ (where n represents an odd integer from 1 to 23) by whole numbers until the closest integer-ratio relationship was derived. For example, the lower frequency of the quartertone dyad uni+ is related to its upper frequency by a number that corresponds to 2 raised to 1/24, or 1.029 (a semitone dyad has a ratio of 1.059, the 12th root of 2). A distance of 21 quartertones or 1.834, for example, equals the ratio of the m7+ interval. By multiplying 1.834 by successive whole numbers, a whole number product is eventually derived describing the integer-ratio relationship of the m7+ in whole numbers. (For example, $1.834 \times 6 = 11.004$, so a ratio of 11:6 was assigned to the m7+. Each interval's ratio was determined once the closest whole integer, after rounding off at the hundredths place, was found.) If evaluated by frequency ratio complexity, all of the MT dyads were notably dissonant; however, some had frequency ratio relationships that were less complex than some of the just-tuned CT dyads used in Exp. 2. (See Table A.III for a complete description of each dyad.)

3. Procedure

The same procedure and instructions used in Experiments 1 and 2 were followed here. No a priori assumptions were made for the roughness of the eight dyads used in the practice session, and so the following were chosen as a representative sample: two M2+s, two m2+s, a M6+, a m6+, a m3+ and a M7+.

B. Results

Analysis of MT dyads proceeded as in Experiment 1. The box plots in Fig. 5 show that mean roughness ratings for most classes of MT dyads centered about the middle of the scale. There were no significant differences between mean musicians' and nonmusicians' MT roughness ratings under a Bonferroni corrected alpha of 0.004. The uni+ interval class came closest to being significantly different ($t(28) = 2.48$, $p = 0.02$), with musicians assessing these dyads as rougher than did the nonmusicians.

Insert Figure 5 About Here

Independent-sample t-tests compared musicians' and nonmusicians' ratings in the 12 classes of MT intervals. Musicians' roughness ratings were not significantly different from nonmusicians' ratings under the Bonferroni corrected alpha of $p \leq 0.004$ for any given interval. One interval, the uni+, approached significance, with musicians rating it as rougher [$t(28) = 2.48$, $p = 0.02$].

Figure 6 shows the mean MT dyad roughness ratings as a function of ERB-rate scale. Only the musician group followed the pattern observed in Experiments 1 and 2 by assigning higher roughness ratings to dyads narrower than a critical band (ERBn difference < 1.0). Nonmusicians' ratings were nearly uniformly flat across the ERBn difference scale. As noted above, nonmusicians tended to rate the uni+ dyads as smoother than the m2+s, in contrast to the musician groups' ratings.

Insert Figure 6 About Here

Compared to when familiar, just-tuned intervals were assessed, the intra-rater reliability decreased significantly for ratings of unfamiliar MT dyads, as seen in Table 5. Individuals rated interval classes of PT dyads [$t(11)_{\mathrm{Mus}} = 2.42$, $p < 0.05$; $t(11)_{\mathrm{NMus}} = 2.44$, $p < 0.05$] and CT dyads [$t(11)_{\mathrm{Mus}} = 2.32$, $p < 0.05$; $t(11)_{\mathrm{NMus}} = 2.51$, $p < 0.05$] more consistently than they did MT dyads. Both groups reported lower coefficient alphas for MT dyads, although there was no significant difference between musicians and nonmusicians [$t(11) = 0.36$, $p = 0.72$]. Musicians' coefficient alphas ranged from –0.59 to 0.87 ($M_{\mathrm{Mus}} = 0.41$, $SD_{\mathrm{Mus}} = 0.50$). Nonmusicians' alphas ranged from –1.12 to 0.76 ($M_{\mathrm{NMus}} = 0.32$, $SD_{\mathrm{NMus}} = 0.53$). Negative alphas were observed in both groups, indicating that individuals were highly inconsistent in rating the six items grouped by interval class. The classes in which these confusions occurred revealed differences between the two participant groups. Musicians showed negative alphas for p4+ and p5+ — intervals representing midpoints between psychologically consonant and dissonant intervals (p4 and tritone in the first example, and p5 and m6 in the second example). Nonmusicians showed negative alphas for M2+ and m3+ — classes containing both narrow and wide dyads in terms of critical bandwidth.

Insert Table V About Here

Inter-rater reliability indicated about the same degree of rater agreement for MT as for CT dyads in terms of which of six items were roughest or smoothest [$t(11)_{\mathrm{Mus}} = 1.41$, $p = 0.19$; $t(11)_{\mathrm{NMus}} = 2.08$, $p = 0.06$]. When MT ICCs were compared with PT ICCs, nonmusicians showed a significantly higher group agreement when rating PT dyads [$t(11)_{\mathrm{NMus}} = 3.55$, $p < 0.01$], but musicians were in no better agreement under the same comparison [$t(11)_{\mathrm{Mus}} = 0.80$, $p = 0.44$]. For MT dyad ratings the mean ICCs between groups showed no significant difference [$t(11) = 0.95$, $p = 0.36$]. Compared with the CT dyad result, fewer negative ICCs were reported for MT dyads, as seen in Table VI. Musicians' ICCs ranged from –0.04 to 0.15 ($M_{\mathrm{Mus}} = 0.04$, $SD_{\mathrm{Mus}} = 0.05$). Nonmusicians' ICCs ranged from –0.02 to 0.07 ($M_{\mathrm{NMus}} = 0.02$, $SD_{\mathrm{NMus}} = 0.03$). Musicians were in significantly high agreement ($p \leq 0.05$) in rating M2+s, m3+s, and tt+s, but nonmusicians did not reach significant agreement for any interval class, as indicated by *F* tests.

Insert Table VI About Here

In summary, the influence of musical interval class and each dyad's associated root note on roughness ratings was to some extent a product of the life experiences of the perceiver. One-way repeated-measures ANOVAs compared the six mean dyad roughness ratings within each of the twelve interval classes. Post hoc tests following significant ANOVAs used a critical value of Tukey's HSD ($\alpha = 0.05$, levels = 6, df = 70) against the observed HSD for pairs of dyads. Significantly different pairs are shown in Table VII.

Insert Table VII About Here

V. COMPARATIVE ANALYSIS

A. Method

To assist the current experiment in describing listener variability, roughness ratings were compared with ratings provided by two software-based auditory roughness analyzers and, in particular cases, against the predicted sensory dissonance ratings from earlier work. We used the Spectral Roughness Analyzer (Vassilakis, 2007, based on data from Sethares, 1998) and PsySound 3's Roughness DW program (Flax, M., and Ferguson, S., v 1.0, alpha, 2007, based on Daniel and Weber, 1997). (Other software-based analyzers were considered and rejected due to incompatibility with our available, more recent, computer operating systems and/or the analyzer's reliance on specific databases.) The appendix offers a more detailed description of each analyzer.

So that evaluated roughness may be discussed in terms of PT dyad sensory dissonance, Exp. 1's ratings were compared against predicted ratings as derived from Plomp and Levelt's (1965, Fig. 9) plot of the sensory consonance band for pure-tone dyads. Plomp and Levelt gathered subjective assessments of PT dyad consonance from 50 male participants. Their Fig. 9 plotted the smallest frequency difference between two tones that resulted in a rating of "consonant" (defined as *beautiful* and *euphonious*), as a function of a dyad's mean frequency. Pure-tone dyads in the current experiment were fitted to their "dissonance band" and a dissonance rating was assigned. Appendix A describes the derivation of these PT dyad dissonance ratings.

Hutchinson and Knopoff (1978) provided a table comparing the dissonance ratings of Helmholtz (1885/1954) and Plomp and Levelt (1965) with the rank orderings of C/D used in musical practice. The CT (just-tuned only) dyad assessments from listeners in the current experiment were compared with Hutchinson and Knopoff's table of "dissonance factors for the intervals within two octaves."

Each analyzer was calibrated according to its designer's recommendations and the roughness values for all 216 (72 × 3) dyads were produced. Within stimulus types (PT, CT, or MT) each participant group's mean ratings were rank ordered from 1 to 72, as were ratings from the analyzers (SRA and DW) and the appropriate model (P&L or H&K). Spearman product-moment correlations (r) and the 95% confidence intervals between pairs of rankings were calculated using a bootstrap method. Samples were drawn 1000 times with replacement from pairs in each stimulus set. The results are illustrated in Fig. 7 and are discussed in order of stimulus type, followed by comparisons within analyzers.

Insert Figure 7 About Here

B. Results

1. Pure-tone, just-tuned dyads

Both analyzers and the P&L model fit the musicians' and nonmusicians' PT data significantly well at $p \leq 0.01$, and did not fit one group's data better than the other's, although there was a small trend for better fit to the nonmusician ratings. Musicians' roughness ratings correlated equivalently well with both analyzers and the model. Nonmusicans' ratings correlated significantly better to the SRA ratings than to the DW or P&L ratings. Scatterplots of the comparisons are shown in Figs. 8(a)–8(f). Figures show that the SRA [see Fig. 8(a) and 8(b)] and P&L [see Fig. 8(e) and 8(f)] correlations were strongest when comparing the roughest dyads against listener ratings. The two obvious outlying data points in both of the SRA comparisons were each m2s at lower root notes in the stimulus set (C#3 and D3). These outliers indicate that the SRA analyzer did not assign a high roughness value to very narrowly spaced PT dyads, unlike these dyads' perceptual assessments. The scatterplot comparing the DW predictions with listeners' ratings [see Fig. 8(c) and 8(d)] shows less of a tendency to agree in any particular region of the scale.

Insert Figure 8 About Here

----

2. Complex-tone, just-tuned dyads

The six CT dyad roughness correlations between participant groups and the analyzers and H&K model were all significantly strong at $p < 0.01$. The SRA provided an equivalent fit to both groups' data (see Fig. 7). The DW analyzer fit the nonmusicians' ratings significantly stronger than it fit the musicians' ratings. Dissonance ratings from the H&K model showed higher correspondence to musicians' ratings but the confidence intervals were short of statistical significance. Of the comparisons to musicians' ratings, the H&K model was a significant improvement over the SRA, which was a significant improvement over the DW. The distinctions were not as sharp between the analyzers/model and the nonmusicians' results. The H&K fit the nonmusicians' data equivalently well to the SRA, which in turn fit the nonmusicians' data equivalently to the DW. Of the best and worst comparisons, the H&K provided a significantly better fit to the nonmusicians' ratings than did the DW analyzer.

Scatterplots [see Figs. 9(a)–9(f)] show correlations between groups' CT dyad rank ordered roughness and the analyzers and model. A comparison of these plots with those in Figs. 8(a)–8(f) reveals the same tendency for better fit at the highest (roughest) rank, but with CT dyads there is also good correlation at the opposite (smoothest) extreme in the SRA and H&K comparisons. The comparatively poorer correlation between these rankings and the analyzers' is evident when Figs. 9(a)–9(f) for CT dyads is compared to Figs. 8(a)–8(f) for PT dyads.

Insert Figure 9 About Here

----

3. Micro-tuned, complex-tone dyads

Correlations between groups' MT dyad roughness ratings and the two analyzers were the poorest of the comparisons here (see Fig. 7). Only one correlation — the SRA and musicians — was significantly strong at the 1% level. The correlation between the SRA and nonmusicians was significant at the 5% level. The SRA's output fit the musicians' rankings significantly better than it did the nonmusicians' rankings. The DW analyzer did not correlate well with either group's MT dyad roughness rankings. Compared to the SRA, the DW results were a significantly poorer fit to musicians' roughness rankings. For nonmusicians, the fit provided by the

DW was not statistically poorer than that provided by the SRA. Scatterplots in Figs. 9(a)–9(d) illustrate the relative lack of fit between the analyzers and listeners' MT dyad assessments. A comparison of the four plots suggests that musicians and the SRA are sensitive to a property of MT dyads that both nonmusicians and the DW analyzer ignored. Only the correlation between the SRA and the musicians' rankings [see Figs. 10(a)–10(d)] shows a linear trend.

Insert Figure 10 About Here

C. Analyzer cross-comparisons

The SRA provided equivalently strong correlations to musicians' dyad rankings regardless of stimulus type (see Fig. 7). For nonmusicians, the SRA provided the highest correlation to this group's PT dyad ratings. This fit was significantly better than to nonmusicians' CT dyad ratings, which in turn was significantly better than the fit to their MT dyad ratings.

This pattern was not observed with the DW analyzer. The DW fit musicians' PT dyad ratings significantly better than it fit their CT dyad ratings, which in turn were significantly better than to musicians' MT dyad ratings. The DW was only marginally better at fitting nonmusicians' PT dyad rankings in comparison with the fit to their CT dyad rankings. The DW's fit to nonmusicians' MT dyad rankings was the poorest of the comparisons made.

VI. DISCUSSION

We addressed the question of how musical expertise influences the evaluated roughness of musical dyads, and compared the results to analyzers and models to learn how they accounted for varying degrees of auditory acuity. The implication of these findings for theories of sensory C/D is assisted by the application of a more controlled experimental protocol and newer methods of statistical analysis than used in earlier studies.

There is much variance among listeners in their capacities for assessing roughness. Individuals' ratings are more reliable when musical events are familiar than when they are unfamiliar, suggesting a role for musical knowledge in the psychophysical scaling of dyad roughness. For listeners with musical expertise, the physical attributes of roughness are diminished by cognitive consonance and augmented by cognitive dissonance. Musicians find it difficult to assess roughness independently of pitch relationships in a musical signal, even in the absence of physical elements responsible for the sensation. Musical expertise may heighten a listener's perceptual awareness of roughness components, but it impedes the ability to ignore attributes of a signal such as frequency ratio relationships that impart meaning in a musical culture.

Lack of agreement among nonexperts increases when pure-tone dyads become complex through the addition of harmonic partials, but musically trained listeners do not show this disparity. These findings indicate that experts are better equipped than nonexperts to isolate and attend to roughness cues from partials. This observation is supported by evidence showing enhanced neurological responses in musicians to upper harmonics, compared with nonmusicians' responses (Lee, Skoe, Kraus, and

Ashley, 2009). Nonexperts are thus more inclined than experts to assess dyad roughness according to the ERB-rate difference between two fundamental frequencies (F0). This comparatively heavier emphasis on F0 relationships has been shown elsewhere; when the fundamental frequency of a tone complex is missing, nonmusicians are much more likely than musicians to regard the lowest available harmonic as fundamental to the tone's pitch (Seither-Preisler *et al.* 2007).

Musicians' assessments of PT dyads narrower than a critical band aligned more closely than did nonmusicians' with Plomp and Levelt's (1965) data on sensory dissonance ratings. Plomp and Levelt's Fig. 10 plots a smooth curve with dyad dissonance rapidly increasing from unison and then smoothly diminishing as the frequency difference between the two tones approaches critical bandwidth. The distribution of averaged roughness ratings from musicians in the present study shows a smooth curve from the narrowest dyad (m2 at C#3) to dyads approaching an ERB-rate scale difference of 1.0 [see Fig. 2(a)], similar to Plomp and Levelt's standardized curve. Nonmusicians, in contrast, show a more linear decrease in roughness ratings in this region [see Fig. 2(b)] and greater variation. The difference between groups is most likely due to experience attending to narrowly spaced intervals as part of musical praxis. Although Plomp and Levelt did not report on the musical training of their participants, only those who gave consistent C/D ratings in a pre-test were used in their experiment, suggesting that their sample could have been composed of an above-average percentage of musically trained listeners.[1]

Averaging the results from listeners can produce the appearance of greater agreement than exists in the data. The smoothness of the roughness curve below critical bandwidth masks strong disagreement among listeners as to which signal generated the absolute roughest sensation. This was evidenced by poor inter-rater agreement (as indexed by the ICC) for the roughness of PT m2s. (Negative ICC values are produced when the scale is reversed, as when one judge applies a score of 10 as a minimum and another applies it to the maximum. A mean score of 5 can hide this discrepancy.) Future studies of roughness or sensory dissonance will benefit from examining listener assessments of simultaneous tones that may be more properly termed as *beating*.

Nonexperts displayed the highest degree of inter-rater agreement for PT dyads — the stimulus set with the least amount of spectral complexity. Spectral information from harmonic partials in the CT and MT dyads somehow increased nonmusicians' response variance, instead of assisting dyad evaluation. Familiar CT dyads were endowed with more cultural relevance than PT dyads, and this may have caused items in successive trials to be perceived as categorically more different from one another, in contrast to the PT dyads. The size of the psychological differences between sequential items is known to affect the precision and consistency of subjective assessments, and when differences are large, disagreement among raters frequently increases (Lockhead, 2004). The observation that nonmusicians showed relatively high *intra*-rater agreement, but lower *inter*-rater agreement for CT dyads suggests that each participant's internal roughness standard, while consistent, did not match the standards adopted by other participants. Because this was not the case for the less ecologically valid PT dyads, this suggests that implicit musical knowledge negatively influences the variance in perceptual assessments.

Despite the relative similarity of the MT dyads, listeners did rate some of these as rougher than others. The rank order of listeners' roughness ratings did not correspond to frequency-ratio complexity. An interesting finding was that nonmusicians' ratings were similar to the normalized rank order of the sensory dissonance of MT (quartertone) dyads reported by Ayers *et al*. (1980). Ayers's participants rated the m2+ and m7+ as the most dissonant, the M3+ and m3+ as the least dissonant, and the uni+ as relatively smooth, as seen in the roughness rankings by nonmusicians in the current experiment. It is not clear what drove this particular ordering. An additional similarity was that Ayers *et al*.'s musically trained participants judged the uni+ to be more dissonant than the adjacent m2+, also reported by the musicians here.

Relationships among frequency spectra are so integral to extracting the pitch of a chord that they provide pitch information even when energy at the root frequency of the chord is missing — the phenomenon that Terhardt (1984) called "virtual pitch." For the most part, the analyzers used here correlated better with listeners' PT dyad ratings over their complex-tone ratings, perhaps reflecting designs that put stronger emphasis on the relationship between the two F0s. The SRA, developed using Sethares's (1998) calculations on the frequency relationship between sine pairs, correlated especially well with ratings from nonexpert listeners. Correlating less well with the SRA were musicians' PT ratings, which were skewed by the tendency to hear (perhaps enhanced) roughness in pitch relationships associated with musical dissonance (e.g., m7s and M7s).

Audio analyzers fine-tuned to detect coordinated activity might be expected to show greater sensitivity toward the *relative smoothness* rather than the roughness components of sounds. PsySound's RoughnessDW analyzer is based on a model by Daniel and Weber (1997), itself a refined version of Aures's (1985) roughness model. Daniel and Weber optimized Aures's model to detect roughness in a wide range of signals from amplitude-modulated tones to noise. The DW analyzer may have mis-estimated the roughness contribution from close partials, due in part to its calculation of critical bandwidth based on the older Bark scale, rather than the more updated ERB-rate scale (Daniel and Weber, 1997). Furthermore, Aures's emphasis on roughness as an acoustic phenomenon was primarily related to amplitude variations over time, and this may have made the model particularly unsuited for evaluating the roughness of musical sounds, where frequency and phase relationships diminish the linearity of the roughness calculation (Cabrera, 2008, Parncutt, 2006, Pressnitzer and McAdams, 1999). Daniel and Weber tested their model in part by gathering subjective roughness assessments for amplitude-modulated tones. The authors reported that the only exception to a good agreement between their model and subjective evaluations of amplitude-modulated tones occurred for carrier frequencies (f_c) below 500 Hz with modulation frequencies (f_{mod}) above 110 Hz — precisely the frequency range of the tones used here.

Nonmusicians are shown to be less adept than musicians at classifying tone complexes by pitch information implied by the tone's harmonic relations, relying instead on the absolute information available in the frequency spectra (Seither-Preisler *et al.* 2007). The observation that the SRA's CT and MT dyad ratings correlated equally well with the musicians' ratings, but strikingly dissimilarly to the nonmusicians' (see Table IV), gives credence to the idea that the ability to extract pitch from dyads accounted for some of the differences observed between musicians and nonmusicians in assigning roughness (Bidelman and Krishnan, 2009; Cariani, 2004; Tsai, 2004). This conclusion is predicated on the assumption that musicians found it easier than nonmusicians to find a pitch for a typical MT dyad. While it is unknown which criterion — pitch stability or number of noncoincidental partials — was used, the similar correlations between musicians' ratings and the SRA for both CT and MT dyads imply consistency, regardless of musical familiarity. The dissimilar correlations between nonmusicians and the SRA for CT and MT dyads imply the opposite — that nonmusicians applied one roughness criterion for familiar intervals and something different for unfamiliar intervals.

VII. CONCLUSION

Listeners' responses to roughness and sensory dissonance are exceptionally task-dependent and prone to contamination from cultural, methodological, and technical inconsistencies (Kreiman *et al*. 2007; Mashinter, 2006; Parncutt, 2006; Plomp and Steeneken, 1968). The current study on evaluated roughness was motivated in part by the opportunity to collect data under a stricter protocol, using more advanced audio technology than available in earlier decades. Replacing the analog tape machine and loudspeaker playback used in much of the seminal C/D work with digital audio sources and high-quality headphones circumvented the sound coloration that may have contributed error to earlier findings. Limiting the frequency range of dyads helped to reduce sharpness (a component of sensory dissonance; Terhardt, 1974b), although this subjective measure was not considered entirely eliminated. Higher pitched dyads were somewhat piercing and how well listeners avoided confounding sharpness with roughness, despite being instructed in the difference, is unknown. Reducing the dissonance factor of loudness by amplitude-normalizing dyads permitted a closer look at the confounded variables of tonality and roughness, thus optimizing the usefulness of the current finding in future models of sensory dissonance.

A limitation of the present finding is that only one sound source (alto saxophone) was used in the complex-tone experiments. This provided necessary control of the phase relationships between partials, but it is noted that the rank order of roughness ratings, as well as the correlations between the analyzers and evaluated roughness may be different with other timbres. Future work should include roughness assessments of other musical as well as nonmusical sounds using participant groups segregated by expertise in music, with the goal of accurately modeling the perceptual differences between sensory and cognitive dissonance.

ACKNOWLEDGEMENTS

This article is submitted in partial fulfillment of the requirements for the Ph.D. in Experimental Psychology by the first author.

This work owes a debt of gratitude to Bennett Smith who made it possible for the first author to use and interpret the RoughnessDW algorithm in PsySound 3. In addition, Bennett put a large amount of effort into our exploration of other software-based models that, for technical reasons beyond our control, were not used.

Densil Cabrera provided insights into the research question and offered comments and instruction on the Roughness DW algorithm. We would like to thank research assistants Mattson Ogg and Sabrina Lytton who were a great help with the data collection. Thanks also go to the following colleagues who offered feedback on interpreting and analyzing our results: Bruno Giordano, Bruno Gingras, Georgios Marentakis, Manuela Marin, Nils Peters, Jake Shenker, Finn Upham, Michel Vallières, and Doug Van Nort.

The authors gratefully acknowledge the support of grants from NSERC and Bell Canada to Daniel J. Levitin (graduate co-advisor), grants from NSERC and a Canada Research Chair to SMc, and a CIRMMT Student Award to SER. Portions of this work were presented at the convention of the Society for Music Perception and Cognition, Indianapolis, IN, August, 2009.

APPENDIX

Current dissonance models have difficulty predicting the roughness of both musical sounds and noise (Leman, 2000; Parncutt, 2006). Leman (2000) described two kinds of roughness models. *Curve fitting* model types have mapped psychoacoustic judgments of sensory dissonance to activity in the auditory periphery (Guirao and Garavilla, 1976; Hutchinson and Knopoff, 1978; Kameoka and Kuriyagawa, 1969; Leman, 2000; Mashinter, 2006; Plomp and Levelt, 1965; Porres and Jônatas, 2007; Sethares, 1993; Skovenborg and Nielsen, 2002). These models attempt to account for the sensory dissonance component of musical dissonance. The purely *auditory model* type is used for theoretical analysis and estimates roughness from human auditory processing characteristics. Unlike curve fitting models, auditory models attempt to predict the roughness of a wide range of sounds including (or in some cases especially) noise (Aures, 1985; Daniel and Weber, 1997; Pressnitzer and McAdams, 1997; Vassilakis, 2007).

PsySound 3 is a software program implementing the roughness algorithm of Daniel and Weber (1997) in MATLAB (MathWorks, Natick, MA) as "RoughnessDW" to analyze sound files. As of this writing, the software is available online at http://psysound.wikidot.com/. The reader is referred to Daniel and Weber (1997) for a thorough description, but briefly, the model sums the energy of beating frequencies in 47 overlapping critical band filters. It calculates the modulation frequency, carrier frequency, and depth of modulation in each channel. Phase effects are taken into account by correlating the outputs of all 47 channels. The model is optimized to describe the roughness of amplitude modulated tones and unmodulated noise but is reported to accurately predict the roughness of frequency modulated tones and white noise as well. The authors of the model report that subjective data from several studies provided a good fit with the model's output in most applications (Daniel and Weber, 1997). The DW analyzer generates four roughness values for each item at, in this case, 0.0, 180, 370, and 500 ms corresponding to points at the onset, middle, and offset of each 500-ms dyad. Based on a suggestion from the algorithm's programmer, we discarded the roughness values at the onset and offset times and derived a roughness value for each dyad by averaging the values at 180 and 370 ms.

The Spectral Roughness Analyzer (SRA) (Vassilakis, 2007) is a web-based application for calculating the roughness component of sound signals, implementing the algorithm of Fulop and Fitz (2006). The SRA model is available online at http://www.acousticslab.org/roughness/index.html. The model calculates the roughness of complex-tone signals by summing the roughness contributions from all sine-pairs. The calculation is based in large part on the subjective roughness of pure-tone pairs reported by Sethares (1998). Compared with other models, SRA places more emphasis on the relative amplitude differences between signal components and less emphasis on the absolute sound pressure level. Phase parameters are not included in the analysis, because the model is optimized for comparing the relative roughness of sounds with phase relationships that are either the same for all members of the set, or entirely random (Vassilakis, 2007). Sound files can be uploaded to the model online and several parameters can be optimized for best accuracy. The SRA model calculated the overall roughness value of each 500-ms dyad from 5 values at 150, 200, 250, 300, and 350 ms.

In addition to the auditory models described above, we extracted roughness values from Plomp and Levelt's (1965) Fig. 9 "dissonance band" for comparison with our listeners' pure-tone roughness ratings. First, the mean frequency of each dyad was calculated. Next, numerical values in Hertz corresponding to the “frequency difference of (the) smallest consonant interval of two simple tones as a function of the mean frequency of the tones” were derived from Plomp and Levelt’s Figure 9, interpolating when necessary. Depending on a dyad’s mean frequency, the appropriate “P&L criterion” was subtracted from the dyad’s frequency difference in Hertz ($f_2 - f_1$), generating either a positive or negative number. In this way each dyad was labeled as falling either outside or inside of Plomp and Levelt’s “dissonant band.”

Because sensory C/D is based on frequency relationships *relative* to critical bandwidth, it was necessary to label each dyad by how far it exceeded or receded into the “dissonance band.” This was accomplished by dividing $f_2 - f_1$ by the P&L criterion to yield a ratio. These ratios could be ranked by size, indicating the degree of a dyad’s sensory consonance or dissonance, or in this case, presumed roughness.

Stimulus sets for Experiments 1, 2, and 3 are fully described in Tables A.I., A.II., and A.III, respectively.

Footnote

1. Plomp and Steeneken (1968) corrected problems with Plomp and Levelt's (1965) methodology and collected sensory dissonance data from more participants ($N = 20$) under conditions closer to the methods used in the present experiment — dyads presented at 60 phon (60 dB SPL), using headphones. The 1968 data should be considered more reliable than the 1965 data (Donald D. Greenwood, 1999, communication on the Auditory List, www.auditory.org/postings/1999/289.html).

TABLE I: Experiment 1: Pure-tone dyads. Intra-rater reliability as indexed by Cronbach's coefficient alpha (ζ) including lower (LL) and upper (UL) confidence limits.

	Musicians			Nonmusicians		
Interval	ζ	LL	UL	ζ	LL	UL
m2	***0.89****	0.78	1.00	*0.80*	0.67	0.93
M2	*0.55*	0.43	0.67	*0.78*	0.61	0.95
m3	*0.76*	0.66	0.86	*0.52*	0.32	0.72
M3	*0.78*	0.64	0.92	*0.56*	0.41	0.71
p4	*0.69*	0.54	0.84	*0.54*	0.37	0.71
tri	*0.78*	0.68	0.88	*0.71*	0.60	0.82
p5	*0.75*	0.61	0.89	*0.32*	0.14	0.50
m6	*0.74*	0.62	0.86	*0.56*	0.39	0.73
M6	*0.72*	0.62	0.82	*0.76*	0.55	0.97
m7	***0.83****	0.74	0.92	*0.74*	0.64	0.84
M7	*0.78*	0.66	0.90	*0.77*	0.61	0.93
oct	***0.86****	0.72	1.00	***0.89****	0.80	0.98

Note. $N = 15$ for each group. Coefficient alpha ≥ 0.70 was adopted as a satisfactory level of modest intra-rater reliability (Nunnally and Bernstein, 1994, pp. 264–265). Bold type and an asterisk highlights cases where the lower limit of the 95% confidence interval was ≥ 0.70. Following Feldt, Woodruff, and Salih (1987), lower case zeta (ζ) was adopted to denote coefficient alpha and avoid confusion with alpha (α) used in significance tests.
*$p < 0.05$.

TABLE II: Experiment 1: Pure-tone dyads. Inter-rater reliability as indexed by the intraclass correlation coefficient (ICC), including the lower (LL) and upper (UL) limits of the 95% confidence intervals.

	Musicians				Nonmusicians			
Interval	ICC	LL	UL	*F*(5, 70)	ICC	LL	UL	*F*(5, 70)
m2	-0.06	-0.07	0.00	0.16	-0.05	-0.06	0.06	0.32
M2	-0.01	-0.05	0.22	0.86	0.04	-0.03	0.38	1.64
m3	**0.10**	0.00	0.51	2.68*	0.05	-0.02	0.41	1.85
M3	-0.00	-0.04	0.25	0.97	**0.19**	0.04	0.64	4.55**
p4	**0.14**	0.02	0.57	3.44**	**0.28**	0.09	0.74	7.00**
tri	0.02	-0.04	0.30	1.23	**0.20**	0.05	0.66	4.85**
p5	0.07	-0.02	0.44	2.13	**0.29**	0.10	0.74	7.13**
m6	**0.37**	0.15	0.80	9.92**	**0.21**	0.05	0.66	4.90**
M6	0.00	-0.04	0.26	1.02	**0.29**	0.09	0.74	7.06**
m7	0.06	-0.02	0.41	1.87	**0.10**	0.00	0.50	2.62*
M7	0.06	-0.02	0.42	1.96	0.06	-0.02	0.41	1.90
oct	0.05	-0.02	0.39	1.76	0.06	-0.02	0.43	2.04

Note. F tests determined whether the observed 95% confidence intervals were within the population value (ρ) of this intraclass correlation; significant ICCs are highlighted in bold type (Shrout and Fleiss, 1979). $N = 15$ for each group.
*$p \leq 0.05$. **$p \leq 0.01$.

TABLE III: Experiment 2: Complex-tone dyads. Intra-rater reliability as indexed by Cronbach's coefficient alpha (ζ) including lower (LL) and upper (UL) confidence limits and tests of equality of means.

	Musicians			Nonmusicians		
Interval	ζ	LL	UL	ζ	LL	UL
m2	***0.86****	0.73	0.99	***0.83****	0.73	0.93
M2	*0.74*	0.62	0.86	*0.57*	0.36	0.78
m3	*0.60*	0.44	0.76	*0.56*	0.38	0.74
M3	*0.62*	0.45	0.79	*0.67*	0.53	0.81
p4	*0.64*	0.50	0.78	*0.41*	0.29	0.53
tri	*0.76*	0.60	0.92	*0.61*	0.38	0.84
p5	*0.77*	0.63	0.91	*0.53*	0.33	0.73
m6	*0.53*	0.37	0.69	*0.76*	0.56	0.96
M6	*0.67*	0.53	0.81	*0.74*	0.64	0.84
m7	*0.78*	0.65	0.91	*0.76*	0.56	0.96
M7	*0.81*	0.69	0.93	*0.77*	0.59	0.95
oct	*0.86*	0.67	1.05	***0.85****	0.72	0.98

Note. $N = 15$ for each group.
*$p \leq 0.05$.

TABLE IV. Experiment 2: Complex-tone dyads. Intraclass correlation coefficients and the lower (LL) and upper (UL) limits of the 95% confidence intervals.

	Musicians				Nonmusicians			
Interval	ICC	LL	UL	*F*(5, 70)	ICC	LL	UL	*F*(5, 70)
m2	0.07	-0.02	0.44	2.06	0.01	-0.04	0.28	1.12
M2	**0.12**	0.00	0.53	2.94*	-0.01	-0.05	0.22	0.88
m3	**0.08**	-0.01	0.47	2.34*	0.03	-0.03	0.36	1.52
M3	-0.03	-0.06	0.13	0.54	-0.04	-0.06	0.08	0.38
p4	-0.05	-0.06	0.06	0.33	0.03	-0.03	0.35	1.47
tri	0.02	-0.04	0.30	1.24	-0.02	-0.05	0.18	0.67
p5	-0.03	-0.06	0.13	0.54	-0.04	-0.06	0.09	0.41
m6	0.04	-0.03	0.38	1.70	0.04	-0.03	0.38	1.65
M6	-0.02	-0.05	0.19	0.74	-0.02	-0.05	0.17	0.64
m7	-0.04	-0.06	0.09	0.84	-0.00	-0.05	0.24	0.95
M7	0.04	-0.03	0.38	1.66	-0.01	-0.05	0.22	0.84
oct	-0.02	-0.05	0.19	0.75	0.01	-0.04	0.27	1.07

Note. $N = 15$ for each group.
$^*p \leq 0.05$.

TABLE V: Experiment 3: Microtuned complex-tone dyads. Intra-rater reliability as indexed by Cronbach's coefficient alpha (ζ) including lower (LL) and upper (UL) confidence limits and tests of equality of means.

	Musicians			Nonmusicians		
Interval	ζ	LL	UL	ζ	LL	UL
uni+	***0.83****	0.71	0.95	*0.76*	0.63	0.89
m2+	*0.79*	0.68	0.90	*0.46*	0.30	0.62
M2+	***0.79****	0.71	0.87	*-1.12*	-1.29	-0.95
m3+	*0.26*	0.14	0.38	*-0.18*	-0.35	-0.01
M3+	*0.53*	0.39	0.67	*0.16*	0.02	0.30
p4+	*-0.57*	-0.75	-0.39	*0.39*	0.22	0.56
tri+	*0.49*	0.35	0.63	*0.29*	0.14	0.44
p5+	*-0.59*	-0.77	-0.41	*0.73*	0.61	0.85
m6+	*0.62*	0.52	0.72	*0.71*	0.58	0.84
M6+	*0.40*	0.24	0.56	*0.69*	0.53	0.85
m7+	*0.48*	0.29	0.67	*0.62*	0.49	0.75
M7+	***0.87****	0.81	0.93	*0.38*	0.26	0.50

Note. $N = 15$ for each group.
*$p \leq 0.05$.

TABLE VI: Experiment 3: Microtuned complex-tone dyads. Intraclass correlation coefficients and the lower (LL) and upper (UL) limits of the 95% confidence intervals.

	Musicians				Nonmusicians			
Interval	ICC	LL	UL	*F*(5, 70)	ICC	LL	UL	*F*(5, 70)
uni+	0.02	-0.04	0.30	1.22	-0.02	-0.05	0.18	0.68
m2+	-0.04	-0.06	0.09	0.41	0.06	-0.02	0.42	1.96
M2+	**0.15**	0.02	0.58	3.60*	0.02	-0.04	0.32	1.33
m3+	**0.09**	-0.01	0.48	2.44*	0.01	-0.04	0.28	1.14
M3+	0.02	-0.04	0.32	1.35	0.01	-0.04	0.30	1.20
p4+	-0.01	-0.05	0.23	0.90	0.06	-0.02	0.43	2.04
tri+	**0.10**	0.00	0.52	2.77*	0.02	-0.04	0.32	1.34
p5+	0.06	-0.02	0.42	1.97	0.03	-0.03	0.34	1.43
m6+	0.01	-0.04	0.28	1.12	-0.01	-0.05	0.22	0.86
M6+	-0.01	-0.05	0.21	0.80	0.07	-0.02	0.44	2.10
m7+	0.05	-0.02	0.40	1.80	0.02	-0.04	0.31	1.25
M7+	0.06	-0.02	0.42	1.93	-0.00	-0.05	0.24	0.95

Note. $n = 15$ for each group.
$*p \leq 0.05$.

TABLE VII. Post-hoc results following significant one-way repeated-measures ANOVAs comparing the six mean dyad roughness ratings within each interval class. Only significantly different pairs are shown.

Pure-tone, just-tuned dyads				
	Musicians		Nonmusicians	
Interval	(rougher) root note	(smoother) root note	(rougher) root note	(smoother) root note
m3	F3	A#4		
M3			D3	D#4, E4, G#4
p4			C3	E3, F3, C4, D4, A4
tri			D#3, A#3	F4
p5			C3, F#3, G3	A4, B4
m6	C3, F#3, B3	D#4, F#4, G4	C3	F#3, D#4, F#4, G4
M6			D#3, A3	F4, F#4, G#4

Complex-tone, just-tuned dyads		
	Musicians	
Interval	(rougher) root note	(smoother) root note
M2	C#3	D4

Microtuned, complex-tone dyads		
	Musicians	
Interval	(rougher) root note	(smoother) root note
M2+	D#3	C4
tri+	C3	A3, B4

TABLE A.I. Stimulus set for Experiment 1: Just-tuned, pure-tone dyads.

musical interval	lower note	upper note	f_1 (Hz)	f_2 (Hz)	$f_2 - f_1$ (Hz)	integer ratio
p5	C3	G3	130.8	196.0	65.2	2:3
m6	C3	G#3	130.8	207.6	76.8	5:8
p4	C3	F3	130.8	174.6	43.8	3:4
M2	C#3	D#3	138.6	155.6	17.0	8:9
M7	C#3	C4	138.6	261.6	123.0	8:15
m2	C#3	D3	138.6	146.8	8.2	15:16
M3	D3	F#3	146.8	185.0	38.2	4:5
m2	D3	D#3	146.8	155.6	8.8	15:16
m7	D3	C4	146.8	261.6	114.8	9:16
tri	D#3	A3	155.6	220.0	64.4	32:45
M6	D#3	C4	155.6	261.6	106.0	3:5
M7	D#3	D4	155.6	293.6	138.0	8:15
m2	E3	F3	164.8	174.6	9.8	15:16
p4	E3	A3	164.8	220.0	55.2	3:4
oct	E3	E4	164.8	329.6	164.8	1:2
p4	F3	A#3	174.6	233.1	58.5	3:4
M2	F3	G3	174.6	196.0	21.4	8:9
m3	F3	G#3	174.6	207.6	33.0	5:6
m6	F#3	D4	185.0	293.6	108.6	5:8
p5	F#3	C#4	185.0	277.2	92.2	2:3
tri	F#3	C4	185.0	261.6	76.6	32:45
m7	G3	F4	196.0	349.2	153.2	9:16
oct	G3	G4	196.0	392.0	196.0	1:2
p5	G3	D4	196.0	293.6	97.6	2:3
m3	G#3	B3	207.6	246.9	39.3	5:6
m7	G#3	F#4	207.6	370.0	162.4	9:16
M3	G#3	C4	207.6	261.6	54.0	4:5
oct	A3	A4	220.0	440.0	220.0	1:2
m3	A3	C4	220.0	261.6	41.6	5:6
M6	A3	F#4	220.0	370.0	150.0	3:5
M6	A#3	G4	233.1	392.0	158.9	3:5
tri	A#3	E4	233.1	329.6	96.5	32:45
M2	A#3	C4	233.1	261.6	28.5	8:9
M7	B3	A#4	246.9	466.2	219.3	8:15
M3	B3	D#4	246.9	311.2	64.3	4:5
m6	B3	G4	246.9	392.0	145.1	5:8
p4	C4	F4	261.6	349.2	87.6	3:4
m7	C4	A#4	261.6	466.2	204.6	9:16
M2	C4	D4	261.6	293.6	32.0	8:9
oct	C#4	C#5	277.2	554.4	277.2	1:2
tri	C#4	G4	277.2	392.0	114.8	32:45
M7	C#4	C5	277.2	523.5	246.3	8:15
tri	D4	G#4	293.6	415.2	121.6	32:45
p4	D4	G4	293.6	392.0	98.4	3:4
m3	D4	F4	293.6	349.2	55.6	5:6
m6	D#4	B4	311.2	493.8	182.6	5:8

M3	D#4	G4	311.2	392.0	80.8	4:5
oct	D#4	D#5	311.2	622.3	311.1	1:2
M2	E4	F#4	329.6	370.0	40.4	8:9
M7	E4	D#5	329.6	622.3	292.7	8:15
M3	E4	G#4	329.6	415.2	85.6	4:5
M6	F4	D5	349.2	587.3	238.1	3:5
oct	F4	F5	349.2	698.4	349.2	1:2
tri	F4	B4	349.2	493.8	144.6	32:45
m7	F#4	E5	370.0	659.2	289.2	9:16
m6	F#4	D5	370.0	587.3	217.3	5:8
M6	F#4	D#5	370.0	622.3	252.3	3:5
m3	G4	A#4	392.0	466.2	74.2	5:6
m2	G4	G#4	392.0	415.2	23.2	15:16
m6	G4	D#5	392.0	622.3	230.3	5:8
M3	G#4	C5	415.2	523.5	108.3	4:5
M6	G#4	F5	415.2	698.4	283.2	3:5
p5	G#4	D#5	415.2	622.3	207.1	2:3
m2	A4	A#4	440.0	466.2	26.2	15:16
p5	A4	E5	440.0	659.2	219.2	2:3
p4	A4	D5	440.0	587.3	147.3	3:4
M7	A#4	A5	466.2	880.0	413.8	8:15
m3	A#4	C#5	466.2	554.4	88.2	5:6
m7	A#4	G#5	466.2	830.6	364.4	9:16
p5	B4	F#5	493.8	740.0	246.2	2:3
M2	B4	C#5	493.8	554.4	60.6	8:9
m2	B4	C5	493.8	523.5	29.7	15:16

TABLE A.II. Stimulus set for Experiment 2: Just-tuned, complex-tone dyads

musical interval	lower note	upper note	f_1 (Hz)	f_2 (Hz)	$f_2 - f_1$ (Hz)	integer ratio
m6	C3	G#3	130.8	207.6	76.8	5:8
m3	C3	D#3	130.8	155.6	24.8	5:6
m7	C3	A#3	130.8	233.1	102.3	9:16
M2	C#3	D#3	138.6	155.6	17.0	8:9
tri	C#3	G3	138.6	196.0	57.4	32:45
M7	C#3	C4	138.6	261.6	123.0	8:15
oct	D3	D4	146.8	293.6	146.8	1:2
p5	D3	A3	146.8	220.0	73.2	2:3
p4	D3	G3	146.8	196.0	49.2	3:4
tri	D#3	A3	155.6	220.0	64.4	32:45
oct	D#3	D#4	155.6	311.2	155.6	1:2
M2	D#3	F3	155.6	174.6	19.0	8:9
M3	E3	G#3	164.8	207.6	42.8	4:5
M2	E3	F#3	164.8	185.0	20.2	8:9
m6	E3	C4	164.8	261.6	96.8	5:8
m7	F3	D#4	174.6	311.2	136.6	9:16
M7	F3	E4	174.6	329.6	155.0	8:15
m2	F3	F#3	174.6	185.0	10.4	15:16
p5	F#3	C#4	185.0	277.2	92.2	2:3
p4	F#3	B3	185.0	246.9	61.9	3:4
oct	F#3	F#4	185.0	370.0	185.0	1:2
p4	G3	C4	196.0	261.6	65.6	3:4
m7	G3	F4	196.0	349.2	153.2	9:16
M6	G3	E4	196.0	329.6	133.6	3:5
m3	G#3	B3	207.6	246.9	39.3	5:6
m2	G#3	A3	207.6	220.0	12.4	15:16
tri	G#3	D4	207.6	293.6	86.0	32:45
m2	A3	A#3	220.0	233.1	13.1	15:16
M6	A3	F#4	220.0	370.0	150.0	3:5
M3	A3	C#4	220.0	277.2	57.2	4:5
M7	A#3	A4	233.1	440.0	206.9	8:15
M3	A#3	D4	233.1	293.6	60.5	4:5
p5	A#3	F4	233.1	349.2	116.1	2:3
M6	B3	G#4	246.9	415.2	168.3	3:5
m6	B3	G4	246.9	392.0	145.1	5:8
m3	B3	D4	246.9	293.6	46.7	5:6
oct	C4	C5	261.6	523.5	261.9	1:2
M7	C4	B4	261.6	493.8	232.2	8:15
m2	C4	C#4	261.6	277.2	15.6	15:16
p4	C#4	F#4	277.2	370.0	92.8	3:4
M3	C#4	F4	277.2	349.2	72.0	4:5
m6	C#4	A4	277.2	440.0	162.8	5:8
M7	D4	C#5	293.6	554.4	260.8	8:15
M2	D4	E4	293.6	329.6	36.0	8:9
m3	D4	F4	293.6	349.2	55.6	5:6
m6	D#4	B4	311.2	493.8	182.6	5:8
p4	D#4	G#4	311.2	415.2	104.0	3:4

m7	D#4	C#5	311.2	554.4	243.2	9:16
m2	E4	F4	329.6	349.2	19.6	15:16
tri	E4	A#4	329.6	466.2	136.6	32:45
p5	E4	B4	329.6	493.8	164.2	2:3
tri	F4	B4	349.2	493.8	144.6	32:45
m6	F4	C#5	349.2	554.4	205.2	5:8
M3	F4	A4	349.2	440.0	90.8	4:5
m3	F#4	A4	370.0	440.0	70.0	5:6
M6	F#4	D#5	370.0	622.3	252.3	3:5
M7	F#4	F5	370.0	698.4	328.4	8:15
M3	G4	B4	392.0	493.8	101.8	4:5
m2	G4	G#4	392.0	415.2	23.2	15:16
M6	G4	E5	392.0	659.2	267.2	3:5
M2	G#4	A#4	415.2	466.2	51.0	8:9
oct	G#4	G#5	415.2	830.6	415.4	1:2
p4	G#4	C#5	415.2	554.4	139.2	3:4
p5	A4	E5	440.0	659.2	219.2	2:3
m3	A4	C5	440.0	523.5	83.5	5:6
M2	A4	B4	440.0	493.9	53.9	8:9
M6	A#4	G5	466.2	784.0	317.8	3:5
m7	A#4	G#5	466.2	830.6	364.4	9:16
oct	A#4	A#5	466.2	932.3	466.1	1:2
m7	B4	A5	493.8	880.0	386.2	9:16
p5	B4	F#5	493.8	740.0	246.2	2:3
tri	B4	F5	493.8	698.4	204.6	32:45

TABLE A.III. Stimulus set for Experiment 3: Microtuned, complex-tone dyads

Musical interval	lower note	upper note	f_1 (Hz)	f_2 (Hz)	$f_2 - f_1$ (Hz)	integer ratio
M6+	C3	A3+	130.8	226.4	95.6	45:26
tri+	C3	F#3+	130.8	190.4	59.6	16:11
M7+	C3	B3+	130.8	254.1	123.3	68:35
p4+	C#3	F#3+	138.6	190.4	51.8	11:8
M3+	C#3	F3+	138.6	179.7	41.1	35:27
m2+	C#3	D3+	138.6	151.1	12.5	12:11
m3+	D3	F3+	146.8	179.7	32.9	11:9
uni+	D3	D3+	146.8	151.1	4.3	35:34
p5+	D3	A3+	146.8	226.4	79.6	37:24
m6+	D#3	B3+	155.6	254.1	98.5	49:30
M2+	D#3	F3+	155.6	179.7	24.1	15:13
m7+	D#3	C#4+	155.6	285.3	129.7	11:6
m2+	E3	F3+	164.8	179.7	14.9	12:11
M6+	E3	C#4+	164.8	285.3	120.5	45:26
M3+	E3	G#3+	164.8	213.7	48.9	35:27
p5+	F3	C4+	174.6	269.3	94.7	37:24
m6+	F3	C#4+	174.6	285.3	110.7	49:30
p4+	F3	A#3+	174.6	239.9	65.3	11:8
tri+	F#3	C4+	185.0	269.3	84.3	16:11
M7+	F#3	F4+	185.0	359.4	174.4	68:35
m3+	F#3	A3+	185.0	226.4	41.4	11:9
M7+	G3	F#4+	196.0	380.8	184.8	68:35
p4+	G3	C4+	196.0	269.3	73.3	11:8
m6+	G3	D#4+	196.0	320.3	124.3	49:30
uni+	G#3	G#3+	207.6	213.7	6.1	35:34
m7+	G#3	F#4+	207.6	380.8	173.2	11:6
M2+	G#3	A#3+	207.6	239.9	32.3	15:13
M3+	A3	C#4+	220.0	285.3	65.3	35:27
m2+	A3	A#3+	220.0	239.9	19.9	12:11
tri+	A3	D#4+	220.0	320.3	100.3	16:11
m7+	A#3	G#4+	233.1	427.4	194.3	11:6
p5+	A#3	F4+	233.1	359.4	126.3	37:24
M6+	A#3	G4+	233.1	403.5	170.4	45:26
M2+	B3	C#4+	246.9	285.3	38.4	15:13
m3+	B3	D4+	246.9	302.2	55.3	11:9
uni+	B3	B3+	246.9	254.1	7.2	35:34
tri+	C4	F#4+	261.6	380.8	119.2	16:11
M2+	C4	D4+	261.6	302.2	40.6	15:13
m3+	C4	D#4+	261.6	320.3	58.7	11:9
M7+	C#4	C5+	277.2	538.5	261.3	68:35
p5+	C#4	G#4+	277.2	427.4	150.2	37:24
tri+	C#4	G4+	277.2	403.5	126.3	16:11
p5+	D4	A4+	293.6	452.9	159.3	37:24
m7+	D4	C5+	293.6	538.5	244.9	11:6
M7+	D4	C#5+	293.6	570.6	277.0	68:35
m2+	D#4	E4+	311.2	339.3	28.1	12:11
M7+	D#4	D5+	311.2	604.4	293.2	68:35

M3+	D#4	G4+	311.2	403.5	92.3	35:27
m6+	E4	C5+	329.6	538.5	208.9	49:30
p4+	E4	A4+	329.6	452.9	123.3	11:8
uni+	E4	E4+	329.6	339.3	9.7	35:34
M6+	F4	D5+	349.2	604.4	255.2	45:26
m3+	F4	G#4+	349.2	427.4	78.2	11:9
m2+	F4	F#4+	349.2	380.8	31.6	12:11
p4+	F#4	B4+	370.0	508.3	138.3	11:8
M6+	F#4	D#5+	370.0	640.6	270.6	45:26
m6+	F#4	D5+	370.0	604.4	234.4	49:30
M2+	G4	A4+	392.0	452.9	60.9	15:13
M3+	G4	B4+	392.0	508.3	116.3	35:27
M6+	G4	E5+	392.0	678.5	286.5	45:26
M3+	G#4	C5+	415.2	538.5	123.3	35:27
m2+	G#4	A4+	415.2	452.9	37.7	12:11
p5+	G#4	D#5+	415.2	640.6	225.4	37:24
m3+	A4	C5+	440.0	538.5	98.5	11:9
uni+	A4	A4+	440.0	452.9	12.9	35:34
m7+	A4	G5+	440.0	807.0	367.0	11:6
uni+	A#4	A#4+	466.2	479.9	13.7	35:34
m6+	A#4	F#5+	466.2	761.7	295.5	49:30
M2+	A#4	C5+	466.2	538.5	72.3	15:13
m7+	B4	A5+	493.8	905.8	412.0	11:6
tri+	B4	F5+	493.8	718.9	225.1	16:11
p4+	B4	E5+	493.8	678.5	184.7	11:8

Figure Captions

FIG. 1. The distribution of roughness ratings for pure-tone musical intervals: musicians (a) and nonmusicians (b). Dark horizontal bars represent the median roughness rating for each group ($N = 15$); the gray areas above and below the medians represent the upper and lower quartiles of the range of responses. The minima and maxima for all intervals were 0.0 and 1.0 respectively, and are omitted here for clarity.

FIG. 2. Mean roughness ratings for pure-tone dyads as a function of ERB-rate difference between the upper and lower tones: musicians (a) and nonmusicians (b). Seventy-two dyads are represented by symbols grouped according to musical interval.

FIG. 3. The distribution of roughness ratings for complex-tone, just-tuned dyads: musicians (a) and nonmusicians (b), (see Fig. 1 caption).

FIG. 4. Mean roughness ratings for complex-tone, just-tuned dyads: musicians (a) and nonmusicians (b).

FIG. 5. The distribution of roughness ratings for microtuned, complex-tone dyads musicians (a) and nonmusicians (b), (see Fig. 1 caption).

***FIG.* 6.** Mean roughness ratings for microtuned, complex-tone dyads: musicians (a) and nonmusicians (b).

FIG. 7. Pearson product-moment correlations (r), including lower (LL) and upper (UL) limits of the 95% confidence intervals, between analyzers or models and participants' ratings, derived using a bootstrap method by sampling 1000 times with replacement.

FIG. 8. Scatterplots of the pure-tone dyad correlations between analyzers or a model and the rank order of mean roughness ratings from each group ($N = 15$) for 72 dyads. The rank is from 1 (smoothest or most consonant) to 72 (roughest or most dissonant). Spearman rank correlation coefficients are included. All values of rho are significant at $p < 0.01$. Top panel: SRA and (a) musicians, (b) nonmusicians; middle panel: DW analyzer and (c) musicians, (d) nonmusicians; bottom panel: Plomp and Levelt's (1965) Fig. 9 sensory dissonance rank order curve and (e) musicians, (f) nonmusicians.

FIG. 9. Scatterplots for complex-tone, just-tuned dyads (see Fig. 8 caption). All values of rho are significant at $p < 0.01$. Top panel: SRA and (a) musicians, (b) nonmusicians; middle panel: DW analyzer and (c) musicians, (d) nonmusicians; Hutchinson and Knopoff's (1978) predicted rank order of sensory dissonance for complex-tone Western dyads and (e) musicians, (f) nonmusicians.

FIG. 10. Scatterplots for complex-tone, microtuned dyad correlations (see Fig. 8 caption). Top panel: SRA and (a) musicians, (b) nonmusicians; bottom panel: DW analyzer and (c) musicians, (d) nonmusicians.

* = Nonsignificant $p < 0.01$; remaining values of rho are significant at this level.

FIG. 1

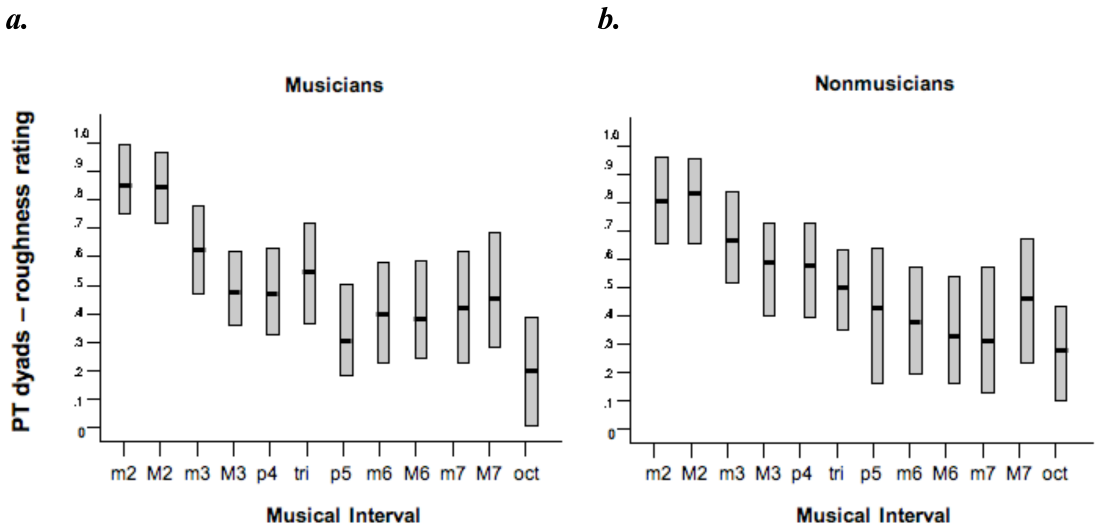
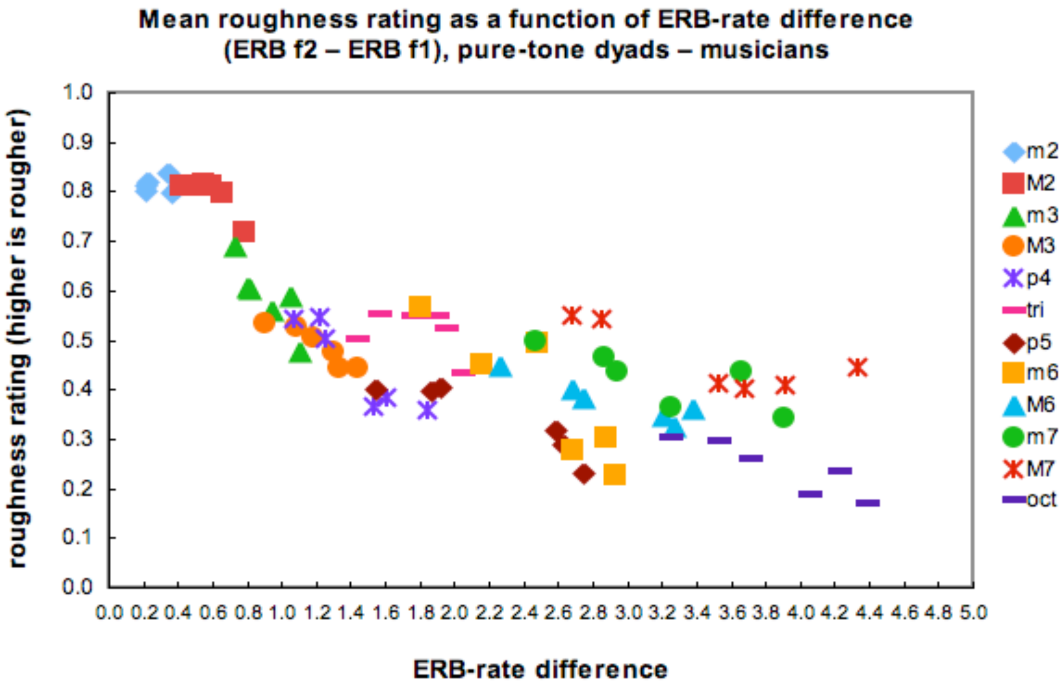

a.
b.
Musicians
Nonmusicians
PT dyads – roughness rating
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
m2 M2 m3 M3 p4 tri p5 m6 M6 m7 M7 oct
Musical Interval

FIG. 2

a.


Mean roughness rating as a function of ERB-rate difference
(ERB f2 – ERB f1), pure-tone dyads – musicians
roughness rating (higher is rougher)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0
ERB-rate difference
m2
M2
m3
M3
p4
tri
p5
m6
M6
m7
M7
oct

b.

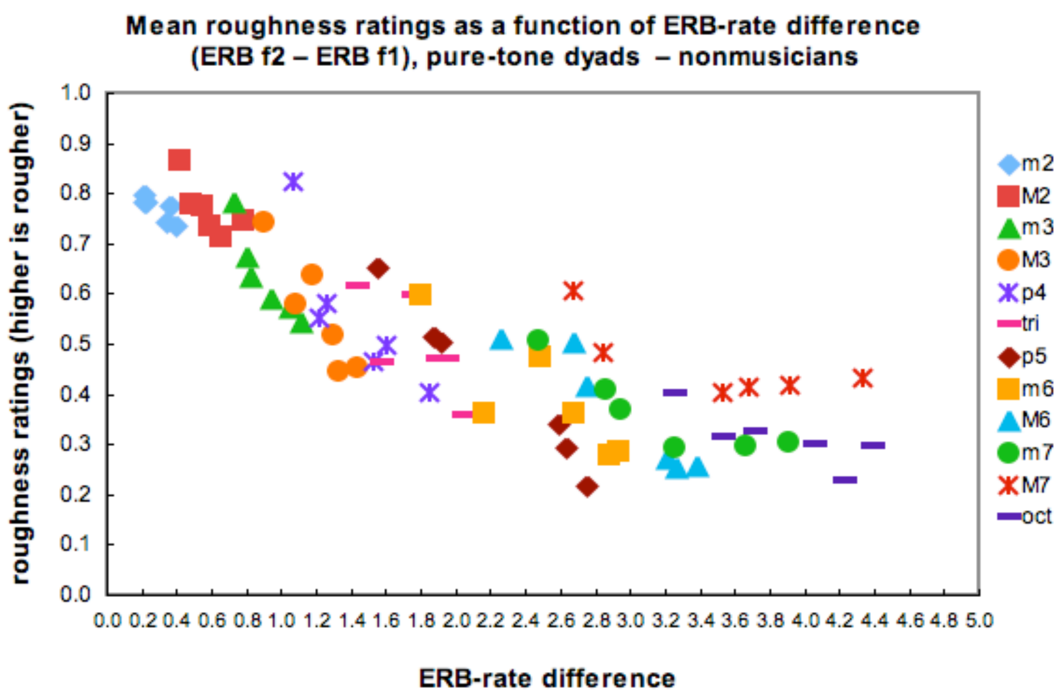

Mean roughness ratings as a function of ERB-rate difference
(ERB f2 – ERB f1), pure-tone dyads – nonmusicians
roughness ratings (higher is rougher)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0
ERB-rate difference
m2
M2
m3
M3
p4
tri
p5
m6
M6
m7
M7
oct

FIG. 3

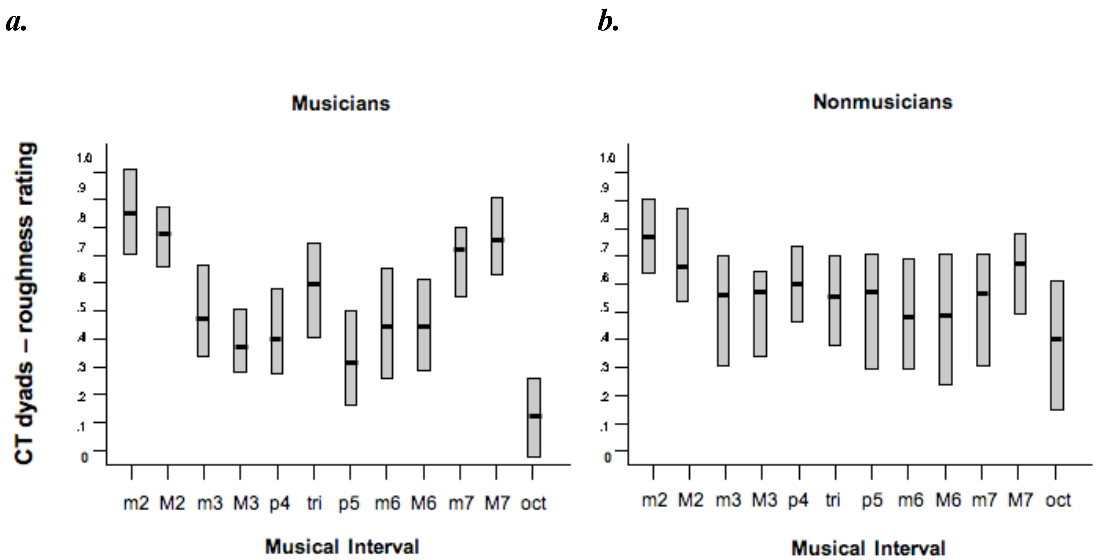
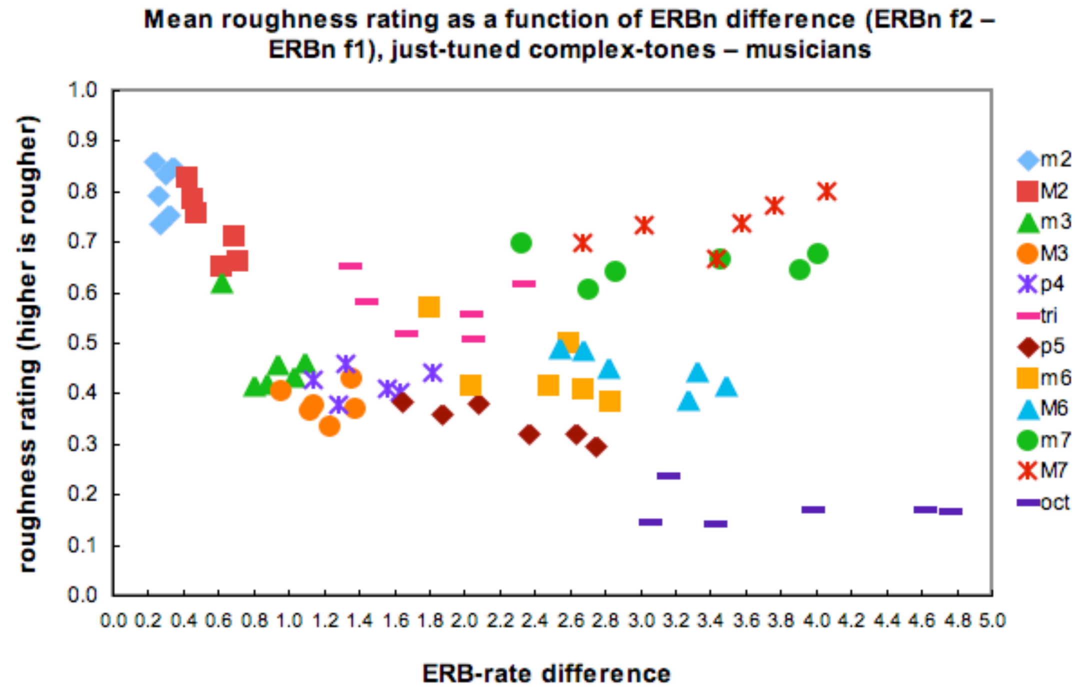

a.
b.
Musicians
Nonmusicians
CT dyads – roughness rating
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
m2 M2 m3 M3 p4 tri p5 m6 M6 m7 M7 oct
Musical Interval

FIG. 4

a.


Mean roughness rating as a function of ERBn difference (ERBn f2 – ERBn f1), just-tuned complex-tones – musicians
roughness rating (higher is rougher)
ERB-rate difference
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0
m2
M2
m3
M3
p4
tri
p5
m6
M6
m7
M7
oct

b.

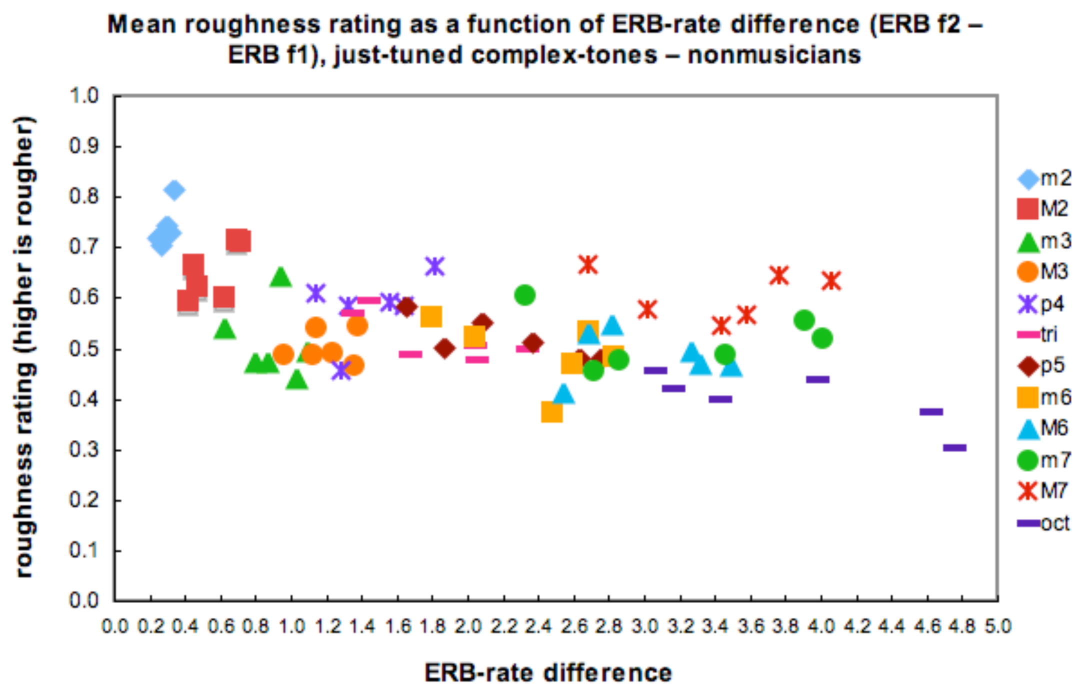
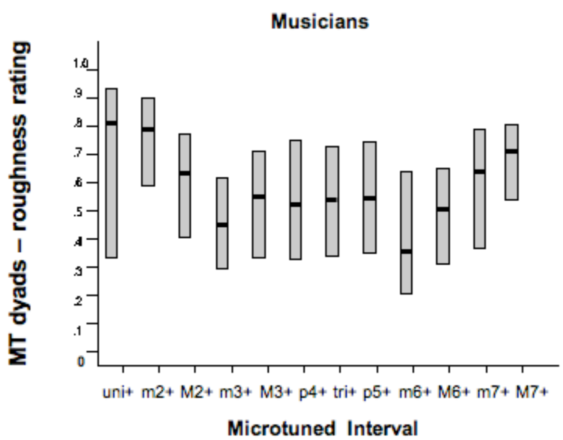

Mean roughness rating as a function of ERB-rate difference (ERB f2 – ERB f1), just-tuned complex-tones – nonmusicians
roughness rating (higher is rougher)
ERB-rate difference
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0
m2
M2
m3
M3
p4
tri
p5
m6
M6
m7
M7
oct

FIG. 5

a.


Musicians
MT dyads – roughness rating
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
uni+ m2+ M2+ m3+ M3+ p4+ tri+ p5+ m6+ M6+ m7+ M7+
Microtuned Interval

b.

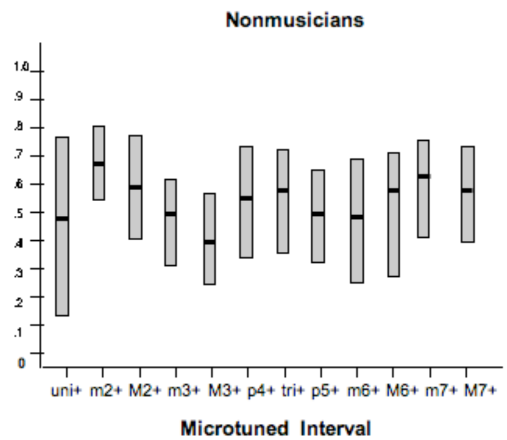
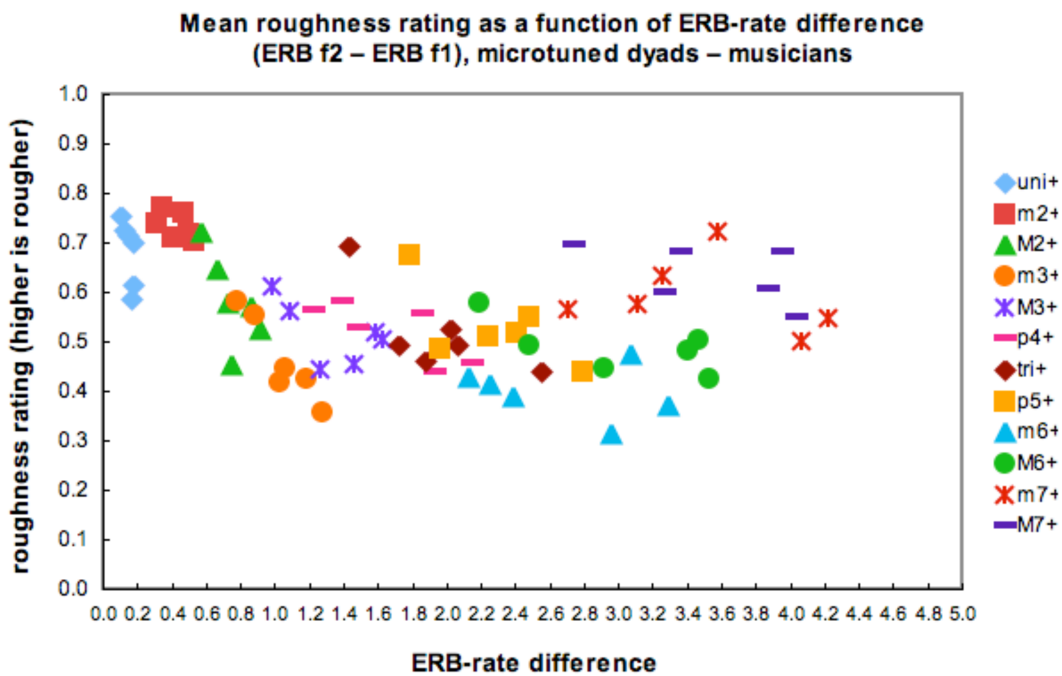

Nonmusicians
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
uni+ m2+ M2+ m3+ M3+ p4+ tri+ p5+ m6+ M6+ m7+ M7+
Microtuned Interval

FIG. 6

a.


Mean roughness rating as a function of ERB-rate difference
(ERB f2 – ERB f1), microtuned dyads – musicians
roughness rating (higher is rougher)
ERB-rate difference
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0
uni+
m2+
M2+
m3+
M3+
p4+
tri+
p5+
m6+
M6+
m7+
M7+

b.

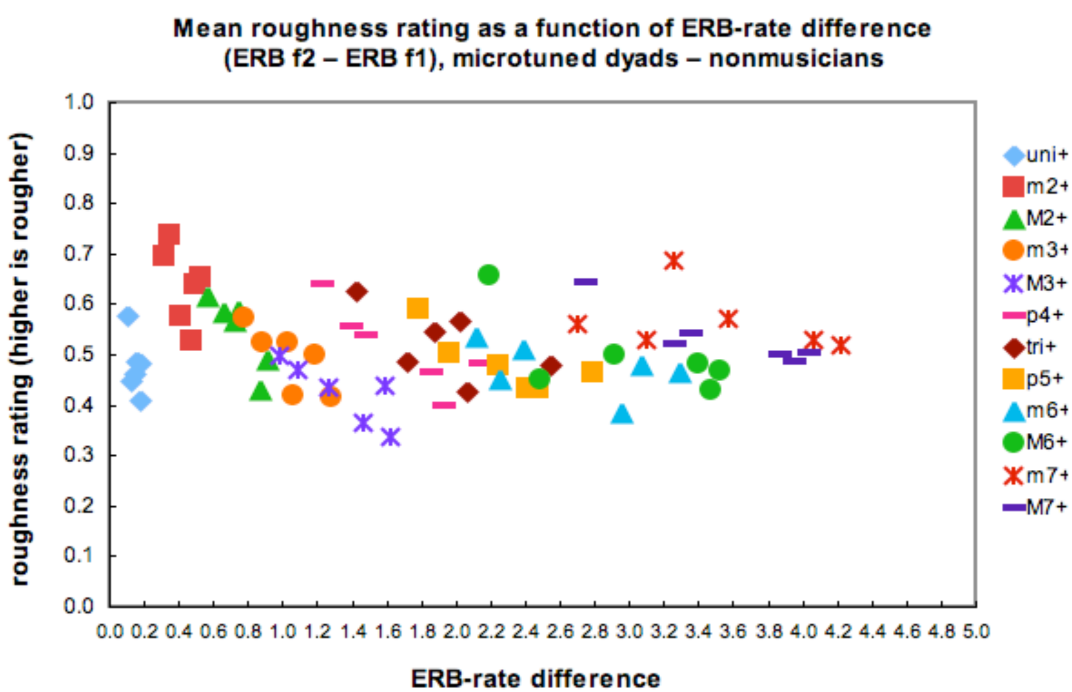
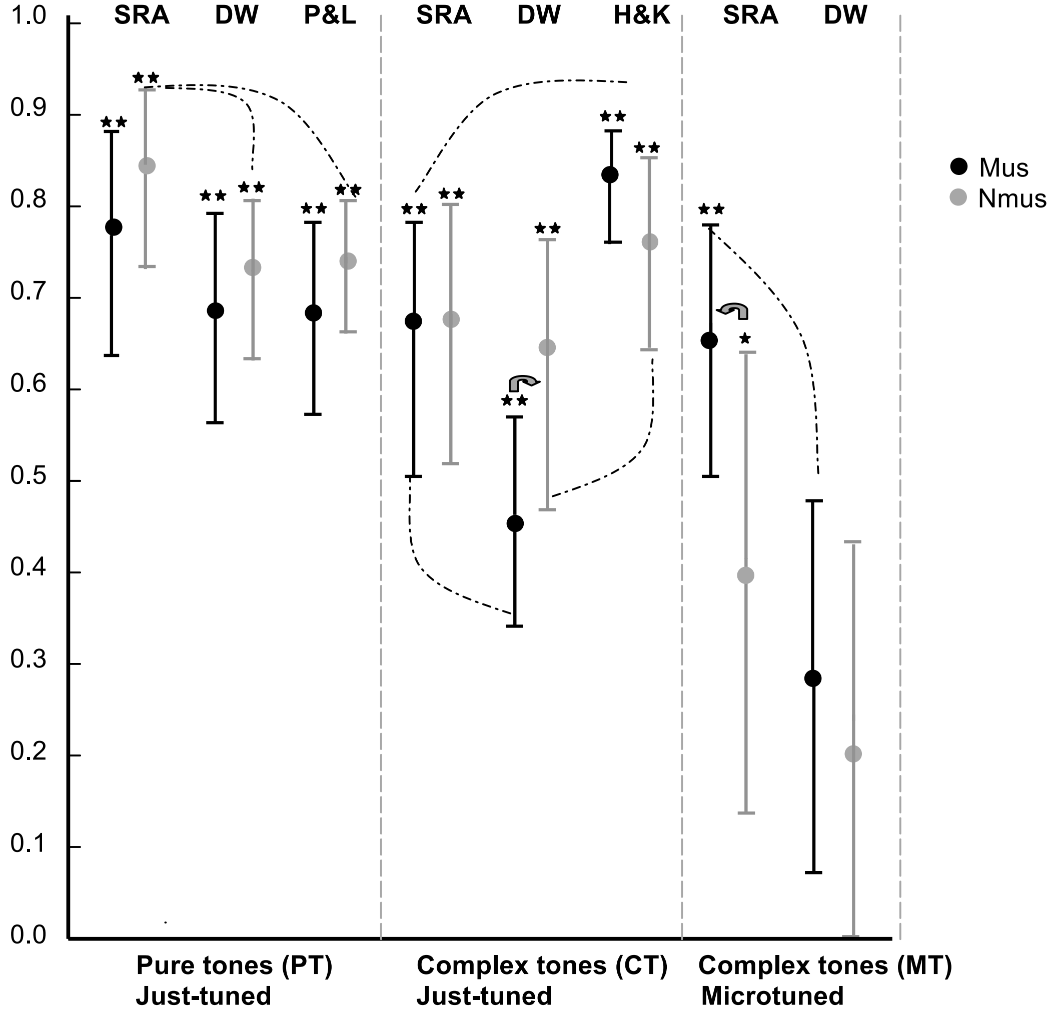

Mean roughness rating as a function of ERB-rate difference
(ERB f2 – ERB f1), microtuned dyads – nonmusicians
roughness rating (higher is rougher)
ERB-rate difference
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6 4.8 5.0
uni+
m2+
M2+
m3+
M3+
p4+
tri+
p5+
m6+
M6+
m7+
M7+

FIG. 7



⋆ Correlation between listener ratings and analyzer/model significant at $p \leq 0.05$
⋆⋆ Correlation between listener ratings and analyzer/model significant at $p \leq 0.001$
Significant difference between mus./nmus. group correlations at $p \leq 0.05$
Significant difference between analyzer/model correlations at $p \leq 0.05$

FIG. 8

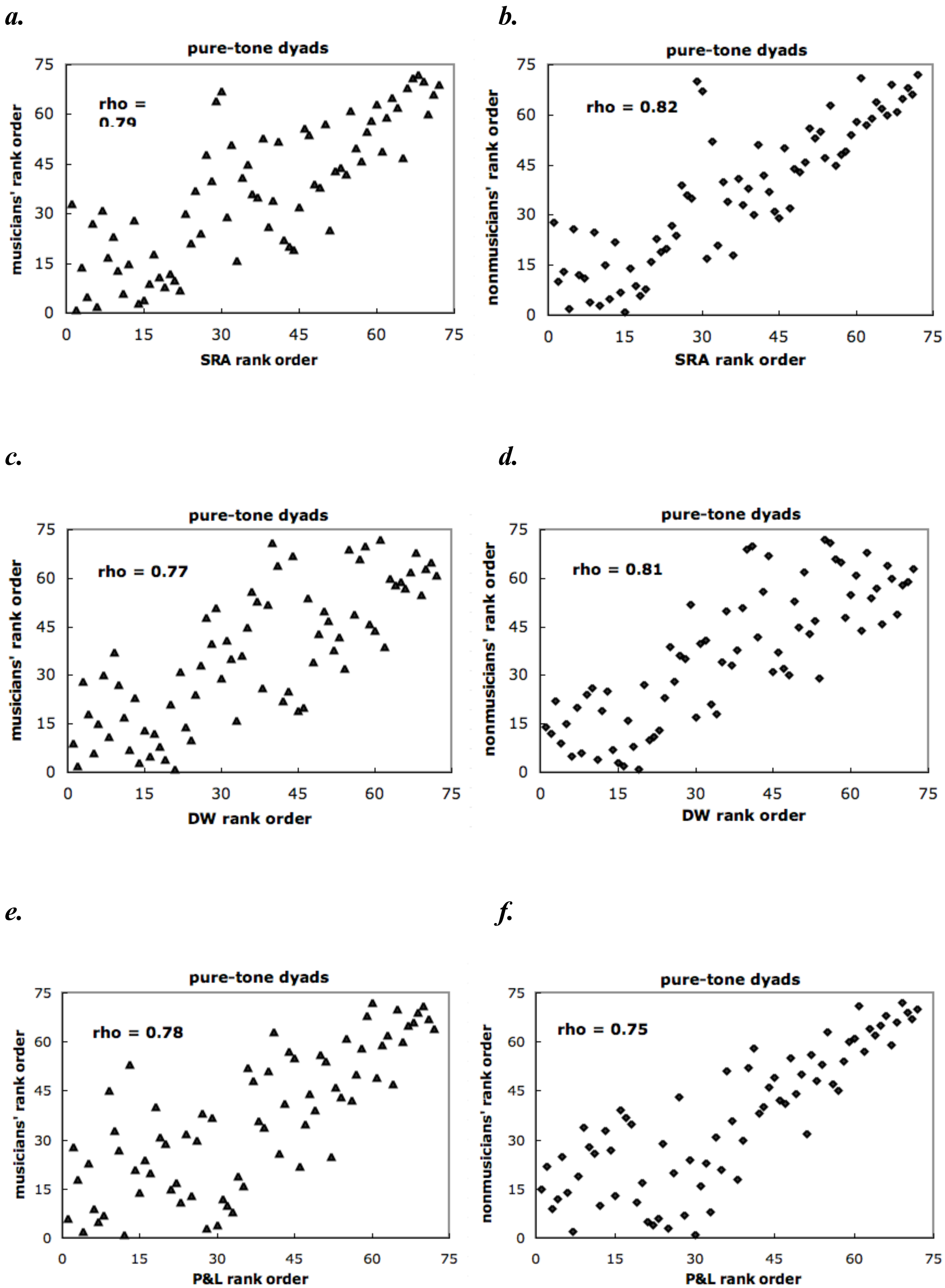

a.
pure-tone dyads
rho =
0.79
musicians' rank order
SRA rank order
b.
pure-tone dyads
rho = 0.82
nonmusicians' rank order
SRA rank order
c.
pure-tone dyads
rho = 0.77
musicians' rank order
DW rank order
d.
pure-tone dyads
rho = 0.81
nonmusicians' rank order
DW rank order
e.
pure-tone dyads
rho = 0.78
musicians' rank order
P&L rank order
f.
pure-tone dyads
rho = 0.75
nonmusicians' rank order
P&L rank order

FIG. 9

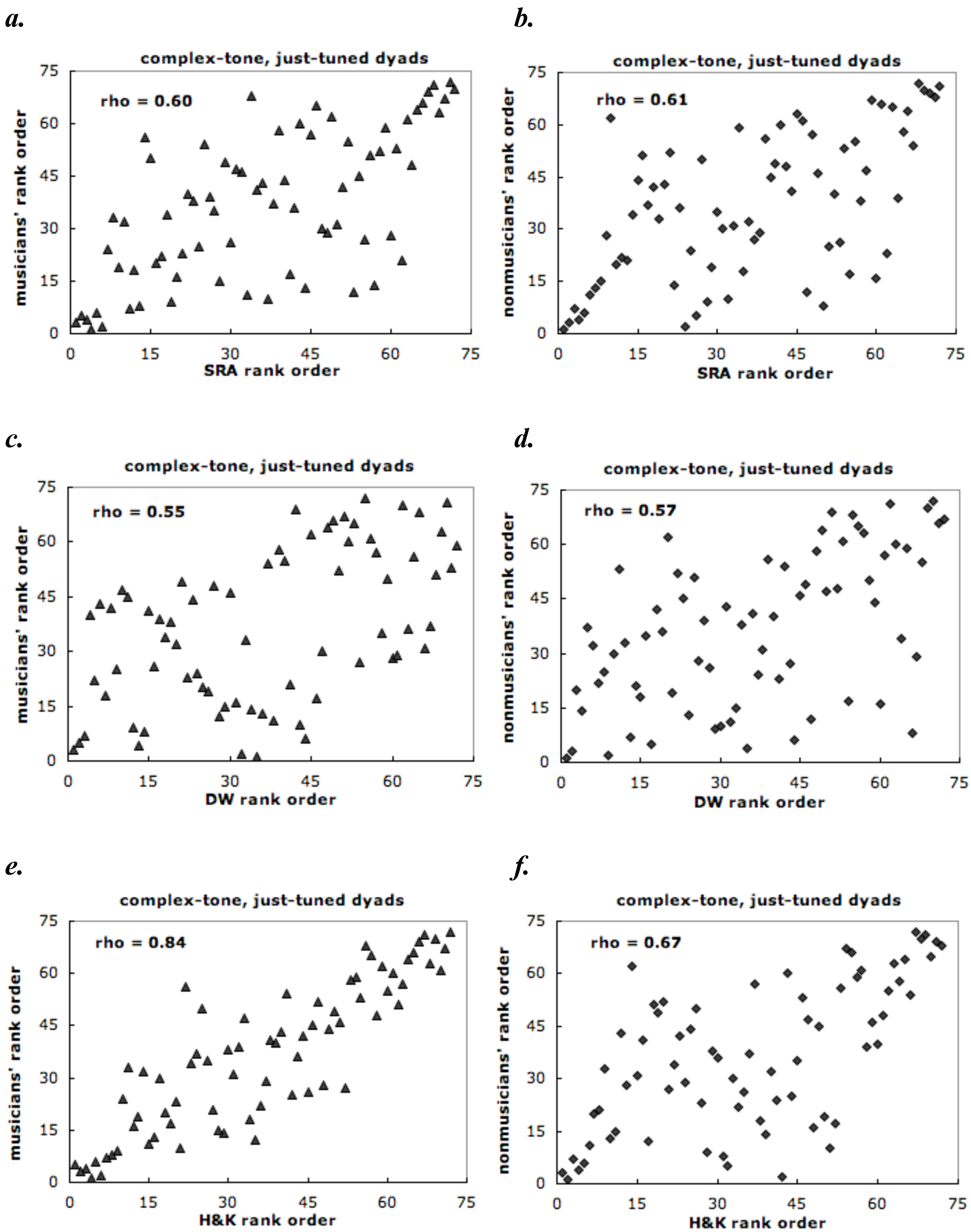
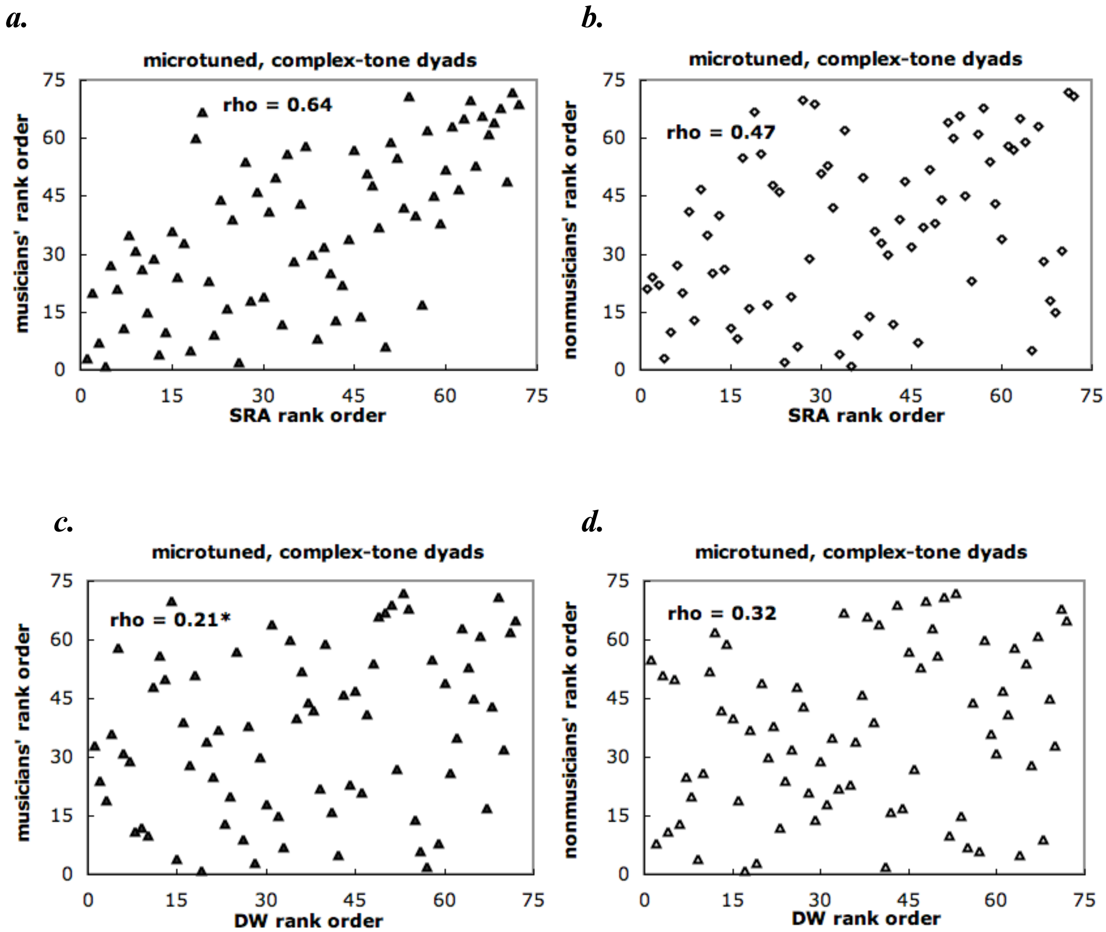

a.
complex-tone, just-tuned dyads
rho = 0.60
musicians' rank order
SRA rank order
b.
complex-tone, just-tuned dyads
rho = 0.61
nonmusicians' rank order
SRA rank order
c.
complex-tone, just-tuned dyads
rho = 0.55
musicians' rank order
DW rank order
d.
complex-tone, just-tuned dyads
rho = 0.57
nonmusicians' rank order
DW rank order
e.
complex-tone, just-tuned dyads
rho = 0.84
musicians' rank order
H&K rank order
f.
complex-tone, just-tuned dyads
rho = 0.67
nonmusicians' rank order
H&K rank order

FIG. 10


a.
microtuned, complex-tone dyads
rho = 0.64
musicians' rank order
SRA rank order
b.
microtuned, complex-tone dyads
rho = 0.47
nonmusicians' rank order
SRA rank order
c.
microtuned, complex-tone dyads
rho = 0.21*
musicians' rank order
DW rank order
d.
microtuned, complex-tone dyads
rho = 0.32
nonmusicians' rank order
DW rank order

Chapter 3: Memory for pure-tone dyads

A sound's degree of consonance or dissonance is mediated by both bottom-up (feature-driven) and top-down (knowledge-driven) processing streams. Chapter 2 investigated the degree to which a musical sound's features may be assessed independently of knowledge of the sound's function in a tonal, musical system. Chapter 3 continues the exploration of the consonance/dissonance distinction from a cognitive perspective. We used a running memory task to learn whether perceptual features or musical knowledge would provide the most reliable cues to accurate auditory short-term memory for musical intervals.

CHAPTER 3

Short-term memory for consonant and dissonant pure-tone dyads

Susan E. Rogers[a)], Daniel J. Levitin, and Stephen McAdams

Author affiliations:
Susan E. Rogers, Daniel J. Levitin
Department of Psychology and Center for Interdisciplinary Research in Music, Media, and Technology, McGill University

Stephen McAdams
Schulich School of Music and Center for Interdisciplinary Research in Music, Media, and Technology, McGill University

[a)] McGill University
1205 Dr. Penfield Avenue, 8th floor
Montreal, Quebec, CA H3A 1B1
Electronic mail: susan.rogers@mail.mcgill.ca

Abstract

In the perception of musical chords, bottom-up processing is not entirely independent of top-down processing; top-down processing involves (implicit or explicit) awareness of how chords function in each listener's musical culture. We explore a cognitive basis for the distinction between the sensory and cognitive properties of musical intervals using a short-term memory paradigm in which pure-tone dyads — two simultaneous tones — were manipulated. Musicians and nonmusicians listened to sequentially presented dyads with a range of sensory and music-theoretic properties. Each dyad was presented twice, separated by a varying number of intervening stimuli; participants judged whether the stimulus was novel or had been presented before. Auditory short-term memory for dyads was accurate well beyond the previously described limit of 30 s for a single pure tone (Winkler et al., 2002), despite interference from incoming sounds and the inability to label the items for long-term storage. Some dyads were recognized with higher accuracy than others, but in no case was there a clear and systematic effect of sensory or cognitive signal attributes on STM as a function of retention interval. Performance dips followed by subsequent improvements were observed for certain categorical subsets. We also found a main effect of musical training: musicians showed better dyad recognition overall and smaller performance differences among interval classes than did nonmusicians. Dyads with salient sensory properties (e.g., Major 2nds, minor 2nds) or presumed cognitive simplicity (e.g., octaves) showed no outstanding advantages over other intervals. This preliminary finding suggests that more than one strategy or mechanism is employed in auditory short-term memory for dyads, mediated by the amount of time afforded for processing. Tracing the time course of differential auditory STM provides information on the distinguishing properties that best serve accurate recognition and suggests optimal periods for feature availability.

keywords: consonance, dissonance, auditory short-term memory, top-down/bottom-up processing, auditory processing

Introduction

When two tones are produced simultaneously, the resulting sound – a *dyad* – may be judged along several continua such as: smooth/rough, pleasant/unpleasant, harmonious/discordant or consonant/dissonant (hereafter abbreviated C/D). A categorical distinction between consonance and dissonance emerges in infancy, presumably as a consequence of perceptual organization and exposure to music and speech (Schwartz, Howe, & Purves, 2003; Terhardt, 1974; Trainor & Heinmiller, 1998; Trainor, Tsang, & Cheung, 2002; Zentner & Kagan, 1998). Human infants display a heightened sensitivity to and preference for consonance over dissonance, leading to speculation that consonant intervals are inherently "natural" and thus easier to process than dissonant ones (Burns & Ward, 1982; Schellenberg & Trehub, 1994a, 1994b; Trainor & Heinmiller, 1998). The C/D distinction may also be regarded as a music-theoretic construct, changing throughout the course of musical history and across musical cultures (Cazden, 1945, 1980). Regardless of whether the C/D distinction is the product of innate or acquired traits, two auditory processing mechanisms determine its perception.

Consonance and dissonance

The terms consonance and dissonance are used to describe horizontal C/D — the harmoniousness of tones in a melodic sequence — as well as vertical C/D — the harmoniousness of simultaneously sounded tones. (The present work focuses on the latter.)

Sensory C/D is described by the psychoacoustical properties of an auditory event. Sensory dissonant signals have one or more "annoying factors" that lead to a relatively discordant, tense, or unpleasant perception (Kameoka & Kuriyagawa, 1969a, 1969b; Terhardt, 1974; Terhardt & Stoll, 1981; Van de Geer, Levelt, & Plomp, 1962). Helmholtz (1885/1954) described the sensory C/D of a pure-tone dyad as a function of the absolute frequency difference between the two tones, noting that unpleasant aural byproducts, and thus dissonance, arise when the tones are narrowly separated. Due to tonotopic (frequency sensitive) mapping on the basilar membrane and auditory nerve, the sensory C/D of a dyad was ultimately shown to relate not only to the frequency separation between two tones, but also to the dyad's mean frequency. Psychoacoustic experiments therefore modeled sensory C/D as a function of relative frequency difference and linked the phenomenon to the critical bandwidth of auditory filtering in the cochlea (Greenwood, 1961, Kameoka & Kuriyagawa, 1969a, 1969b; Plomp & Levelt, 1965, Plomp & Steeneken, 1968). Thus empirical evidence was provided for the centuries-old observation that a given musical dyad sounded more consonant played in the upper register of some instruments than did the same interval played in a lower register (Piston, 1978; Rameau, 1971/1722, pp. 119-123).

Within a single critical bandwidth, two simultaneous tones of equal amplitude interact to produce amplitude modulations termed *beating* when the fundamental frequency difference is less than 15 Hz (Zwicker & Fastl, 1990). At approximately 15-300 Hz frequency difference, corresponding to 10%-50% of the critical bandwidth, depending on the mean frequency of the interval, a turbulent sensation of *roughness* appears (Greenwood, 1961, 1991; Mayer, 1894). Plomp and Levelt (1965, Figure 10) plotted the sensory dissonance of pure-tone dyads as a smooth curve where maximal dissonance was experienced when narrowly spaced intervals were just wider than unison. Minimal dissonance was reported at the frequency separation point corresponding to a marked reduction in perceived roughness. Plomp and Levelt's (1965) model predicted that the average adult listener should judge pure tone dyads with frequency separation wider than a critical band as sensory consonant. In fact, contrary evidence challenges this prediction by suggesting that cultural and experimental differences bias the sensory C/D perception.

Terhardt (1984) redefined sensory C/D with respect to musical sounds and intervals. Sounds were sensory consonant in the absence of two potentially annoying properties: roughness and sharpness (a piercing quality — the weighted loudness of narrow band noise as a function of its spectral center), and the presence of toneness (a quality of periodicity — the opposite of noise). Later research showed that roughness was perhaps the most critical of these due to the effectiveness with which it reduces auditory pleasantness (Blood, Zatorre, Bermudez, & Evans, 1999) and correlates with musical tension (Pressnitzer, McAdams, Winsberg, & Fineberg, 2000).

Sensory C/D is subsumed under what Terhardt (1984) called *musical C/D*, referring to the combined sensory and harmonic or music-theoretic aspects of a musical sound's consonance or dissonance. Here we propose that the more precise and inclusive term *cognitive C/D* be adopted to reflect the phenomenon's conceptual encoding by higher-level cortical processes (Bigand & Tillmann, 2005; Regnault, Bigand, & Besson, 2001). Cognitive C/D is shaped by an individual's life experiences with sounds of all type: music, speech, and environmental noise. Thus the term includes both sensory "pleasantness" as defined by Helmholtz (1885/1954) and Terhardt (1984), and the hierarchical factors of tone-affinity, root-note relationships, and frequency of occurrence and usage in the auditory environment.

The cognitive C/D of a dyad is associated with the size of the integers describing the frequency ratio relationship between the two tones. The integer ratio identifies the number of coincidental harmonics present in a complex-tone dyad. Small-integer ratio dyads such as octaves (1:2) and perfect 5^{th}s (2:3) have a high proportion of coincidental harmonics. Most adult listeners rate small-integer ratio dyads as more pleasant, stable, and consonant than large-integer ratio dyads (Ayres, Aeschbach, & Walker, 1980; Butler & Daston, 1968; Guernsey, 1928; Kameoka & Kuroyagawa, 1969a, 1969b; Malmberg, 1918; Plomp & Levelt, 1965; Plomp & Steeneken, 1968; Van de Geer et al., 1962). Large-integer ratio dyads such as minor 7^{th}s (9:16) and Major 2^{nd}s (8:9) feature a number of harmonics separated by less than half of a critical bandwidth. Narrowly separated harmonics interfere on the basilar membrane in a manner that gives rise to the roughness characteristic associated with large-integer ratio dyads. Although cognitive C/D is based upon principles described by sensory processing, for purposes of the current experiment we will refer to sensory and cognitive C/D as separate notions — the former distinguishing dyads by frequency distance in relation to critical bandwidth and the latter referring to integer ratio complexity.

Early research presumed that the sensory properties of musical events presented in isolation (outside of a musical, tonal context) could be critically assessed without cultural or knowledge bias. A number of studies have asked listeners to rank order the sensory C/D of pure-tone dyads in isolation; lack of agreement in the results indicates that this presumption may be wrong. Table 1 shows that pure-tone dyad assessment by listeners of Western tonal music reflects the traditional music-theoretic relationship between two tones, in many cases despite the absence of sensory dissonance components (Ayers et al*.,* 1980; Guernsey, 1928; Kameoka & Kuriyagawa, 1969a; Malmberg, 1918; Plomp & Levelt, 1965; Van de Geer et al., 1962). Internalized tonal schemas, developed from passive exposure to complex-tone intervals in speech and music, can be unconsciously referenced when pure-tone dyads are evaluated by either expert or nonexpert (musically untrained) listeners (Bigand & Tillmann, 2005; Deutsch, Henthorn, & Dolson, 2004; Itoh et al*.,* 2003; Schwartz et al., 2003; Tramo, Cariani, Delgutte, & Braida, 2003).

Neuropsychological studies with musical intervals show some degree of independence between sensory and cognitive C/D processing, whether intervals are presented in a musical, tonal context (Regnault et al., 2001; Tillmann, Janata, & Bharucha, 2003) or in isolation (Itoh, Suwazono, & Nakada, 2003; Passynkova, Neubauer, & Scheich, 2007; Passynkova, Sander, & Scheich, 2005). An outstanding question is the extent to which this independence is reflected in higher cognitive processes such as short-term memory (STM). The present study aims to elucidate the role of signal properties such as auditory roughness and frequency-ratio relationships in STM for pure-tone dyads — two simultaneously sounded sine waves. A finding of differential memory for dyads would provide clues to the robustness of specific auditory features over time. Two populations — musicians and nonmusicians — are examined separately to acknowledge the growing body of evidence that musical expertise shapes the processing location, speed, and acuity for auditory stimuli (Brattico et al., 2009; Foss, Altschuler, & James, 2007; Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Kreiman, Gerratt, & Berke, 1994; Lee, Skoe, Kraus, & Ashley, 2009; Minati et al., 2009; Passynkova et al., 2007; Regnault et al., 2001; Schlaug, Norton, Overy, & Winner, 2005; Schön, Regnault, Ystad, & Besson, 2005; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005).

Insert Table 1 About Here

Auditory STM

Details of the theoretical distinction between long- and short-term memory systems are the topic of ongoing debate, but a wealth of evidence demonstrates that they are functionally distinct processes. Short-term memory refers to active traces above threshold — meaning that elements of the memory are active, unlike those in long-term storage are thought to be, but are not necessarily the focus of attention (Cowan, 1988; Engle, Tuholski, Laughlin, & Conway, 1999). The fine details in an auditory STM signal are available for mental operations (termed “working memory”), but only for a short time – roughly 30 s as shown in behavioral and physiological experiments (Winkler et al., 2002). Decay over time and interference from new incoming stimuli cause rapid degradation of the active STM trace (Baddeley & Hitch, 1974). The detrimental effect of interference is magnified when incoming stimuli are processed as members of the same category (Dehaene, 1997; Stewart & Brown, 2004). Similarity among musical items in STM presumably depends upon on their familiarity or relevance to individual listeners.

Individual differences appear in STM tasks due to STM's association with long-term memory (LTM). Auditory STM recruits and employs long-term categorical representations of sounds to assist in the formation of a percept, permitting recognition and identification of familiar sounds (Crowder, 1989; Demany & Semal, 2007; Näätänen & Winkler, 1999). Although more robust in terms of resistance to decay, auditory memories retrieved from LTM have poorer feature resolution than do active traces in STM. Memories retrieved from LTM are most useful when the to-be-recognized items are very distinct (Durlach & Braida, 1969).

Top-down control by rehearsing or "looping" an active signal in STM has only a small effect on extending STM duration, unless additional mnemonics are employed (Keller, Cowan, & Saults, 1995). Rehearsing a signal in auditory STM typically involves memory strategies such as attaching a verbal label or visualizing the object, forming new associations, or "chunking" by grouping several events into a single object. “Genuine auditory STM” is said to result when the capacity to rehearse, label, or visualize items can be ruled out by the stimuli or task (Crowder, 1993).

The present study examined how differences between musically expert and nonexpert populations were manifested in the automatic processing of dyads in auditory STM. The dyads used here were from familiar musical categories and easily discriminable by listeners with normal sensitivity to pitch differences (i.e., persons not reporting amusia), therefore no a priori difference in memory capacity per se between populations was assumed. Expert listeners reporting absolute pitch perception — the capacity to immediately identify and label a pitch — were excluded from participation.

Under the present paradigm, both bottom-up and top-down encoding could serve accurate recognition and the results will provide information concerning which processing strategy predominated. Outstanding sensory features of dyads may provide salient recognition cues. If bottom-up encoding based on sensory C/D is used to facilitate dyad recognition, the roughness characteristic of the most sensory dissonant dyads should provide a stronger recognition cue than any found in sensory consonant dyads. Evidence suggests, however, that top-down encoding will exert more influence on dyad recognition than bottom-up encoding (Gaab & Schlaug, 2003; Jacoby, 1983). Should cognitive C/D mediate auditory STM, then large-integer ratio dyads may be recognized more accurately than small-integer ratio ones. This hypothesis is consistent with a recent fMRI study showing increased neural activation for unexpected dissonant targets (dyads) compared with anticipated consonant targets (Tillmann et al., 2003). Increased neural activation for dissonant dyads may promote enhanced memory processing — particularly under heavy cognitive load where additional resources are of value.

Method

Participants

Participants (*N* = 30; 12 men and 18 women; 19-54 years; *M* = 27, *SD* = 8.8) were recruited from a classified ad or volunteered from the Schulich School of Music and Psychology Department at McGill University. Recruits were paid $10 for their time, and volunteers served without pay. All reported normal hearing. Fifteen participants (musician group) had five or more years of formal music training (*M* = 14, *SD* = 7.0); the remaining 15 participants (nonmusician group) had two or fewer years of training (*M* = 0.7, *SD* = 0.9). None had absolute pitch by self-report. Musical training and music listening habits were assessed using a modified version of the Queen's University Music Questionnaire (Cuddy, Balkwill, Peretz, & Holden, 2005).

Apparatus and Stimuli

Sessions were conducted in a soundproof listening environment with participants seated at a Macintosh G5 PowerPC computer with a 30” Apple Cinema Display screen (Apple Computer, Cupertino, CA). Tone sequences were delivered from the computer’s digital audio output, converted to analog by a Mytek Digital Stereo 96 digital-to-analog converter (Mytek Digital, New York, NY), and presented to listeners diotically through AKG K240DF headphones (AKG Acoustics, Nashville, TN).

The software package Signal (Engineering Design, Berkeley, CA) was used to create 72 pure tone dyads by summing in cosine phase the lower frequency sine wave ($f1$) with the higher frequency sine wave ($f2$). Dyads were amplitude normalized so that each stimulus was delivered at a sound pressure level of 57 dB $\pm$ 0.75 dBA SPL at the headphone as measured with a Brüel & Kjær 2203 sound level meter and Type 4153 Artificial Ear headphone coupler (Brüel & Kjær, Naerum, Denmark). Each dyad was 500 ms in duration, including 10-ms raised-cosine onset and offset ramps.

The 72 dyads formed the twelve musical intervals of the Western chromatic scale: minor 2nd (m2), Major 2nd (M2), minor 3rd (m3), Major 3rd (M3), perfect 4th (p4), tritone, perfect 5th (p5), minor 6th (m6), Major 6th (M6), minor 7th (m7), Major 7th (M7), and octave. Frequency ratio relationships corresponded to the just-tuned scale rather than the equal-tempered scale used in contemporary musical performance. Just-tuning was chosen so that frequency ratios would be expressible as a ratio of small whole numbers (Burns & Ward, 1982), in keeping with previous research (Schellenberg & Trehub, 1994, 1996; Trainor, 1997; Tramo et al., 2003; Van de Geer et al., 1962). (The results reported here are assumed to apply to slight mistunings from these ratios, including equal-tempered.) The lower frequencies ($f1$) for each musical interval were assigned by random number table and ranged from C3 (130.8 Hz) to B4 (493.8 Hz). The upper frequencies ($f2$) ranged from C#3 (146.8 Hz) to A5 (880 Hz).

Three unique intervals were created at each of the 24 root notes; the assignment to root notes was random. The pitch of C3 (130.8 Hz), for example, was used to create p5, m6, and p4 and the pitch of C4 (261.6 Hz) was used to create p4, m7, and M2. The reciprocal nature of the design allowed each musical interval to be represented at six different root notes. The m2s, for example, had root notes at C#3, D3, E3, A4, G4, and B4.

Following Schellenberg and Trehub’s (1994b) classification scheme for cognitive C/D, interval complexity was derived by taking the reciprocal of the natural logarithm of the sum of each interval's integers. Next, the 72 dyads were assigned to one of four C/D levels as follows: musically consonant or "MC" (octave, p5, p4), moderately consonant or "mc" (M6, M3, m3), moderately dissonant or "md" (m6, M2, M7), and musically dissonant or "MD" (m7, m2, tritone). The stimulus set was also labeled according to four levels of sensory C/D, based on Plomp & Levelt's (1965, Fig. 9) plot of the subjective sensory consonance of pure-tone dyads as a function of mean frequency and frequency separation. The four levels were labeled as follows: sensory consonant or "SC," somewhat consonant or "sc," somewhat dissonant or "sd," and sensory dissonant or "SD." (See Appendix A for a complete description of each dyad and the derivation of C/D levels.)

Procedure

The 72 dyads were subdivided and presented in three blocks in which participants heard 24 dyads twice – first as a novel item and then later as a familiar item – comprising a grand total of 144 trials (72 × 2). Each block contained a unique subset of dyads: two from each musical interval, with one from the lower octave and one from the higher octave. Thus listeners heard 48 items in a single block [12 intervals × 2 octaves × 2 presentations each (novel/familiar) = 48 trials]. For purposes of balancing the design, item order was carefully controlled within each block as detailed below. The order of presentation of the three blocks however, was randomized for each participant. (See Appendix B for the arrangement of dyads in blocks.)

Stimuli were presented at a fixed rate; time for a single trial was 8.25 s. A *trial* consisted of a 250-ms pre-stimulus alert, a 500-ms stimulus, a 3-s response window, a 500-ms feedback window, and a silent (unfilled) 4-s wait period. Time for each block was 6 min 36 s. Participants were allowed a pause between blocks, but voluntarily completed the experiment in less than 25 min.

Seven possible *retention periods* — the time from when a novel dyad (**S-**) first appeared to when it reappeared as a familiar dyad (**S+**) — served to vary the cognitive load and induce differential memory performance. A given **S-** could have from zero to six intervening stimuli before it reappeared as an **S+**, corresponding to retention periods of 7.75, 16.00, 24.25, 32.50, 40.75, 48.00, or 56.25 s respectively. (Design constraints permitted relatively few of the 56.25-s periods and so data from these periods were not analyzed due to low predictive power.) Novel/familiar pairs with zero intervening stimuli served as catch trials, because performance was expected to be near ceiling under this easy condition.

Consider the following sequence using the letters **W**, **X**, **Y**, and **Z** to represent stimuli, and "**-**" and "**+**" to represent novel and familiar tokens, respectively: **W-, X-**, **Y-**, **Z-**, **Z+**, **W+**, **Y+, X+**. With no intervening stimuli between the novel/familiar pair labeled **Z**, presumably the familiar dyad **Z+** would be easier to recognize than the familiar dyad **X+** with five intervening dyads. Thus correct recognition of dyad **X+** would be performed under a heavier cognitive load than correct recognition of **Z+**. (Although if dyads **Z+** and **X+** were both correctly recognized, dyad **Z+** would have been recognized while the memory trace for **X-** was still active, meaning that **Z+** was recognized under some proactive interference).

To control for repetition bias — the likelihood of predicting whether an upcoming trial would be novel or familiar — novel and familiar trials were carefully distributed within blocks (Todd & Mackintosh, 1990). Because a familiar stimulus must always follow a novel stimulus, the first **S** in a block was always an **S-** and the last was always an **S+**. Within blocks, subsequences from short (- +) to long (- - - - - - + + + + + + +) were arranged to minimize participants' reliance on sequential cues. To best control predictability, the ideal probability of the opposite sign stimulus occurring immediately after a given trial would be 24 out of 47 times or 0.51 (a proportion of 0.50 was impossible, because nothing followed the last stimulus). Achieving sign change ratios close to 0.51 was only possible by using a relatively large number of short sub-sequences (e.g., - +) and a relatively small number of the lengthiest (e.g., - - - - - + + + + + +). Thus, although levels of C/D were distributed across all retention periods, it was not possible to have an equal number of dyad attributes at the longest periods. At moderate retention periods, however, the design was well balanced in terms of the incidence of C/D properties.

Participants were tested individually using an original computer program to adapt the auditory memory paradigm of Cowan, Saults, and Nugent (1997). The experimenter instructed participants that they would hear musical sounds, each presented twice within a block of 48, and that their task was to listen carefully and to use the computer keyboard to match *exact* pairs of sounds from memory. They were told that the familiar presentation would occur no more than one minute or so after a given dyad first appeared. Participants were informed that there would be visual feedback and that the task difficulty was such that they could expect a moderate percentage of wrong answers. They were instructed not to worry about the wrong answers, but to try to focus on memorizing sounds. The experimenter familiarized participants with the task by running a short (eight trial) practice sequence before testing began. Results from the practice session were not analyzed.

The color of the screen changed from yellow to red to alert participants that a trial was about to begin. The question "HAVE YOU HEARD THIS BEFORE?" appeared on the red screen. The screen stayed red as a dyad was played through the headphones, and then turned green for the response phase. Participants had 3 s to respond by pressing a "NO" key (the "-" key on the number keypad) if they believed they were hearing a novel stimulus or a "YES" key (the "+" key on the number keypad) if they thought it was familiar. After the 3-s response window, the screen stayed green while "CORRECT" or "WRONG" appeared for 500 ms. If no answer was entered, the trial was scored as incorrect, and the "WRONG" feedback appeared. The screen turned yellow during a 4-s (silent) inter-stimulus period. After the waiting period, the screen turned red for the next stimulus.

At the end of blocks 1 and 2, a "Please take a rest" message appeared on the screen for 3 s. This was followed by a "Hit any key to continue" message that allowed the participant to initiate the next block when ready. After the third block, a "Finished – Thank you!" message appeared on the screen. Upon exiting the booth, participants were asked to describe any strategies they used to recognize dyads and to give any other impressions of the task.

Results

Data Analysis

Because the data were rich with respect to the number of variables and their possible combinations, a brief description of the analysis methods used is followed by specific results. A variety of statistical tests explored several hypotheses, specifically: comparative memory capacity of musicians and nonmusicians, the effect of retention period on STM for dyads, and the effects of C/D attributes on recognition accuracy.

Analysis methods from Signal Detection Theory (SDT; Macmillan & Creelman, 2005) were used to compare accuracy, confidence, and response biases between participant groups. Stepwise logistic regression and trend analyses addressed whether or not systemic effects of sensory or cognitive C/D were a function of retention period. Hypotheses of differential memory for properties of C/D were tested through separate repeated-measures ANOVAs, as dictated by the variable(s) of interest and their interaction. Where assumptions of sphericity were not met, significance tests were corrected using either Huynh-Feldt estimates of sphericity when epsilon was large ($\varepsilon > .75$) or Greenhouse-Geisser corrections otherwise. (The Greenhouse-Geisser correction tends to underestimate ε when ε is close to 1, but is otherwise an accurate adjustment. The Huynh-Feldt correction tends to overestimate sphericity and is most effective when ε is large; Stevens, 2002, p. 502.) In these instances, the original degrees of freedom (*df*), epsilon and corrected *p* value were reported. Post-hoc tests used Tukey's Honestly Significant Difference (HSD) with a Type I error rate of 0.05 as the criterion for pairwise comparisons. For post hoc tests in cases where the sphericity assumption of the ANOVA was violated, *t*-tests with Bonferroni corrections of $.05/k$ were used for the significance criteria because this correction is more robust to violations of sphericity than Tukey's HSD (Stevens, 2002, p. 509). Because the primary research question addressed the fate of STM traces over time, hypothesis testing of the effect of C/D analyzed only familiar trials.

Unequal distribution of some C/D dyad variables at some retention periods was an unavoidable consequence of the design (see Appendix B), so raw scores for each variable of interest could not be used. Rather, each participant's proportion of correct answers was used as the dependent variable. (Satisfying the demand of equal distribution across all combinations of variables would have required a much larger dataset spanning several octaves to include, for example, sensory consonant m2s at very high pitches and sensory dissonant octaves at very low pitches. A large increase in the number of stimuli would have lengthened the session time considerably, increasing the likelihood of participant fatigue and experimental (t)error.) Each participant's proportion correct was then adjusted for guessing (Macmillan & Creelman, 2005) by the following formula: $p(c)^* = 0.5[P(\text{hits}) + (1 - P(\text{false alarms})]$. Friedman's nonparametric ANOVAs were conducted in addition to the inferential tests (Conover, 1971, p. 265) to account for the possibility that C/D recognition scores were not normally distributed (violating assumptions of ANOVA). Nonparametric post-hoc pairwise comparisons were conducted using the Wilcoxon test. The significance of recognition accuracy under the heaviest cognitive load (48.00 s retention period) was analyzed using *t*-tests that compared the mean $p(c)^*$ score at each C/D classification to chance performance.

Overall performance and the effect of musical training.

Participants displayed a high aptitude for the task, recognizing the novel/familiar status of dyads better than chance. Overall the mean percentage correct was 77.5%; correct responses on familiar trials were significantly higher than on novel trials ($M_{Familiar}$ = 0.80, *SD* = 0.06; M_{Novel} = 0.75, *SD* = 0.08) as confirmed by a paired-samples *t* test, $t(29) = 3.13$, $p < .01$. The proportion of hits and false alarms were calculated for each participant; the proportion of hits was then adjusted for guessing as described above. Corrected hit rates (p(c)*), sensitivity to novel/familiar status (d'), and response biases (c) for each group (musician and nonmusician) are displayed in Table 2. Musicians were more capable than nonmusicians of discriminating between novel and familiar status, as indicated by higher d' values. The criterion value (c) for each class of dyads was calculated to measure participants' decision rules or likelihood of responding either "yes" or "no" (Macmillan & Creelman, 2005, pp. 29-31). Response biases were low except for MC dyads where participants displayed a tendency to respond "no" (i.e., "I have not heard that before"). The values of c ranged from zero or near zero for MD and sc dyads (meaning no bias) to a moderately low value of 0.44 for MC dyads by nonmusicians who were biased towards regarding these items as novel. The distributions of individual d' values were compared in a Mann-Whitney *U* test to evaluate whether the two groups varied in recognition acuity. The result was significant – musicians were more sensitive to recognizing novel/familiar status, and less likely to guess, than nonmusicians, $U = 49.00$, $Z = 2.76$, $p < .01$.

Insert Table 2 About Here.

The effect of retention period.

Retention period was expected to have the largest effect on recognition accuracy. This was confirmed by a stepwise logistic regression analysis using six predictor variables: retention period (7 levels), sensory C/D (4 levels: SC, sc, sd, SD), cognitive C/D (12 levels), root note (24 levels), top note (33 levels), and block order (3 levels: first, second, or last). The dependent variable was the raw score on familiar trials (novel trials were excluded so that retention period could be a factor). Retention period and sensory C/D had the most predictive power on correct scores ($p < .001$). A Hosmer-Lemeshow goodness-of-fit test indicated that a model with only these variables, however, did not improve the intercept-only model, $\chi^2(8, N = 2160) = 22.34$, $p = .004$ and accounted for only 12% of the variance in the data as indicated by Nagelkerke's R^2.

Although recognition accuracy for levels of C/D decreased over time, the patterns were nonsystematic and mostly nonlinear, as shown in Table 3 and Figures 1 – 4. Trend analyses were performed to quantify these changes. Analyses of linear, quadratic, cubic, and 4th order trends were conducted on four levels of cognitive C/D (MC, mc, md, MD) and four levels of sensory C/D (SC, sc, sd, SD) over five retention periods (16.00, 24.25, 32.50, 40.75, and 48. 00 s). (Data from the 7.75 s period were excluded from the analysis because performance was at ceiling and therefore did not contribute anything to the research questions.) Significant trends were seen with some levels of C/D, but no consistent pattern emerged.

Insert Table 3 About Here.

The effect of consonance and dissonance.

The main research question concerned the effect of signal attributes on STM accuracy over time. We grouped dyads by type according to their sensory or cognitive C/D properties and used inferential statistical analysis to explore the significance of these qualitative differences. Two separate three-way repeated-measures mixed-design ANOVAs (one for each C/D property) were initially conducted to evaluate the between-subjects factor of Expertise (2 levels: musician and nonmusician) and two within-subject factors, Retention Period (5 levels: 16.00, 24.25, 32.50, 40.75, and 48.00 s) and C/D (4 levels in each test: MC, mc, md, and MD *or* SC, sc, sd, and SD). The dependent variable was the proportion correct adjusted for guessing (p(c)*) from familiar presentations. The two ANOVAs each revealed a significant main effect of expertise: ANOVA with Cognitive C/D, $F(1, 28) = 6.48$, $p = .02$; ANOVA with Sensory C/D, $F(1, 28) = 4.80$, $p = .04$. The interactions of Expertise with Retention and/or C/D were not significant. Subsequent tests of the effect of C/D segregated participant groups (musician vs. nonmusician) to reduce error (within-cell) variance and increase statistical power (Stevens, 2002, p. 323). The following tests were two-way repeated-measures ANOVAs with Retention (5 levels, 16.00-48.00 s) and C/D (4 levels) as within-subject independent variables and p(c)* as the dependent variable.

Cognitive C/D. Musicians' recognition of dyads decreased with increasing retention period as expected, but the effect of Cognitive C/D attributes on recognition accuracy differed across periods. The effect was most striking for large-integer ratio dyads (MD). Performance accuracy for MD dyads dipped at a moderate retention period but was stronger after longer retention, as shown in Figure 1. The ANOVA revealed a significant Retention × C/D interaction, $F(12, 168) = 1.87$, $p < .05$, and so the simple main effect of Cognitive C/D was explored at each retention period. The only significant effect was observed at 24.25 s, $F(3, 42) = 3.03$, $p = .04$. Pairwise comparisons of the means at this period used Tukey's HSD and revealed a single significant difference: musicians recognized md dyads ($M = 0.88$, $SD = 0.21$) more often than MD dyads ($M = 0.66$, $SD = 0.21$).

Insert Figure 1 About Here.

Nonmusicians' results showed a more pronounced effect of retention period on dyad recognition by Cognitive C/D type, as shown in Figure 2. A dip and improvement for MD dyads was observed, but a second drop in MC dyad recognition appeared at the longest retention periods. The Retention × C/D interaction was stronger than observed in the musician group, $F(12, 168) = 3.14$, $p <.001$. The simple main effects of Cognitive C/D were significant at two retention periods: 24.25 s, $F(3, 42) = 3.00$, $p =.04$, and 40.75 s, $F(3, 42) = 4.89$, $\varepsilon = .55$, $p = .02$. (The Greenhouse-Geisser ε is shown for corrected cases.) Pairwise comparisons of means at 24.25 s revealed that mc dyads ($M = 0.87$, $SD = 0.28$) were recognized significantly more often than MD dyads ($M = 0.60$, $SD = 0.21$), using a Tukey HSD c.v. of 0.251. Mean recognition differences among C/D levels at 40.75 s were compared using a

Bonferroni-adjusted *t*-test and rather than Tukey's HSD due to a violation of sphericity at this retention period (see note regarding test applicability under *Data Analysis*). Differences among levels of Cognitive C/D at 40.75 s just missed significance under a conservative alpha of .008, although one difference approached this criterion. Nonmusicians recognized large-integer ratio (MD) dyads with greater accuracy than small-integer ratio (MC) dyads under heavy cognitive load, $t(14) = 3.01$, $p = .009$.

Insert Figure 2 About Here.

Sensory C/D. Musicians supported the hypothesis that sensory differences among dyads would have a smaller impact than their cognitive differences on recognition accuracy. As predicted, sensory dissonant dyads were recognized with slighter greater accuracy overall than sensory consonant dyads, as shown in Figure 3. Musicians showed significant main effects of Retention, $F(4, 56) = 5.17$, $p = .001$ and Sensory C/D, $F(3, 42) = 8.57$, $p < .001$, and no interaction between the two variables, $F(12, 168) = 0.66$, $p = .78$. Due to the lack of interaction, post hoc tests of Sensory C/D collapsed each musician's mean p(c)* scores across retention periods. Recognition of SD ($M = 0.85$, $SD = 0.03$) and sd ($M = 0.80$, $SD = 0.02$) dyads was significantly more accurate than for SC dyads ($M = 0.67$, $SD = 0.03$) using a Tukey HSD c.v. of 0.089.

Insert Figure 3 About Here.

The sensory distinctions among dyads had a larger influence on nonmusicians' recognition accuracy at specific retention periods than was seen in the musicians' results, as shown in Figure 4. The interaction between Retention, $F(4, 56) = 7.25$, $p < .001$, and Sensory C/D, $F(3,42) = 3.79$, $p = .02$ was marginally significant, $F(12, 168) = 1.69$, $p = .07$, for nonmusicians and so the simple main effects of C/D were examined at all retention periods. Significant performance differences for levels of Sensory C/D appeared at two retention periods: 16.00 s, $F(3, 42) = 4.09$, $p = .01$, and 40.75 s, $F(3, 42) = 5.02$, $\varepsilon = 0.60$, $p = .01$. Post hoc tests of the means at 16.00 s revealed that sc dyads ($M = 0.97$, $SD = 0.02$) were recognized more accurately than both SC ($M = 0.75$, $SD = 0.06$) and sd dyads ($M = 0.79$, $SD = 0.05$), using a Tukey HSD c.v. of 0.176. The differences between nonmusicians' means at 40.75 s were compared using a *t*-test and Bonferroni correction ($\propto = 0.008$) because, as noted, sphericity was not met at this retention period. The most sensory consonant (SC) dyads were recognized significantly less often than sc dyads at 40.75 s, $t(14) = 3.30$, $p = .005$.

Insert Figure 4 About Here.

The inferential statistical tests indicated that STM for dyads is equivalent across most attribute distinctions with exceptions for some classes at specific retention periods. Departures from normal distribution or homogeneity of variance can affect the power of ANOVA tests (Stevens, 2002, pp. 256-267). In cases where homogeneity was not met, adjusting the degrees of freedom using either a Greenhouse-Geisser or Huynh-Feldt correction compensated for these violations and provided a more accurate test. Nevertheless, to ensure better accuracy of these findings we performed Shapiro-Wilk tests of normality using four levels of C/D at the five retention periods of interest. A total of 80 tests of normality were performed — one for each C/D type at the five retention periods of interest — for Sensory and Cognitive C/D by participant group. The result showed that approximately one quarter of all score distributions were significantly skewed from normal under a Bonferroni corrected alpha of $p = .05/20 = .0025$.

A Friedman nonparametric repeated-measures ANOVA examined differences among the C/D classifications and retention periods by participant group. As before, the dependent variable used the average p(c)* for each participant. Five Friedman tests were significant at $p < .05$. The nonparametric results were essentially the same as the parametric ones except for a single more precise distinction at 24.25 s. The Friedman test indicated that musicians recognized SD dyads significantly better than sd and SC dyads at this retention period. The Wilcoxon test was chosen as a post hoc to the significant Friedman tests. The pairwise comparisons showed no significant differences between dyads under the conservative alpha of $p = .05/6 = .008$ but several pairs, all classified by Sensory C/D, came close. The results of the significant Friedman tests and the follow-ups with the largest differences are listed in Table 4.

Insert Table 4 About Here.

Tests of auditory memory duration.

The tests described above compared recognition accuracy among dyads with various properties of consonance and dissonance. An unexpected finding was that recognition accuracy was quite robust for some dyads at the longest retention period —48.00 s. It was of interest to learn whether or not these recognitions were significantly better than chance — the expected level of recognition after so much delay and interference. Paired-sample *t*-tests compared the recognition scores for all classes of dyads against chance (.50) performance; eight tests were conducted for each participant group. The dependent variable was participants' p(c)* scores for each class at 48.00 s. It was not assumed that the eight tests were independent, so a p value of $.05/8 = .006$ was chosen for significance to avoid Type I error. Five tests were significant (see Table 5). Musicians' recognition of the most dissonant (MD, SD) dyads was better than chance at 48.00 s of retention. For nonmusicians under the same conditions, dissonance also aided recognition memory under heavy cognitive load and showed better-than-chance performance for md, MD, and sd dyads. It was noted that although MD dyads were recognized well at 48.00 s, they were recognized significantly worse than other classes at 24.25 s.

Insert Table 5 About Here.

The effect of secondary variables.

The secondary variables of root note, stimulus block, and presentation order had a smaller impact on recognition memory than did retention period and C/D. These variables were examined using one-way ANOVAs. Because cell sizes for these variables were equally distributed across the design, the unadjusted raw scores (correct/incorrect) from both novel and familiar trials ($N = 30$) were used as the dependent variables in these three tests.

As described in the Methods section and the Appendices, each of the 24 root notes was presented to listeners a total of six times (in three novel and in three familiar trials), thus a score of six was the maximum possible at each note. There was a slight tendency for dyads with higher root notes to be recognized more often than dyads with lower root notes. Mean correct scores ranged from 5.3 for note A4 (SD = 0.66) to 4.1 for note G3 (SD = 0.94), as shown in Figure 5. A one-way repeated-measures ANOVA was performed with root note (24 levels) as the independent variable. The omnibus ANOVA for root pitch was significant, $F(23,667) = 3.96$, $p < .001$. Post hoc tests using a Tukey HSD c.v. of 0.862 indicated that 20 of the 276 comparisons were significantly different at $p \leq .05$. Dyads with root notes A4, A#4, and D#4 were more likely to be recognized than dyads with root notes F#3, C#3, D#3, and G3. In addition, A4 and A#4 root notes were recognized more often than root notes A#3, E3, and F4. Root note A4 was frequently recognized and performed significantly better than F#4 in addition to the above-mentioned notes. Root note C3 was recognized more often than root note G3.

Of the three stimulus blocks, Block 2 was expected to report slightly lower average scores because it contained four instances of six intervening dyads between the novel/familiar presentations, corresponding to 56.25 s of retention (not analyzed). Block differences were examined in a one-way repeated-measures ANOVA using each participant's correct score (out of a potential maximum of 48) as the dependent variable. A significant main effect of stimulus block was found, $F(2, 58) = 4.75$, $p = .012$. Post hoc tests used a Tukey HSD c.v. of 2.02 for significance at $p \leq .05$. Scores in Block 1 were significantly higher than in Block 2 ($M_{Block\ 1} = 38.6$, $SD = 3.34$; $M_{Block\ 2} = 36.4$, $SD = 4.05$; $M_{Block\ 3} = 36.8$, $SD = 3.29$).

The order of presentation (first, second, last) of the stimulus blocks was randomized for each participant and showed no primacy, practice, or other effect of order. A one-way repeated-measures ANOVA used raw scores (out of 48) from each participant's first, second, and last block as the dependent variable. Presentation order did not affect response scores, $F(2,58) = 0.40$, $p = .674$, ($M_{First} = 37.3$, $SD = 3.40$; $M_{Second} = 37.7$, $SD = 3.40$; $M_{Last} = 36.9$, $SD = 4.23$).

Discussion

The influence of sensory (as indexed by roughness) and cognitive (as indexed by frequency-ratio complexity) properties distinguishing consonance from dissonance in STM is nonsystematic or at least too subtle to reveal with the current methodology. Recognition accuracy in general decreases with time, but the mediating effects of retention period and the number of intervening items on memory processing are not linearly related to C/D distinctions. Performance dips followed by subsequent improvement after lengthier retention periods were observed for certain dyad classes. These patterns could have been artifacts of the design, including the similarity of any given dyad to the dyads immediately preceding or following its presentation (Stewart & Brown, 2004, 2005). Regardless of the cause of performance dips, accurate recognition was observed for most classes of dyads beyond 30-s retention — the presumed limit of auditory STM for single pitches (Winkler et al., 2002) — despite the intervening effects of new incoming dyads.

Cognitively consonant, small-integer ratio dyads were not recognized more accurately than dyads with more complex frequency-ratio relationships. In fact, they were the least well recognized under the heaviest cognitive load. This finding is at odds with the notion of a processing advantage for so-called natural intervals (Schellenberg & Trehub, 1996), presuming that such an advantage assists STM. However, harmonic partials that contribute to C/D were absent in the pure-tone dyads used here. The cognitive C/D distinction among pure-tone dyads is, practically speaking, subtler than for complex-tone dyads. Nevertheless, differential STM performance by degree of cognitive C/D was observed at some retention periods.

Memory for dyads in the most complex integer-ratio relationships was more affected by retention period than for those in simpler relationships. Recognition accuracy was poor at a moderate retention period but better at a long retention period for the most cognitively dissonant dyads. The hypothesis that dissonance would assist STM was therefore conditionally supported. If, however, high recognition accuracy was due to additional neural resources recruited to process dissonance (Bigand & Tillmann, 2005), these resources were selective. Although *cognitive* dissonance was poorly recognized at 24.25 s, *sensory* dissonance was recognized more accurately than other attributes at the same period. A follow-up experiment (Chap. 4) with complex-tone dyads under the same paradigm will allow a more in-depth examination of the effect of cognitive C/D on recognition STM and will answer the question of whether or not the dip at 24.25 s was merely an artifact of this particular arrangement of dyads.

Sensory C/D distinctions were more consistent predictors of dyad recognition across retention periods than were cognitive C/D distinctions. The roughness characteristic of SD dyads was expected to provide a strong perceptual cue that could be used for recognition (Goldstone & Barsalou, 1998) but musicians were more sensitive to this cue than were nonmusicians (see Table 2). Nonmusicians displayed only marginally higher sensitivity to rougher compared to smoother dyads, suggesting that untrained listeners did not employ bottom-up encoding for processing musical intervals. On the surface this finding is counterintuitive, because it suggests that nonmusicians were more inclined than musicians to rely on knowledge-based encoding. Formal musical training, however, increases the individual's capacity to process a musical sound analytically by its component parts, and this may have been engaged automatically (Seither-Preisler et al., 2007; Tervaniemi et al., 2005). A comparison of Figs. 3 and 4 shows that nonmusicians were less consistent than musicians in recognizing dyads by their sensory attributes.

The current experimental protocol afforded listeners a single opportunity to establish a memory trace for each dyad, and according to multiple-trace memory theory (Hintzman, 1986; Goldinger, 1998), this is done automatically. The lack of repetition or a tonal melodic pattern discouraged the establishment of long-term contextual memories to assist dyad recognition, although it is not assumed that LTM was not a factor.

Memory traces for target pitches can be reactivated after 30-s silent intervals (Winkler et al., 2002). A familiar target pitch presented after 30-s retention was presumed to either reactivate a dormant memory trace in a temporary buffer (e.g., STM) or reinstate a contextual memory (assumed to have been) established in LTM through repeated presentations of a standard pitch. Memory for the standard pitch, however, was only robust enough to permit accurate discrimination from the target pitch when the frequency difference between the two was sufficiently large. If accurate recognition memory in the current experiment is due to reactivated traces in STM rather than LTM, at least some auditory features persist in STM beyond 30-s retention, despite interference from incoming items. What remains is often sufficient for accurate recognition.

Familiar trials showed higher recognition accuracy than novel trials. This result is inconsistent with other behavioral studies showing recognition memory to be better for novel than for familiar information (Habib, McIntosh, Wheeler, & Tulving, 2003; Kormi-Nouri, Nilsson, & Ohta, 2000; Tulving & Kroll, 1995). The result is consistent, however, if practical familiarity with musical intervals is considered. The likelihood of an item being encoded for LTM storage is positively correlated with its degree of novelty (Kormi-Nouri et al., 2000; Tulving, Markowitsch, Kapur, Habib, & Houle, 1994). Where the present result conformed to other studies (i.e., novel > familiar items) was with dyads classified as MC. The response criteria *c* presented in Table 2 show that participants — especially nonmusicians — were paradoxically inclined to respond "novel" to common dyads with the most simple integer ratio relationships (octaves, p5s, p4s). This tendency caused a high proportion of misses on familiar MC trials, especially at long retention periods (MC recognition dropped sharply after 32.50 s; see Figs. 1 and 2). The relative commonness of MC dyads and participants' over-familiarity with them may have contributed detrimentally to their encoding during novel trials, causing poor recognition of late familiar presentations. Dyads with a higher degree of novelty outside of the experimental context such as tritones and m2s might have benefited from enhanced encoding during their novel trials. When the novelty of items outside of the experimental context is considered, the difference between novel and familiar trials can be reassessed.

The hypothesis that musicians would recognize dyads more accurately overall than nonmusicians was supported. It is assumed that the differences were not due to enhanced STM processing per se, but to musicians' heightened sensitivity to differences between chords (Tervaniemi et al., 2005) revealing itself in dyads (which are in fact, 2/3rds of a standard triad chord). This was supported by the higher d' values for musicians compared to nonmusicians. Given that d' reflects the magnitude of the participant's sensitivity and thus decision ability (Pastore, Crawley, Berens, & Skelly, 2003), this suggests that a higher capacity for processing auditory stimuli in general contributed to dyad recognition. It is tempting to conclude that because nonmusicians recognized dyads nearly as well as musicians did, the underlying processing was also similar. Yet tests of recognition memory tap a wide range of confidence in the responses (Yonelinas, 2001). The difference between the grand means of d' values (see Table 2), rather than proportions correct, should be considered an accurate marker of the effect size of musical expertise on STM for dyads.

Conclusion

Peretz and Zatorre (2005, pg. 94) urged researchers to explore the stages of auditory processing in which sensory C/D is transduced into cognitive C/D in order to understand the differences between built-in auditory constraints and learned auditory associations. The current work aimed to uncover C/D distinctions in an automatic, nonmusical, cognitive task. The lack of a systematic relationship between C/D attributes and memory integrity suggests that auditory features do not decay uniformly over time. The finding of robust auditory STM beyond 30 s of retention for some types of sounds can provide clues as to which auditory features serve accurate STM recognition and the time course of feature availability. The present finding thus contributes to understanding stages of auditory cognitive processing, including the organization of musical intervals and the nature of the distinction between consonance and dissonance.

Acknowledgements

The work reported herein was submitted in partial fulfillment of the requirements for the Ph.D. in Psychology at McGill University by the first author. We gratefully acknowledge the technical assistance of Bennett Smith and Karle-Philip Zamor, the support of grants from NSERC and Bell Canada to DJL, grants from NSERC and the Canada Research Chair program to SMc, and a CIRMMT Student Award to SER. Evan Balaban and Al Bregman provided valuable comments on earlier drafts of this paper. Richard Pastore offered helpful insights on applying detection theory methods to the analysis. Portions of this work were presented at conventions of the Society for Music Perception and Cognition, Montréal, QC, August, 2007; the Audio Engineering Society, New York, NY, October, 2007; and the Canadian Acoustic Association, Montréal, QC, August, 2007.

Appendix A

Deriving sensory consonance/dissonance classifications

The consonance ratings of pure-tone dyads as collected by Plomp and Levelt (1965, Fig. 9) were used to assign four levels of sensory C/D to the present stimulus set. First, the mean frequency of each dyad was calculated (Table A1, col. 7, "**mean freq**."). Next, numerical values in Hertz corresponding to the "frequency difference of (the) smallest consonant interval of two simple tones as a function of the mean frequency of the tones" were derived from Plomp and Levelt's Fig. 9, interpolating when necessary (Table A1, col. 9, "**P&L criterion**"). Depending on a dyad's mean frequency, the appropriate "P&L criterion" was subtracted from the dyad's frequency difference in Hertz (Table A1, col. 8, "$\mathbf{f_2 - f_1}$"), generating either a positive or negative number (Table A1, col. 10, "**C/D?**"). In this way each dyad was labeled as falling either outside or inside of Plomp and Levelt's "dissonant band."

Because sensory C/D is based on frequency relationships *relative* to critical bandwidth, it was necessary to label each dyad by how far it exceeded or receded into the "dissonance band." This was accomplished by dividing $f_2 - f_1$ by the P&L criterion to yield a ratio (Table A1, col. 11, "**relative C/D**"). These ratios were ranked by size, indicating the degree of a dyad's sensory consonance or dissonance. Dyads within the P&L dissonance band were labeled sensory dissonant or 'SD.' Dyads slightly outside of the dissonance band but within a critical bandwidth as defined by Zwicker and Fastl (1990) were labeled "somewhat dissonant" (sd). The remaining dyads were divided such that the widest were labeled sensory consonant (SC) and the rest were labeled "somewhat consonant" (sc) – wider than a critical band but by not as much as dyads in the SC category. The categorization scheme is shown in Table A2.

Table A1

Stimulus blocks (see Appendix A for a description of each column).

Block 1.

dyad	intvl.	lower note	upper note	f1 Hz	f2 Hz	mean freq.	f2–f1 Hz	P&L criterion	C/D?	relative C/D	cog. class	sen. class	freq. ratio
A	p5	C3	G3	130.8	196.0	163.4	65.2	60	5.2	1.09	sd	MC	2:3
B	M2	C#3	D#3	138.6	155.6	147.1	17.0	70	-53.0	0.24	SD	md	8:9
C	M3	D3	F#3	146.8	185.0	165.9	38.2	60	-21.8	0.64	SD	mc	4:5
D	tri	D#3	A3	155.6	220.0	187.8	64.4	60	4.4	1.07	sd	MD	32:45
E	m2	E3	F3	164.8	174.6	169.7	9.8	60	-50.2	0.16	SD	MD	15:16
F	p4	F3	A#3	174.6	233.1	203.9	58.5	60	-1.5	0.98	SD	MC	3:4
G	m6	F#3	D4	185.0	293.6	239.3	108.6	50	58.6	2.17	sd	md	5:8
H	m7	G3	F4	196.0	349.2	272.6	153.2	40	113.2	3.83	sc	MD	9:16
I	m3	G#3	B3	207.6	246.9	227.3	39.3	60	-20.7	0.66	SD	mc	5:6
J	oct	A3	A4	220.0	440.0	330.0	220.0	20	200.0	11.00	SC	MC	1:2
K	M6	A#3	G4	233.1	392.0	312.6	158.9	30	128.9	5.30	SC	mc	3:5
L	M7	B3	A#4	246.9	466.2	356.6	219.3	20	199.3	10.97	SC	md	8:15
M	p4	C4	F4	261.6	349.2	305.4	87.6	30	57.6	2.92	sc	MC	3:4
N	oct	C#4	C#5	277.2	554.4	415.8	277.2	30	247.2	9.24	SC	MC	1:2
O	tri	D4	G#4	293.6	415.2	354.4	121.6	20	101.6	6.08	SC	MD	32:45
P	m6	D#4	B4	311.2	493.8	402.5	182.6	30	152.6	6.09	SC	md	5:8
Q	M2	E4	F#4	329.6	370.0	349.8	40.4	20	20.4	2.02	sd	md	8:9
R	M6	F4	D5	349.2	587.3	468.3	238.1	50	188.1	4.76	sc	mc	3:5
S	m7	F#4	E5	370.0	659.2	514.6	289.2	60	229.2	4.82	sc	MD	9:16
T	m3	G4	A#4	392.0	466.2	429.1	74.2	40	34.2	1.86	sd	mc	5:6
U	M3	G#4	C5	415.2	523.5	469.4	108.3	50	58.3	2.17	sd	mc	4:5
V	m2	A4	A#4	440.0	466.2	453.1	26.2	40	-13.8	0.66	SD	MD	15:16
W	M7	A#4	A5	466.2	880.0	673.1	413.8	70	343.8	5.91	SC	md	8:15
X	p5	B4	F#5	493.8	740.0	616.9	246.2	70	176.2	3.52	sc	MC	2:3

Block 2.

dyad	intvl.	lower note	upper note	f1 Hz	f2 Hz	mean freq.	f2–f1 Hz	P&L criterion	C/D?	relative C/D	cog. class	sen. class	freq. ratio
a	m6	C3	G#3	130.8	207.6	169.2	76.8	60	16.8	1.28	sd	md	5:8
b	M7	C#3	C4	138.6	261.6	200.1	123.0	60	63.0	2.05	sd	md	8:15
c	m2	D3	D#3	146.8	155.6	151.2	8.8	60	-51.2	0.15	SD	MD	15:16
d	M6	D#3	C4	155.6	261.6	208.6	106.0	60	46.0	1.77	sd	mc	3:5
e	p4	E3	A3	164.8	220.0	192.4	55.2	60	-4.8	0.92	SD	MC	3:4
f	M2	F3	G3	174.6	196.0	185.3	21.4	60	-38.6	0.36	SD	md	8:9
g	p5	F#3	C#4	185.0	277.2	231.1	92.2	50	42.2	1.84	sd	MC	2:3
h	oct	G3	G4	196.0	392.0	294.0	196.0	30	166.0	6.53	SC	MC	1:2
i	m7	G#3	F#4	207.6	370.0	288.8	162.4	40	122.4	4.06	sc	MD	9:16
j	m3	A3	C4	220.0	261.6	240.8	41.6	50	-8.4	0.83	SD	mc	5:6
k	tri	A#3	E4	233.1	329.6	281.4	96.5	40	56.5	2.41	sc	MD	32:45
l	M3	B3	D#4	246.9	311.2	279.1	64.3	40	24.3	1.61	sd	mc	4:5
m	m7	C4	A#4	261.6	466.2	363.9	204.6	20	184.6	10.23	SC	MD	9:1
n	tri	C#4	G4	277.2	392.0	334.6	114.8	20	94.8	5.74	SC	MD	32:45
o	p4	D4	G4	293.6	392.0	342.8	98.4	20	78.4	4.92	SC	MC	3:4
p	M3	D#4	G4	311.2	392.0	351.6	80.8	20	60.8	4.04	sc	mc	4:5
q	M7	E4	D#5	329.6	622.3	476.0	292.7	50	242.7	5.85	SC	md	8:15
r	oct	F4	F5	349.2	698.4	523.8	349.2	60	289.2	5.82	SC	MC	1:2
s	m6	F#4	D5	370.0	587.3	478.7	217.3	50	167.3	4.35	sc	md	5:8
t	m2	G4	G#4	392.0	415.2	403.6	23.2	30	-6.8	0.77	SD	MD	15:16
u	M6	G#4	F5	415.2	698.4	556.8	283.2	60	223.2	4.72	sc	mc	3:5
v	p5	A4	E5	440.0	659.2	549.6	219.2	70	149.2	3.13	sc	MC	2:3
w	m3	A#4	C#5	466.2	554.4	510.3	88.2	60	28.2	1.47	sd	mc	5:6
x	M2	B4	C#5	493.8	554.4	524.1	60.6	60	0.6	1.01	sd	md	8:9

Block 3.

dyad	intvl.	lower note	upper note	f1 Hz	f2 Hz	mean freq.	f2–f1 Hz	P&L criterion	C/D?	relative C/D	cog. class	sen. class	freq. ratio
aa	p4	C3	F3	130.8	174.6	152.7	43.8	60	-16.2	0.73	SD	MC	3:4
ab	m2	C#3	D3	138.6	146.8	142.7	8.2	70	-61.8	0.12	SD	MD	15:16
ac	m7	D3	C4	146.8	261.6	204.2	114.8	60	54.8	1.91	sd	MD	9:16
ad	M7	D#3	D4	155.6	293.6	224.6	138.0	60	78.0	2.30	sc	md	8:15
ae	oct	E3	E4	164.8	329.6	247.2	164.8	50	114.8	3.30	sc	MC	1:2
af	m3	F3	G#3	174.6	207.6	191.1	33.0	60	-27.0	0.55	SD	mc	5:6
ag	tri	F#3	C4	185.0	261.6	223.3	76.6	60	16.6	1.28	sd	MD	32:45
ah	p5	G3	D4	196.0	293.6	244.8	97.6	50	47.6	1.95	sc	MC	2:3
ai	M3	G#3	C4	207.6	261.6	234.6	54.0	50	4.0	1.08	sd	mc	4:5
aj	M6	A3	F#4	220.0	370.0	295.0	150.0	30	120.0	5.00	SC	mc	3:5
ak	M2	A#3	C4	233.1	261.6	247.4	28.5	50	-21.5	0.57	SD	md	8:9
al	m6	B3	G4	246.9	392.0	319.5	145.1	30	115.1	4.84	SC	md	5:8
am	M2	C4	D4	261.6	293.6	277.6	32.0	40	-8.0	0.80	SD	md	8:9
an	M7	C#4	C5	277.2	523.5	400.4	246.3	30	216.3	8.21	SC	md	8:15
ao	m3	D4	F4	293.6	349.2	321.4	55.6	30	25.6	1.85	sd	mc	5:6
ap	oct	D#4	D#5	311.2	622.3	466.8	311.1	50	261.1	6.22	SC	MC	1:2
aq	M3	E4	G#4	329.6	415.2	372.4	85.6	20	65.6	4.28	sc	mc	4:5
ar	tri	F4	B4	349.2	493.8	421.5	144.6	30	114.6	4.82	sc	MD	32:45
as	M6	F#4	D#5	370.0	622.3	496.2	252.3	60	192.3	4.21	sc	mc	3:5
at	m6	G4	D#5	392.0	622.3	507.2	230.3	60	170.3	3.84	sc	md	5:8
au	p5	G#4	D#5	415.2	622.3	518.8	207.1	60	147.1	3.45	sc	MC	2:3
av	p4	A4	D5	440.0	587.3	513.7	147.3	60	87.3	2.46	sc	MC	3:4
aw	m7	A#4	G#5	466.2	830.6	648.4	364.4	70	294.4	5.21	SC	MD	9:16
ax	m2	B4	C5	493.8	523.5	508.7	29.7	60	-30.3	0.50	SD	MD	15:16

Table A2
Categorization scheme for sensory consonance/dissonance.

Category	Relative C/D (Table A.1. col. 11)	Number of dyads
SD	≤ 1.00	17
sd	$0.99 < sd \leq 2.18$	18
sc	$2.17 < sc \leq 4.83$	19
SC	> 4.82	18

Deriving cognitive consonance/dissonance levels

Schellenberg and Trehub (1994b) quantified the simplicity of frequency ratios by taking the reciprocal of the natural logarithm of the sum of a dyad's integers (X and Y): $[\ln(X + Y)]^{-1}$. The frequency ratios correspond to the *just tuned* scale, where instruments are tuned so that all notes of the scale are related by whole integers. This index was used to assign four levels of cognitive consonance/dissonance to dyads in the present experiment. Table A3 shows Schellenberg and Trehub's C/D ordering by integer-ratio complexity.

Table A3
Quantitative ordering in terms of cognitive consonance/dissonance.

Musical interval	Integer ratio (X:Y)	$\ln(X+Y)^{-1}$
octave	1:2	0.910
p5	2:3	0.621
p4	3:4	0.514
M6	3:5	0.481
M3	4:5	0.455
m3	5:6	0.417
m6	5:8	0.390
M2	8:9	0.353
M7	8:15	0.319
m7	9:16	0.311
m2	15:16	0.291
tritone	32:45	0.230

Appendix B

Table B1

The arrangement of dyads in three blocks shows trial number, dyad name, novelty status, and the number of intervening dyads between each novel/familiar pair. The symbol "–" stands for a novel presentation and "+" stands for familiar.

Block 1.

Trial #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
dyad	N	F	X	P	R	N	P	F	X	R	M	M	T	V	I	T	I	V	L	C	J	J	L	C
nov./fam.	–	–	–	–	–	+	+	+	+	+	–	+	–	–	–	+	+	+	–	–	–	+	+	+
# intvn.	*	*	*	*	*	4	2	5	5	4	*	0	*	*	*	2	1	3	*	*	*	0	3	3
Trial #	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
dyad	H	H	B	U	B	U	K	K	E	S	O	G	Q	O	E	S	G	Q	D	A	W	W	D	A
nov./fam.	–	+	–	–	+	+	–	+	–	–	–	–	–	+	+	+	+	+	–	–	–	+	+	+
# intvn.	*	0	*	*	1	1	*	0	*	*	*	*	*	2	5	5	4	4	*	*	*	0	3	3

Block 2.

Trial #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
dyad	e	m	v	v	m	r	e	n	u	n	l	r	u	l	o	o	l	w	h	a	q	b	q	i
nov./fam.	–	–	–	+	+	–	+	–	–	+	–	+	+	+	–	+	–	–	–	–	–	–	+	+
# intvn.	*	*	*	0	2	*	5	*	*	1	*	5	3	2	*	0	*	*	*	*	*	*	1	6
Trial #	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
dyad	w	h	a	b	g	g	p	d	x	s	x	p	d	s	j	j	k	c	f	t	t	c	k	f
nov./fam.	+	+	+	+	–	+	–	–	–	–	+	+	+	+	–	+	–	–	–	–	+	+	+	+
# intvn.	6	6	6	5	*	0	*	*	*	*	1	4	4	3	*	0	*	*	*	*	0	3	5	5

Block 3.

Trial #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
dyad	al	ab	ae	ai	al	ai	ab	ax	ae	an	ax	aq	aq	an	ao	av	ar	at	at	av	ao	ar	as	ah
nov./fam.	–	–	–	–	+	+	+	–	+	–	+	–	+	+	–	–	–	–	+	+	+	+	–	–
# intvn.	*	*	*	*	3	1	4	*	5	*	2	*	0	3	*	*	*	*	0	3	5	4	*	*
Trial #	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
dyad	as	ag	ah	ac	aa	af	ag	aa	ad	ac	af	ad	ap	ap	am	au	am	ak	aj	au	aj	aw	aw	ak
nov./fam.	+	–	+	–	–	–	+	+	–	+	+	+	–	+	–	–	+	–	–	+	+	–	+	+
# intvn.	1	*	2	*	*	*	4	2	*	5	4	2	*	0	*	*	1	*	*	3	1	*	0	5

Table 1
Evaluated consonance for pure-tone dyads ranked by musical interval.

Intervals			Consonance rankings (from most to least)					
freq. ratio	size (semitones)	musical interval	P&L[a] 1965	M[b] 1918	G[c] 1928	V,L,&P[d] 1962	K&K[e] 1969	A,A,&W[f] 1980
1:2	12	**oct.**	oct.	oct.	M6	oct.	oct.	oct.
8:15	11	**M7**	M7	p5	M3	m6	M7	p5
9:16	10	**m7**	m7	M6	oct.	p5	m7	m3
3:5	9	**M6**	M6	p4	p4	M6	p5	p4
5:8	8	**m6**	m6	M3	m6	p4	M6	M3
2:3	7	**p5**	p5	m6	m3	tri.	m6	M6
32:45	6	**tri.**	tri.	tri.	p5	m7	p4	m7
3:4	5	**p4**	p4	m3	m7	M3	tri.	m6
4:5	4	**M3**	M3	m7	tri.	m2	M3	tri.
5:6	3	**m3**	m3	M7	M2	M7	m3	M2
8:9	2	**M2**	M2	M2	m2	m3	m2	M7
15:16	1	**m2**	m2	m2	M7	M2	M2	m2

[a] Sine-tone generator; dyads rated for “beauty,” “consonance,” and “euphoniousness.” Plomp & Levelt, 1965.
[b] Tuning forks; dyads rated for “blending” and “purity.” Malmberg, 1918.
[c] Helmholtz resonators; dyads rated for “pleasantness.” Guernsey, 1928.
[d] Electromotor with tone wheel; dyads rated for “rough” vs. “smooth.” Van de Geer, Levelt, & Plomp, 1962.
[e] Sine-tone generator; dyads rated for “clearness” vs. “turbidity.” Kameoka & Kuriyagawa, 1969a.
[f] Sine-wave oscillator; dyads rated on a 7-point scale for “consonance” vs. “complexity.” Ayers, Aeschbach, & Walker, 1980.

Table 2
Discriminability index (d'), response bias (c), and mean adjusted proportion correct (p(c)), averaged over five retention intervals for pure-tone dyads.*

Musicians	MC	mc	md	MD	SC	sc	sd	SD	Grand mean	Std. dev.
d'	1.70	1.71	1.52	1.58	1.53	1.37	1.52	1.96	1.61[b]	0.18
c	0.26	–0.17	–0.22	0.00	0.14	0.00	–0.15	–0.15	–0.04	0.17
*mean p(c)**	0.77	0.82	0.80	0.77	0.69[a]	0.77	0.80	0.85[a]	0.78	0.05
std. dev.	0.03	0.03	0.03	0.03	0.03	0.02	0.02	0.03		

Nonmusicians	MC	mc	md	MD	SC	sc	sd	SD	Grand mean	Std. dev.
d'	1.52	1.51	1.02	1.36	1.22	1.11	1.35	1.47	1.32[b]	0.19
c	0.44	–0.23	–0.24	0.03	0.11	–0.04	–0.22	0.03	–0.01	0.23
*mean p(c)**	0.67	0.79	0.75	0.72	0.62	0.74	0.79	0.74	0.73	0.06
std. dev.	0.04	0.04	0.02	0.04	0.04	0.04	0.04	0.03		

Note. Discriminability index (d') is the inverse of the normal distribution function (z) for the hit rate (H = familiar correct/number of familiar trials) minus the false alarm rate (F = novel incorrect/number of novel trials), i.e., $d' = z(H) – z(F)$. Criterion (c) is the response bias where $c = –0.5*(z(H) + z(F))$. A positive value of c indicates a tendency to respond "no" (i.e., "I have not heard that before"). A negative value of c indicates a tendency to respond "yes." When there is no response bias, the value is 0.0. The proportions correct adjusted for guessing (*p(c)**) are calculated from each participants' proportion of hits and false alarms by the formula $p(c)* = 0.5*[P(H) + (1 – P(F))]$ (Macmillan & Creelman, 2005, pp 3-31).
[a] Significantly different at $p < .05$. Other means were compared at specific retention periods due to sig. Retention by C/D interaction.
[b] Significantly different at $p < .01$.

Table 3
Trend analysis for eight classes of cognitive and sensory consonant/dissonant dyads across five retention periods from 16.00 to 48.00 s.

	Cognitive C/D				Sensory C/D			
Musicians	MC	mc	md	MD	SC	sc	sd	SD
trend	linear	-----	-----	-----	linear	linear	cubic	----
$F(1,3)$	9.82				6.88	7.69	4.80	
p value	.03				.01	.007	.001	
Nonmusicians	MC	mc	md	MD	SC	sc	sd	SD
trend	linear	linear	cubic	-----	-----	cubic	-----	linear
$F(1,3)$	23.69	6.86	4.80			15.82		9.81
p value	.001	.01	.007			.001		.003

Table 4
Significant nonparametric test results for classes of dyads at specific retentions.

			Friedman test			Wilcoxon test		
Group	C/D type	retention period	χ^2	*sig.*	*W*	pair	*z*	*sig.*
Mus.	cog.	24.25s	9.98	0.019	0.22	mc-MD	2.08	0.037
						md-MD	1.97	0.049[a]
	sen.	24.25s	8.20	0.042	0.18	SD-sd	2.49	0.013
						SD-SC	2.05	0.041
Nmus.	cog.	24.25s	11.06	0.011	0.25	MC-MD	2.14	0.033
						mc-MD	2.01	0.044[a]
	sen.	16.00s	10.91	0.012	0.24	sc-SC	2.52	0.012[a]
						sc-sd	2.37	0.018[a]
	sen.	40.75s	8.34	0.040	0.18	sc-SC	2.55	0.011[a]
						sd-SC	2.16	0.031

For brevity only the Friedman tests with a significance value of $p < .05$ are shown. Wilcoxon signed-rank comparisons with pairwise differences greater than a significance value of $p = .05$ are not shown. None of the pairwise differences are statistically significant under the conservative alpha of $p \leq .008$.
[a] Significantly different in post hoc tests following significant parametric ANOVAs.

Table 5

Mean recognition scores at 48.00 s in comparison to chance (0.50) performance.

	Musicians				Nonmusicians			
class	*M*	*SD*	*t*(14)	*sig.*	*M*	*SD*	*t*(14)	*sig.*
MC	0.62	0.23	2.09	*ns*	0.48	0.21	–0.36	*ns*
mc	0.77	0.40	2.62	*ns*	0.64	0.47	1.17	*ns*
md	0.61	0.38	1.27	*ns*	0.71	0.25	3.23	.006
MD	0.73	0.14	6.40	.000	0.69	0.19	3.77	.002
SC	0.45	0.51	–0.36	*ns*	0.57	0.48	0.56	*ns*
sc	0.60	0.28	1..40	*ns*	0.50	0.22	–0.03	*ns*
sd	0.67	0.25	2.57	*ns*	0.71	0.25	3.21	.006
SD	0.79	0.14	8.20	.000	0.63	0.30	1.70	*ns*

Note: $p \leq .05/8 = .006$ required for significance.

Figure Captions

Figure 1. Mean proportion of correct recognitions across retention periods for four levels of cognitive consonance/dissonance by musicians. At 48.00 s, musicians recognized MD dyads significantly better than chance.

Figure 2. Mean proportion of correct recognitions across retention periods for four levels of cognitive consonance/dissonance by nonmusicians.

Figure 3. Mean proportion of correct recognitions across retention periods for four levels of sensory consonance/dissonance by musicians.

Figure 4. Mean proportion of correct recognitions across retention periods for four levels of sensory consonance/dissonance by nonmusicians.

Figure 5. Raw number of correct trials (out of 6) for dyads based on their root note. Twenty out of 276 paired comparisons were significantly different. The four root notes at the left end of the x-axis were correctly recognized significantly more often than the eight root notes at the right end.

Figure 1

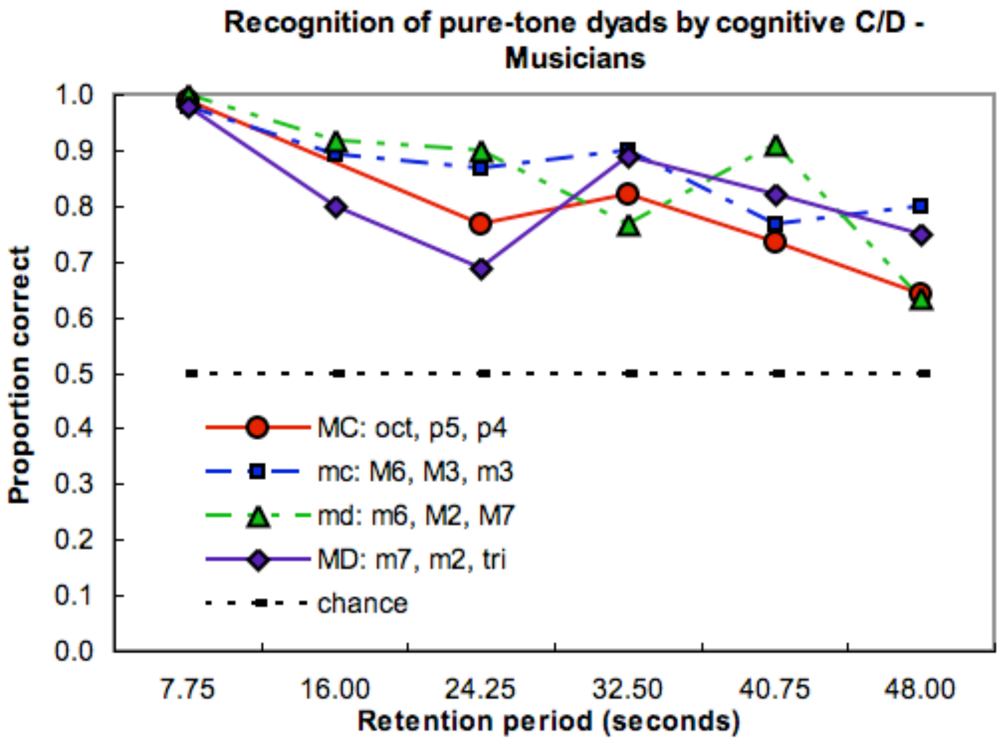

Recognition of pure-tone dyads by cognitive C/D - Musicians
Proportion correct
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
MC: oct, p5, p4
mc: M6, M3, m3
md: m6, M2, M7
MD: m7, m2, tri
chance
7.75
16.00
24.25
32.50
40.75
48.00
Retention period (seconds)

Figure 2

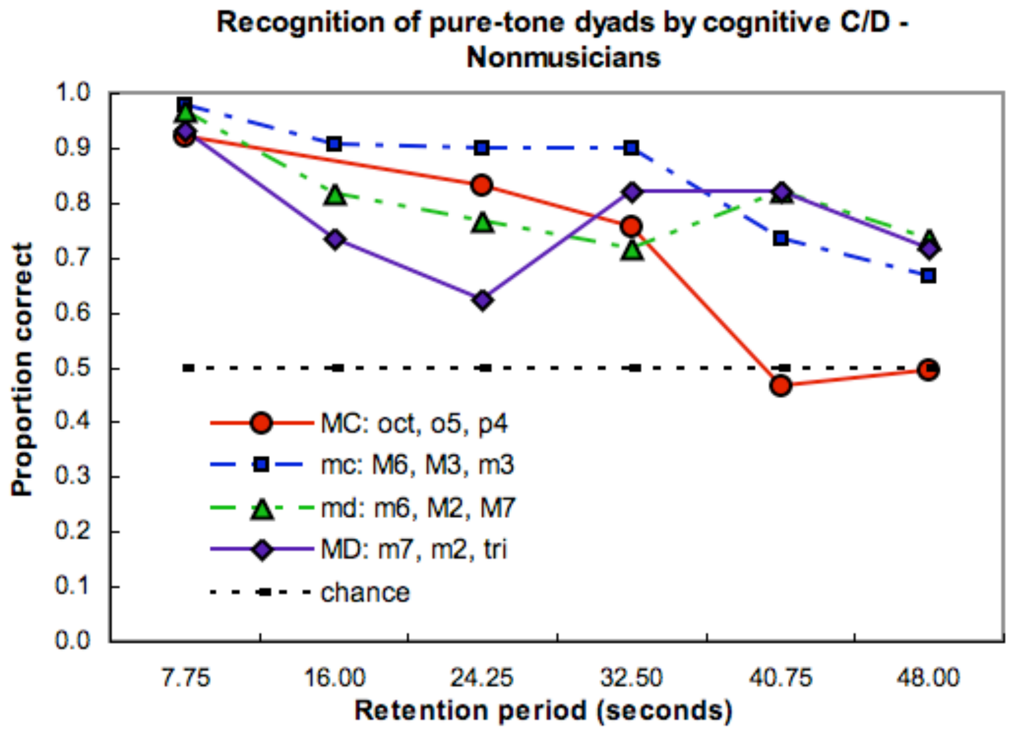

Recognition of pure-tone dyads by cognitive C/D - Nonmusicians
Proportion correct
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
MC: oct, o5, p4
mc: M6, M3, m3
md: m6, M2, M7
MD: m7, m2, tri
chance
7.75
16.00
24.25
32.50
40.75
48.00
Retention period (seconds)

Figure 3

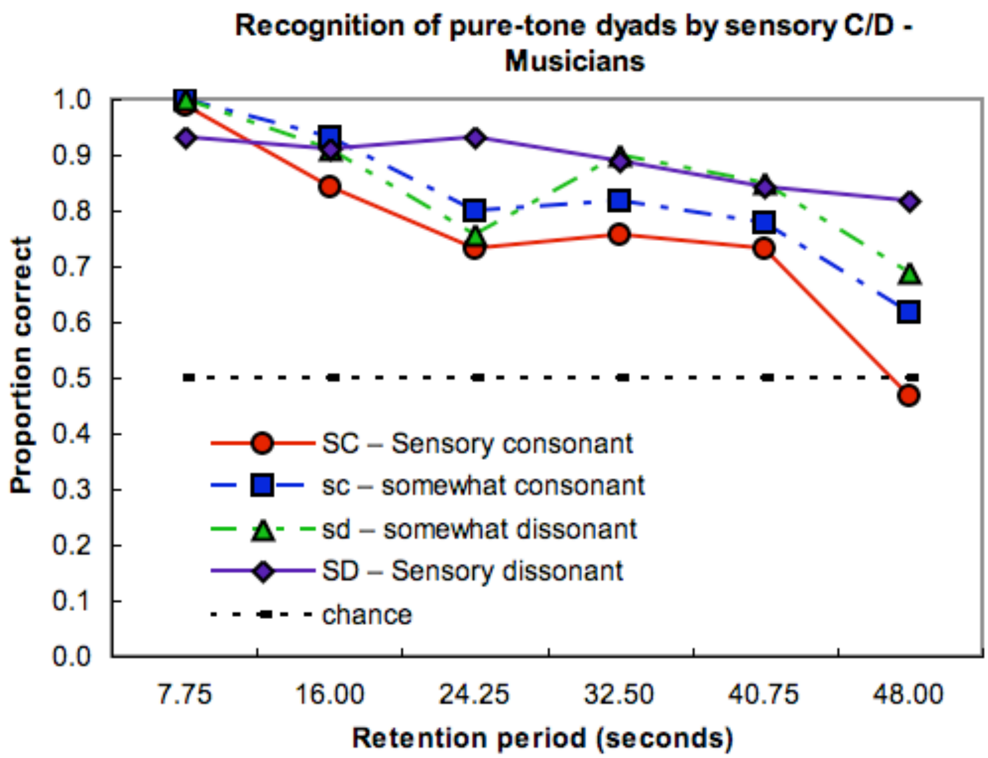

Recognition of pure-tone dyads by sensory C/D - Musicians
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Proportion correct
SC – Sensory consonant
sc – somewhat consonant
sd – somewhat dissonant
SD – Sensory dissonant
chance
7.75
16.00
24.25
32.50
40.75
48.00
Retention period (seconds)

Figure 4

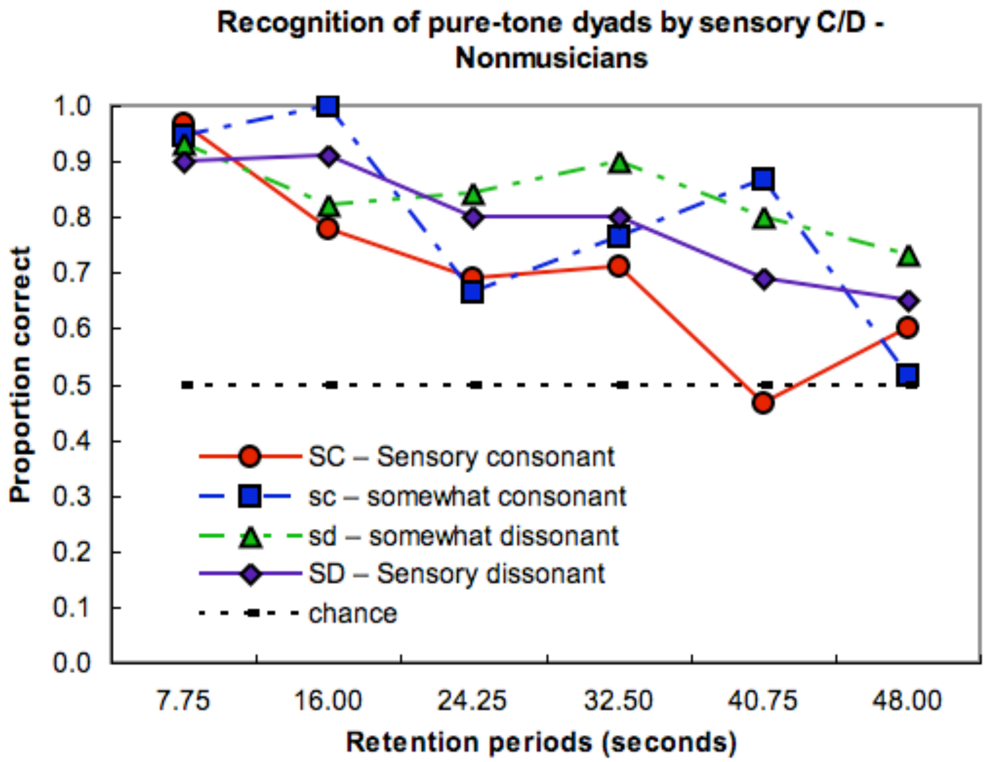
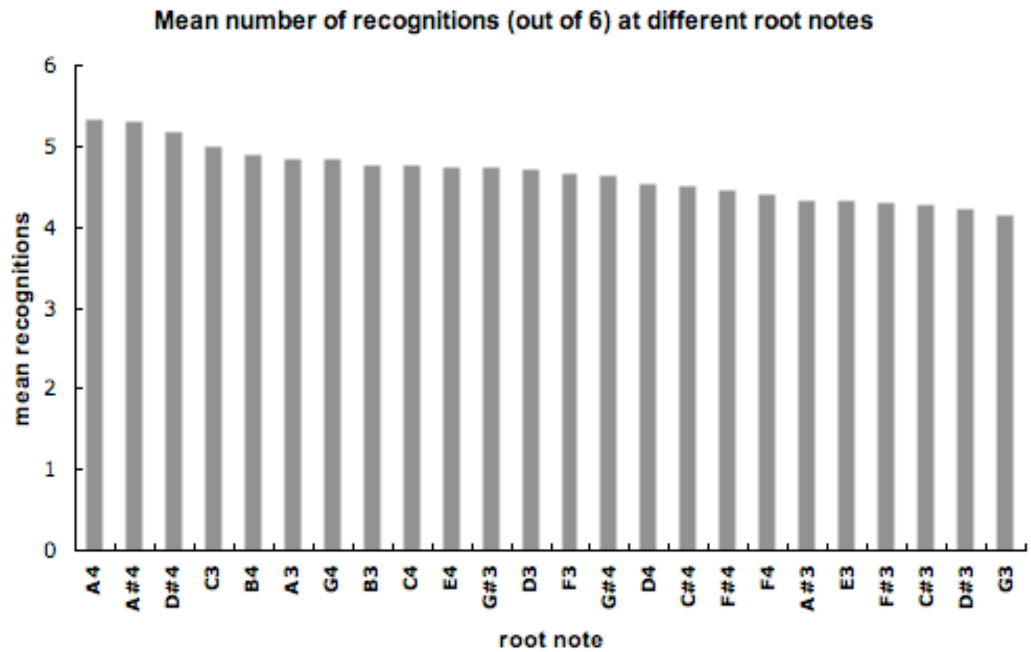

Recognition of pure-tone dyads by sensory C/D - Nonmusicians
Proportion correct
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
SC – Sensory consonant
sc – somewhat consonant
sd – somewhat dissonant
SD – Sensory dissonant
chance
7.75
16.00
24.25
32.50
40.75
48.00
Retention periods (seconds)

Figure 5



Mean number of recognitions (out of 6) at different root notes
mean recognitions
0
1
2
3
4
5
6
A4
A#4
D#4
C3
B4
A3
G4
B3
C4
E4
G#3
D3
F3
G#4
D4
C#4
F#4
F4
A#3
E3
F#3
C#3
D#3
G3
root note

Chapter 3 used a novel/familiar recognition memory paradigm to explore the cognitive underpinnings of the distinction between consonance and dissonance. The influence of consonance/dissonance on short-term memory (STM) for pure-tone *dyads* — two simultaneous pitches — was nonsystematic, but robust and accurate for a longer duration than expected. The work reported in Chapter 4 extends the findings of Chapter 3 by examining the influence of musical exposure on auditory STM. Experiment 1 uses the method reported in Chapter 3 but with complex-tone dyads that have greater ecological validity than the pure-tone dyads used in the previous report. Experiment 2 uses dyads mistuned from the familiar Western semitone standard to learn the extent to which knowledge of a musical, tonal system contributes to accurate STM for musical intervals.

CHAPTER 4

Short-term memory for consonant and dissonant complex-tone dyads — just- and microtuned

Susan E. Rogers[a)], Stephen McAdams, and Daniel J. Levitin

Author affiliations:

Department of Psychology, Schulich School of Music, and
Center for Interdisciplinary Research in Music, Media, and Technology, McGill University

[a)] McGill University
1205 Dr. Penfield Avenue, 8th floor
Montreal, Quebec, CA H3A 1B1
Electronic mail: susan.rogers@mail.mcgill.ca

Abstract

Two experiments use an auditory short-term memory paradigm to explore how the consonance and dissonance of simultaneous, complex-tone musical intervals (*dyads*) contribute to their robustness against interference and decay. Sixty listeners (30 musicians, 30 nonmusicians) of Western tonal music performed a novel/familiar recognition memory task featuring either common just-tuned dyads or unfamiliar microtuned dyads (mistuned from common musical intervals by a quarter tone). As seen in an earlier finding with pure-tone dyads, auditory short-term memory persisted longer than previously reported for single tones (Winkler et al*.,* 2002), regardless of dyad classification or tuning system. Musicians achieved higher recognition scores than nonmusicians for both just- and microtuned dyads. The earlier observation that short-term memory for some classes of dyads was relatively poor at moderate retention periods and improved at longer retention periods was replicated. Small-integer ratio dyads ("natural intervals") conveyed no innate memory advantage; musicians' and nonmusicians' recognition of these was no better or worse than for large-integer ratio dyads. Short-term memory for microtuned dyads was essentially the same as that for just-tuned complex-tone dyads in terms of overall accuracy. This suggests that categorical exemplars of musical intervals retrieved from long-term memory were unlikely to have contributed to accurate recognition. These findings contribute to the study of auditory signal processing by mapping the fate of musical interval memory traces over time.

keywords: auditory short-term memory, consonance, dissonance, auditory processing, microtuning

Introduction

Auditory short-term memory

In his 1993 chapter on auditory memory, Crowder wrote, "auditory processing results in auditory memory" (p. 116). This simple statement followed a passage from Squire (1987) on vertebrate habituation relating behavioral changes to modifications of existing neural architecture. Crowder used this point to argue for a so-called procedural approach to memory research, regarding memory less in terms of a mental faculty or capacity but more as the "persistence that is a by-product of some original mental activity" comprising a learning episode. For auditory memory in particular he promoted the use of a variety of experimental techniques to track what he termed "authentic" auditory memory — where verbal and visual encoding are prohibited by the stimuli or task. Crowder believed that differential memory for two auditory events reflect underlying processing differences, i.e., that existing neural architecture is modified to regard one event as fundamentally or perhaps categorically different from the other. This procedural approach to memory research necessitates a variety of experimental techniques and Crowder noted that inconsistent results might be expected in the early stages of the work. The current experiment adopts Crowder's procedural approach by using a behavioral measure to examine short-term memory (STM) for dyads, in hopes of revealing organizing principles in the cognitive processing of musical intervals.

Demany and Semal (2007) agreed with Crowder that psychoacousticians have paid comparatively little attention to auditory memory. They attributed this to the field's (false) assumption that auditory memory is strongly dependent on attention. In describing the role of memory in auditory perception, Demany and Semal review the compelling evidence from a number of studies showing that STM for nonlinguistic sounds is automatic and does not necessarily depend on attention or the rehearsal strategies described in Baddeley and Hitch's (1974) phonological loop. For example, in tests of STM for a single pitch, rehearsal by humming is shown to degrade STM rather than to enhance it (Kaernbach & Schlemmer, 2008; Massaro, 1970). The automaticity of auditory STM processing allows researchers to test Crowder's idea that different types of signals recruit nonidentical information processing that could lead to differential retention. Differential memory for types of sounds as displayed in behavioral paradigms can help map the temporal course of auditory feature extraction and categorization.

The two experiments described here are the second and third in a study of STM for *dyads* (two simultaneous tones) based on the premise that attributes of sounds contribute either positively or negatively to their robustness against interference from incoming sounds and decay over time. We have taken as a starting point the concept of "natural intervals" — those based on tone combinations that are in small-integer frequency-ratio relationships found between nearby partials at lower ranks in the harmonic series — to include the proposition that humans show a processing advantage for such intervals (Schellenberg & Trehub, 1994a,b, 1996; Trainor & Heinmiller, 1998). The human preference for small-integer ratio (consonant) intervals is deemed to be a musical universal rooted in our exposure to speech and other critical periodic sound stimuli (Krumhansl, 1990; Schwarz, Howe, & Purves, 2003; Terhardt, 1974b). If innate processing advantages exist for consonance, a cognitive difference between consonant and (their opposite) dissonant intervals should appear outside of a musical, tonal context in persons with and without formal musical training. To our knowledge this experimental approach — STM for dyads in a nontonal context — has not been used to study processing differences among categories of musical intervals.

The consonance/dissonance (C/D) distinction ensures that a set of dyads will be perceived as different from one another, but items that differ in their sensory attributes do not necessarily have corresponding differences in their cognitive attributes. This is especially true for sounds that involve learning or cultural transmission, such as musical scale systems. Whether dyads' sensory (e.g., rough versus smooth) or cognitive (e.g., large-integer versus small-integer frequency ratio) properties have an effect on STM persistence is unknown. The processing of the sensory C/D of chords is known to be functionally distinct from their cognitive, music-theoretic processing both in a musical context (Bigand & Tillmann, 2005; Regnault, Bigand, & Besson, 2001) and when presented at random (Itoh, Suwazono, & Nakada, 2003). Given that sensory and cognitive attributes are separable, the use of an STM task that is independent of conscious remembering is ideal for exploring differences between classes of sounds.

Accurate STM recognition is accomplished through either of two processing strategies. One permits recognition via the conceptual reprocessing of *stimulus meaning* (the observer "knows" the item) and the other permits recognition via enhanced perceptual fluency resulting from reprocessing the *stimulus form* (a sensation is recalled) (Roediger, 1990; Wagner & Gabrieli, 1998; Yonelinas, 2001). These two strategies allow nonexperts to recognize dyads as well as expert musicians do, based on information present in the stimuli at hand. Unlike working memory tasks where mental operations are performed on active traces, a STM novel/familiar task can be performed without employing a rule or an abstract concept (Engle, Tuholski, Laughlin, & Conway, 1999). Familiar recognition is a relatively simple cognitive task that tests the ability to decide whether or not a trace was left by a recently encountered object or event and thus can be served by either implicit or explicit memory processes (Roediger, 1990).

Previous findings

The main finding from the first experiment in this series (see Chapter 3) showed that memory for pure-tone dyads was significantly better than chance performance at retention lengths exceeding 30 s, regardless of musical training. Our hypothesis of differential memory along a cognitive (frequency ratio) axis of C/D was not supported because small-integer ratio dyads were not recognized more accurately than large-integer ratio dyads. No systematic effects of C/D classification on recognition memory as a function of time were revealed. A slight memory advantage for sensory dissonant (rougher) over sensory consonant (smoother) dyads was observed but the difference was just below statistical significance. An unexpected finding was that recognition scores for cognitively dissonant (large-integer ratio) dyads dipped and then improved as retention periods increased. Musicians outperformed nonmusicians by a slight but significant margin.

A limitation of the study was that only pure-tone dyads, the sum of two sine tones with no overtones or partials, were used. Pure-tone dyads are rarely encountered outside of acoustics laboratories. The relatively low ecological validity of the stimuli might have suppressed tendencies to process the dyads as meaningful musical events and encouraged processing based on the dyads' sensory qualities. The uniqueness of pure-tone dyads may have caused participants to engage in alternate modes of processing that are not normally employed for standard musical events.

Will complex-tone dyads, such as produced by musical instruments, demonstrate more resilient auditory STM? With higher ecological validity and additional stimulus features from harmonic partials imparting supplementary musical information, memory traces may persist longer. Consequently, overall recognition accuracy could exceed the mean 77.5% correct observed with pure-tone dyads. Although recognition accuracy is argued to depend primarily on fluency for the perceptual features of a stimulus (as opposed to conscious recollection or "knowing"; Jacoby & Dallas, 1981), when conceptual encoding is employed, recognition accuracy is enhanced (Wagner & Gabrieli, 1998). The (presumably) greater familiarity of complex-tone over pure-tone dyads is expected to stimulate conceptual encoding by virtue of familiarity and support improved recognition over the pure-tone dyad finding.

Sensory and cognitive consonance/dissonance of complex-tone dyads

Exploring the theoretical distinction between sensory consonance and tonal affinity is hindered by difficulties listeners sometimes have in isolating either quality (see Chapters 2 and 3 for a review). A perception of *sensory C/D* arises when attributes of a sound or interval produce either pleasantness or annoyance. Terhardt (1984) identified three attributes that alone or in combination contribute to the sensory dissonance of musical sounds: roughness and sharpness (a piercing quality — the weighted loudness of narrow band noise as a function of its spectral center), and the presence of toneness (the loudness weighted mean frequency on the Bark scale). Of these, roughness is considered the most important for its inverse relationship to sensory consonance (Helmholtz, 1885/1954; Terhardt, 2000) (although it has not been established to what degree auditory roughness is truly "unpleasant", musically or otherwise). Roughness results from the physical activity of two or more tones interacting in the auditory periphery and producing amplitude modulations in the range of 15-300 Hz (Van de Geer, Levelt, & Plomp, 1962; Zwicker & Fastl, 1990).

For music theorists, C/D is an evolving, conceptual distinction because Western tonal intervals termed dissonant in past centuries are evaluated as less so in the modern era (Cazden, 1980). *Cognitive C/D* (termed *musical C/D* by Terhardt, 1984) is linked to sensory C/D but the term typically refers to how sounds or intervals function in a tonal music system and it thus includes practical and cultural components. Complex-tone dyads are perceived as cognitively consonant when the relationship between the two fundamental frequencies can be described using small integers, e.g., 2:1 and 3:2 (octave and perfect 5th, respectively). As the complexity of the numerical expression increases, cognitive dissonance increases with the increasing number of noncoincidental partials present in the dyad. For example, octaves with root notes higher than C2 (65.4 Hz) or thereabouts cause perturbations in separate frequency filtering mechanisms in the cochlea; these intervals are subsequently judged as harmonious or consonant. In contrast, the just-tuned tritone (45:32 ratio) has many partials with closely spaced frequencies that cause modulated activity within a single auditory filter. The physical interaction caused by two or more frequencies within a filter's critical bandwidth are linked to the inharmoniousness typically reported for the tritone and other cognitively dissonant musical intervals (Helmholtz, 1885/1954; Plomp & Levelt, 1965; Terhardt, 1984). Beyond the physical interaction among auditory components, internalized tonal schemas developed through exposure to musical tuning systems bias individuals to regard some intervals as more consonant than others, even when intervals are presented outside of a musical, tonal context (Guernsey, 1928; Hutchinson & Knopoff, 1978; Kameoka & Kuriyagawa, 1969a, 1969b; Malmberg, 1918; Schellenberg & Trehub, 1994). Exposure to speech sounds has also been implicated in shaping the cognitive C/D perception of musical intervals (Deutsch, Henthorn, & Dolson, 2004; Ross, Choi, & Purves, 2007; Schwartz, Howe, & Purves, 2003; Terhardt, 1974b).

For the complex-tone dyads used in the two experiments reported here, cognitive C/D was assigned according to the frequency-ratio relationship between tones. Each dyad's sensory C/D was determined by its subjective roughness as evaluated in a previous study (see Chapter 2) and sensory C/D values were assigned accordingly.

Musical experience and musical interval processing

Exposure to musical intervals and knowledge of their function influences C/D perception by moderating an individual's expectations for what he hears (Bigand & Tillmann, 2005). The extent to which listeners can consciously disambiguate bottom-up (data-driven, sensory) and top-down (knowledge-driven, cognitive) processing of musical intervals is not fully known (Demany & Semal, 2007; Terhardt, 2000). Regnault et al. (2001) used the method of event-related potentials (ERP) to map chord processing as it transitioned temporally from sensory to cognitive C/D. Those individuals with musical training were shown to be faster than nonmusicians at differentiating and categorizing the C/D of intervals. Musicians also demonstrated a greater sensitivity to smaller differences between chords. Similar disparity to musical stimuli between musicians and nonmusicians has been reported elsewhere (Bigand, Madurell, Tillmann, & Pineau, 1999; Bigand & Pineau, 1997; Lee, Skoe, Kraus, & Ashley, 2009; Magne, Schön, & Besson, 2006; Neuhaus, Knösche, & Friederici, 2006; Strait, Kraus, Parbery-Clark, & Ashley, 2010; Tervaniemi, 2003; Zatorre & Halpern, 1979; see Chapter 2 for a review).

In a related fMRI study, the sensory and cognitive C/D of chords in a tonal context was independently manipulated (Tillmann, Janata, & Bharucha, 2003). A finding of increased neural activation in the inferior frontal areas for "deviant" targets — those that violated auditory expectancies — was reported. Activation was particularly strong for the most dissonant targets in the stimulus set. Taken together with evidence from language studies, this suggests that "the processing of deviants, or more generally of less frequently encountered events, may then require more neural resources than processing of more familiar or prototypical stimuli" (Bigand & Tillmann, 2005). What is "deviant" depends on expectations provided by context and, as shown in language paradigms, on the scope of participants' implicit and explicit experiences with the stimulus set.

In the absence of tonal context, fMRIs of musicians and nonmusicians reveal regional processing differences for consonant versus dissonant dyads (Foss, Altschuler, & James, 2007). When participants listened passively to dyads presented in isolation (without a melodic or tonal context), they exhibited neural activation in dissimilar brain regions depending on whether or not they had had musical training. The strength of the activity — proportional to the degree of C/D — was similar. Musicians showed neural activation correlated with degree of consonance in areas of the frontal, temporal, and parietal lobes, whereas nonmusicians displayed differential C/D activation only in the right inferior frontal gyrus. How these processing differences for C/D observed under passive listening conditions are manifested in other types of cognitive activity is the focus of the current investigation.

We report the outcome of presenting musicians and nonmusicians with common dyads from the Western tonal scale outside of a tonal context, in a novel/familiar short-term memory (STM) paradigm. If the most sensory or cognitively dissonant dyads recruit additional processing resources, these could show greater resistance to decay over time and the degrading effects of interference from incoming sounds. If, on the other hand, the human affinity for consonance conveys a processing advantage, consonant dyads could show greater recognition accuracy over dissonant dyads. Musicians are expected to show less differential STM across cognitive classes of dyads, due to their (presumably) greater familiarity with all types of musical interval classes.

To examine the effect of exposure to musical intervals, a second experiment (the third in this series) presents listeners with microtuned dyads — mistuned from just-tuned intervals by a quartertone. These intervals are not part of just- or equal-tempered tunings used in Western music and have the experimental advantage of being unfamiliar to the participants recruited here. To listeners unfamiliar with them, these microtuned dyads are all dissonant from both a sensory and cognitive C/D perspective. The use of unfamiliar dyads could reduce a performance gap between musicians and nonmusicians. If musicians' higher auditory acuity extends to microtuned dyads (and there is no a priori reason to believe it does not), this group could outperform the nonmusician group to the same degree as observed with just-tuned, familiar dyads.

Experiment 1: Just-tuned dyads: Method

Participants

Participants ($N = 30$; 13 men and 17 women; 20-55 years; $M = 30$, $SD = 11.9$) were recruited from a classified ad or were volunteers from the Schulich School of Music and Psychology Department at McGill University. Recruits were paid $10 for their time, and volunteers served without pay. Fifteen participants (musician group) had five or more years of formal music training ($M = 14$, $SD = 6.8$); the remaining 15 participants (nonmusician group) had two or fewer years of training ($M = 0.7$, $SD =$

0.8). All had normal hearing and none had absolute pitch perception by self-report. Musical training and music listening habits were assessed using a modified version of the Queen's University Music Questionnaire (Cuddy, Balkwill, Peretz, & Holden, 2005).

Two persons reported having tone-deafness. They completed the experiment but displayed an unacceptably high false-alarm rate (more than 2 standard deviations below the nonmusicians' mean proportion correct). Data from these persons are not reported here.

Stimuli and Apparatus

The digital audio processor ProTools (DigiDesign, Daly City, CA) was used to create 72 dyads by summing two complex tones. The McGill University Master Samples (MUMS volume 3, track 16, sample 16-03) provided a sample of an alto saxophone playing the note D#4. The sample was digitally transferred into the audio processor ProTools (Digidesign, Daly City, CA) and pitch-shifted either higher or lower as needed using the ProTools plug-in “Pitch & Time - Algorithm B” to create 33 upper notes and 24 lower notes. This technique preserves the relative amplitudes of the spectral components upon transposition. Dyads were created by combining in phase a lower frequency note ($f1$ or root note) with an upper frequency note ($f2$ or top note), each matched in amplitude. Dyads were normalized in amplitude so that each stimulus was presented to listeners at a sound pressure level of 57 dB $\pm$ 0.75 dBA SPL at the headphone as measured with a Brüel & Kjær 2203 sound level meter and Type 4153 Artificial Ear headphone coupler (Brüel & Kjær, Naerum, Denmark). Each dyad was 500 ms in duration, including 10-ms raised cosine onset and offset ramps.

Each dyad had its root note assigned at random. Root notes ranged from C3 (130.8 Hz) to B4 (493.8 Hz). The top notes ranged from D#3 (155.6 Hz) to A#5 (932.3 Hz). In keeping with previous research in the psychoacoustics of dyad perception, the frequency-ratio relationship of the intervals corresponded to the just-tuned scale. The stimulus set contained an equal number of small- and large-integer ratio dyads as well as a range of sensory C/D dyads as detailed below.

The 72 dyads formed the 12 musical intervals of the Western chromatic scale: minor 2nd (m2), Major 2nd (M2), minor 3rd (m3), Major 3rd (M3), perfect 4th (p4), tritone (tt), perfect 5th (p5), minor 6th (m6), Major 6th (M6), minor 7th (m7), Major 7th (M7), and octave (oct). Three unique intervals were created at each of the 24 root notes. The pitch of C3 (130.8 Hz), for example, was used to create m3, m6, and m7, and the pitch of C4 (261.6 Hz) was used to create octave, M7, and m2. The reciprocal nature of the design allowed each musical interval to be represented at six different root notes. The m2s, for example, had root notes at F3, G#3, A3, C4, E4, and G4. (See Appendix A for a complete description of each dyad.)

Following Schellenberg and Trehub (1994b), a rating of cognitive C/D based on frequency-ratio complexity was derived by taking the reciprocal of the natural logarithm of the sum of each musical interval’s integers. For example, an octave (2:1) produces a value of $1/\ln(2+1) = 0.91$, whereas a tritone (45:32) produces a value of 0.23 (see Table A3). The 72 dyads were each assigned to one of four frequency-ratio levels as follows: musically consonant or “MC” (octave, p5, p4), moderately consonant or “mc” (M6, M3, m3), moderately dissonant or “md” (m6, M2, M7), and musically dissonant or “MD” (m7, m2, tritone). (The frequency-ratios for the just-tuned complex-tone dyads were the same as for the just-tuned pure-tone dyads in the earlier experiment, so the labeling scheme was kept.) The 72 dyads were also assigned one of four labels for sensory C/D, based on the mean value of the roughness rating of each dyad obtained from 30 participants (15 musicians, 15 nonmusicians) in a separate experiment (see Chapter 2). The four classes were labeled from the "smoothest" to the "roughest" as follows: smooth complex tone or “SCT,” somewhat smooth complex tone or “sct,” somewhat rough complex tone or “rct,” and rough complex tone or “RCT.” (This labeling scheme was applied to dyads in the current experiment to discourage direct comparisons between stimulus sets. Roughness values were obtained independently for each set of dyads in this series of experiments. A just-tuned complex-tone dyad labeled “somewhat rough” is not necessarily equivalent to a “somewhat rough” microtuned dyad from a sensory perspective.)

Sessions were conducted in a soundproof booth (IAC, Manchester, U.K.) with participants seated at a Macintosh G5 PowerPC computer (Apple Computer, Cupertino, CA). Tone sequences were delivered from the Macintosh’s digital output to a Grace m904 (Grace Design, Boulder, CO) digital audio converter and presented to listeners in stereo through Sennheiser HD280 pro 64 Ω headphones (Sennheiser, Wennebostel, Germany).

Procedure

(Note: The paradigm used in the present experiment is identical to the one described in Chapter 3 and is included here for completeness.)

The 72 dyads were subdivided and presented in three blocks in which participants heard 24 dyads twice — first as a novel item and then later as a familiar item — comprising a grand total of 144 trials (72 × 2). Each block contained a unique subset of dyads: two representing each musical interval, with one from the lower octave and one from the higher octave. Thus listeners responded to 48 items in a single block [12 intervals × 2 octaves × 2 presentations each (novel/familiar) = 48 trials]. For purposes of balancing the design, item order was carefully controlled within each block as detailed below. The order of presentation of the three blocks was randomized for each participant. (See Appendix C for the arrangement of dyads in blocks.)

Stimuli were presented at a fixed rate; time for a single trial was 8.25 s. A *trial* consisted of a 250-ms pre-stimulus alert, a 500-ms stimulus, a 3-s response window, a 500-ms feedback window, and a silent (unfilled) 4-s wait period. Elapsed time for each block was 6 min 36 s (8,250 ms × 48 trials). Participants were allowed a pause between blocks but voluntarily completed the experiment in less than 25 min.

Task difficulty was designed to range from very easy to very difficult to provide various levels of cognitive load and elicit differential performance. The *retention period*, the time from when a novel dyad (**S–**) first appeared to when it reappeared as a familiar dyad (**S+**), should, in theory, influence the listener’s ability to recognize familiar stimuli, so dyad distribution across retention periods was balanced as well as possible. A given **S–** could have zero, one, two, three, four, five, or six intervening stimuli before it reappeared as an **S+**, corresponding to retention durations of 7.75, 16.00, 24.25, 32.50, 40.75, 48.00, or 56.25 s respectively. Constraints of stimulus sequence arrangement necessitated four instances of 56.25 s retention periods (see Appendix C). Data from these periods were not analyzed due to low predictive power from having too few exemplars.

Presumably, the familiar presentation of a dyad shortly after its novel presentation is cognitively easier than if the familiar presentation occurs more than 32.50 s after the dyad's novel presentation. Thus dyads recognized after 7.75 s of retention served as catch trials because performance was expected to be at or near ceiling in these instances (in a sequential task, however, proactive interference could degrade recognition accuracy at short retention durations). As much as possible, the distribution of dyads was evenly spaced among the seven possible retention periods so that memory performance for each type of dyad could be explored at every level of cognitive load.

Participants were tested individually using an original computer program to adapt the auditory memory paradigm of Cowan, Saults, and Nugent (1997). They sat at the computer wearing headphones and used the keyboard to respond. Participants were instructed that they would hear musical sounds, each presented twice within a block of 48, and that their task was to listen carefully and use the computer keyboard to indicate whether they were hearing a sound for the first or second time. They were told that the second presentation would occur no more than six trials after the first so that the number of sounds they needed to remember would never exceed six. Participants were informed that there would be visual feedback and that the task difficulty was such that they could expect a moderate percentage of wrong answers. They were instructed not to worry about wrong answers, but to try to remember as many sounds as possible. The experimenter familiarized the participant with the task by running a short (eight trial) practice sequence before testing began. Results from the practice session were not analyzed.

The color of the screen changed from yellow to red to alert participants that a trial was about to begin. The question "HAVE YOU HEARD THIS BEFORE?" appeared on the red screen. The screen stayed red as a dyad was played through the headphones, then turned green for the response phase. Participants had 3 s to respond by pressing a “NO” key (the “–” key on the number keypad) if they believed they were hearing a novel stimulus or a “YES” key (the “+” key on the number keypad) if they thought it was familiar. After the 3-s response window, the screen stayed green while "CORRECT" or "WRONG" appeared for 500 ms. If no answer was entered, the trial was scored as incorrect, and the “WRONG” feedback appeared. The screen turned yellow during a 4-s (silent) inter-stimulus period. After the waiting period the screen turned red for the next stimulus.

At the end of blocks 1 and 2, a “Please take a rest” message appeared on the screen for 3 s. This was followed by a “Hit any key to continue” message that allowed the participant to initiate the next block when ready. After the third block, a “Finished — Thank you!” message appeared on the screen. Upon exiting the booth, participants were asked to describe any strategies they used to recognize dyads and to give their impression of the task.

Results

Data Analysis

The current study reproduced the paradigm used with pure-tone dyads (see Chapter 3) with the aim of further examining some unexpected findings, namely: robust recognition accuracy beyond 30 s of retention, differential memory performance at moderate retention periods that disappeared at longer retention periods, and performance differences between musicians and nonmusicians. The current data analysis followed the methods used in the earlier experiment to afford some comparisons between these and the previous results. As before, data was analyzed under two separate classifications of C/D: cognitive and sensory.

Analysis methods from Signal Detection Theory (SDT) compared accuracy, sensitivity, and response biases between participant groups. Hit rates and false alarms were plotted on receiver-operating characteristic (ROC) curves (not shown) so that floor and ceiling effects could be avoided (which distort SDT methods of analysis; Yonelinas, 2001). An outcome of the ROC plot was to discard data from two subjects, as described in the Participants section.

A stepwise logistic regression was used to verify that retention period length indeed had a significant effect on recognition accuracy.

Hypothesis testing of C/D effects was explored through repeated-measures ANOVAs as dictated by the variable(s) of interest and their interaction. These tests needed to compensate for unequal cell sizes at some retention periods (see Appendix C), so each participant's proportion of correct answers for the variable of interest, rather than his or her raw score, was used as the dependent variable. Each proportion correct was adjusted for guessing by the following formula: $p(c)^* = 0.5[P(hits) + (1 - P(false\ alarms)]$ (Macmillan & Creelman, 2005). Where assumptions of sphericity were not met, significance tests were corrected using either a Huynh-Feldt correction when epsilon was large ($\varepsilon > 0.75$) or a Greenhouse-Geisser correction otherwise. In these instances the original degrees of freedom (*df*), epsilon and corrected *p* value were reported. Post-hoc tests used Tukey's Honestly Significant Difference (HSD) with a Type I error rate of 0.05 as the criterion for pairwise comparisons. For post hoc tests in cases where the sphericity assumption of the ANOVA was violated, *t*-tests with Bonferroni corrections of $0.05/k$ were used for the significance criteria because this correction is more robust to violations of sphericity than Tukey's HSD (Stevens, 2002, p. 509). As in the previous study, tests of C/D analyzed data from familiar trials only. The process of recognizing a novel item (for which no trace exists in the brain) was not explored with regard to C/D.

To account for the possibility that C/D recognition scores were not normally distributed, Friedman's nonparametric ANOVAs were conducted in addition to the parametric tests (Conover, 1971). Pairwise comparisons following significant Friedman tests used the Wilcoxon test. To examine recognition accuracy under the heaviest cognitive load (corresponding to 48.00 s retention), single-sample *t*-tests compared the mean $p(c)^*$ scores for each C/D classification against chance performance.

The secondary variables of root note, stimulus block, and presentation order were examined through one-way ANOVAs using participants' raw scores (correct/incorrect) as the dependent variable.

Overall performance and comparison of musicians and nonmusicians.

Short-term memory for complex-tone dyads approximated that of the pure-tone dyads. Most of the departures in the present results were observed in the nonmusicians' data. Participants averaged 77.1% correct in the task overall (overall performance was 77.5% for pure-tone dyads). The hypothesis that STM for complex-tone dyads would be more accurate than for pure-tone dyads was not supported, $t(29) = 0.503$, $p = .62$. The mean proportion correct in familiar trials was once again significantly higher than in novel trials ($M_{familiar} = 0.81$, $SD = 0.07$; $M_{novel} = 0.73$, $SD = 0.11$) as confirmed by a paired-samples *t* test, $t(29) = 3.34$, $p < .01$. The proportion of hits and false alarms was calculated for each participant; the proportion of hits was then adjusted for guessing as described above. Mean corrected hit rates (p(c)*), sensitivity to novel/familiar status (d'), and response biases (c) for each group (musician and nonmusician) are displayed in Table 1.

Insert Table 1 About Here.

As seen in an earlier experiment (see Chapter 3), a Mann-Whitney *U* test showed that the probability distribution of musicians' and nonmusicians' *d'* values were significantly different, indicating that musicians were likelier than nonmusicians to discriminate novel from familiar dyads, $U = 59.00$, $Z = –2.22$, $p = .03$.

The difference between groups was larger in the current experiment than seen with pure-tone dyads. The mean d' index for musicians was equivalent to the pure-tone dyad result, $t(7) = 0.16$, $p = .88$, but nonmusicians' mean d' dropped significantly, $t(7) = 2.75$, $p = .03$. This result suggests that complex-tone dyads are less memorable or perhaps more similar to one another for nonmusicians than are pure-tone dyads (cf. Table 2, Chapter 3). Musicians once again showed a relatively high sensitivity to sensory dissonant dyads (RCT) but nonmusicians were most sensitive to cognitively consonant dyads (MC and mc). The criterion value (c) for each dyad class was calculated to measure participants' decision rules or likelihood of responding either "yes" or "no" (see Macmillan & Creelman, 2005, pp. 29-31). Response biases were low or moderately low, ranging from –0.20 (musicians, MD) to 0.24 (musicians, RCT), indicating that musicians tended to regard MD dyads as familiar and RCT dyads as novel.

The effect of retention period.

Recognition performance was expected to decline over time but duration of retention period was not predicted to be the sole influence on response accuracy. A binary logistic regression confirmed that of the many variables analyzed retention period had the largest impact on recognition score. A model with only the retention period variable was not a significant improvement over the null (intercept-only) model. The forward, stepwise regression analysis entered the following variables: Retention Period, Cognitive C/D (4 levels: MC, mc, md, MD), Sensory C/D rating (4 levels: SCT, sct, rct, RCT), Root Note, Top Note, and Block Order (first, second, or last). The dependent variable was the response score (correct/incorrect) on familiar trials. A Hosmer-Lemeshow goodness-of-fit test indicated that the data did not fit the model well, $\chi^2(5, N = 2160) = 45.51, p < .001$. The Nagelkerke R^2 showed that only 10% of the variance in the data was explained by retention period alone.

The effects of consonance and dissonance.

As noted above, the distinction between cognitive and sensory C/D can be discussed in theory with greater confidence than can be observed in behavior. The following analyses regard STM for dyads along these two axes but do not assume that listeners perceived each dyad as having a dual nature.

For averaged scores across retention periods, the relationship between expertise and accurate dyad recognition was ordinal — musicians achieved higher corrected scores in all eight dyad classifications over nonmusicians (see Table 1). An omnibus ANOVA was conducted with $N = 30$ to explore the combined effects of musical expertise, retention period, and C/D classification. Two three-way (2 × 5 × 4) repeated-measures mixed-design ANOVAs were conducted using Expertise as a between-subjects factor (2 levels: musician and nonmusician), plus two within-subject factors — Retention (5 levels: 16.00, 24.25, 32.50, 40.75, and 48.00 s) and C/D (4 levels in each test: MC, mc, md, and MD, *or* SCT, sct, rct, and RCT). Data from the 7.75 s retention period were omitted because performance was at ceiling and did not contribute anything meaningful to the research questions. The ANOVAs each revealed a significant effect of Expertise: ANOVA with Cognitive C/D, $F(1, 28) = 10.96, p < .01$; ANOVA with Sensory C/D, $F(1, 28) = 8.13, p < .01$. There were no significant interactions of Expertise with either Retention and/or C/D. Regardless, subsequent tests segregated the data from the two participant groups to reduce error (within-cell variance) and increase statistical power (Stevens, 2002, p. 323). The following tests were performed on musician and nonmusician data using two-way (5 × 4) repeated-measures ANOVAs with Retention and C/D as within-subject independent variables and proportion correct adjusted for guessing (p(c)*) as the dependent variable.

Cognitive C/D. Musicians' recognition of cognitive C/D dyads decreased with increasing retention period, as expected (see Fig. 1). This result, unlike that observed with pure-tone dyads, showed no performance dips at 24.25 s. The class of MD dyads, poorly recognized by musicians in the pure-tone experiment, was well recognized with the current dyads having harmonic partials. The effect of Retention was significant, $F(4, 56) = 5.82$, $\varepsilon = .68$, $p < .01$, but Cognitive C/D, $F(3, 42) = 2.54$, $\varepsilon = .61$, $p = .10$, and the C/D × Retention interaction, $F(12, 168) = 0.88$, $p = .57$, were not. Because C/D approached significance, post hoc tests were run to compare the mean p(c)* scores in four Cognitive C/D classes across retention periods. Musicians' recognition accuracy for MD dyads ($M = 0.84$, $SD = 0.03$) was significantly higher than for md dyads ($M = 0.73$, $SD = 0.03$) using a *t*-test and Bonferroni-corrected alpha of $p = .05/6 = .008$ to control for Type I error. (As noted above, Tukey's HSD was only used as a post hoc test when comparing means in cases where sphericity assumptions were met.)

Insert Figure 1 About Here.

Nonmusicians' recognition of dyads classified by cognitive C/D was more strongly moderated by retention period (see Fig. 2). This group recognized small-integer (MC, mc) more accurately than large-integer ratio dyads (md, MD). The effect of Retention was significant, $F(4, 56) = 3.88$, $\varepsilon = .69$, $p = .03$, as was the Retention × Cognitive C/D interaction, $F(12, 168) = 2.33$, $p < .01$. There was no global effect of C/D, $F(3, 42) = 0.30$, $p = .82$. Tests of simple main effects of Cognitive C/D at each retention period followed the significant interaction. Recognition accuracy was poor at 24.25 s for mc dyads but this class improved at 32.50 s retention. A significant effect of Cognitive C/D class was observed at 24.25 s, $F(3, 42) = 3.89$, $\varepsilon = .68$, $p < .03$. Post hoc tests used a *t*-test and Bonferroni adjustment to control for Type I error, $p = .05/6 = .008$ required for significance. At 24.25 s retention nonmusicians recognized md dyads ($M = 0.84$, $SD = 0.14$) and MD dyads ($M = 0.83$, $SD = 0.21$) significantly better than mc dyads ($M = 0.57$, $SD = 0.23$). Other comparisons fell short of significance under the corrected alpha.

Insert Figure 2 About Here.

Sensory C/D. Dyad recognition as a function of sensory C/D showed a stronger effect of retention period than observed with the cognitive C/D distinction. This observation contrasted with the pure-tone dyad finding. Musicians' average recognition scores in sensory C/D classes spanned a wider range than seen in cognitive C/D classes (see Table 1). Recognition of the most dissonant dyads (RCT) dipped sharply at 48.00 s — only one-third of the familiar responses were correct (see Fig. 3). Musicians showed significant main effects of Retention, $F(4, 56) = 7.99$, $p < .001$, and Sensory C/D, $F(3, 42) = 3.03$, $p < .05$. There was a significant Retention × C/D interaction, $F(12, 168) = 2.15$, $\varepsilon = .72$, $p = .03$. Following the significant interaction, the simple main effects of C/D were explored at each of five retention periods. A significant effect of Sensory C/D was observed at 48.00 s , $F(3, 42) = 7.57$, $\varepsilon = .66$, $p < .01$. Pairwise comparisons of the means at 48.00 s used a *t*-test with a Bonferroni correction of 0.008 for significance. Recognition accuracy was significantly higher for sct dyads ($M = 0.77$, $SD = 0.05$) than for RCT dyads ($M = 0.32$, $SD = 0.12$) at 48.00 s. The difference between SCT ($M = 0.78$, $SD = 0.06$) and RCT dyads just missed significance by the corrected alpha ($p = .009$). (Note: Although the mean for SCT dyads was slightly higher than the mean for sct dyads, the power of the test was reduced due to fewer SCT than sct dyads at 48.00 s.)

Insert Figure 3 About Here.

Nonmusicians' dyad recognition by sensory C/D class clearly showed a strong effect of retention period (see Fig. 4). Although the relationship between sensory C/D class and retention period was significant, it showed no systematic pattern. There was a significant main effect of Retention, $F(4, 56) = 5.15$, $p = .001$, and a significant Sensory C/D x Retention interaction, $F(12, 168) = 3.34$, $p < .001$. The main effect of Sensory C/D was not significant, $F(3, 42) = 1.79$, $p = .164$. The simple main effects of Sensory C/D were tested at each of five retention periods. Tests at three retention periods were significant: 24.25 s, $F(3, 42) = 3.47$, $p = .02$; 32.50 s, $F(3, 42) = 3.99$, $p = .01$; and 48.00 s, $F(3, 42) = 3.39$, $\varepsilon = .66$, $p = .05$. Post hoc comparisons of the means at 24.25 s used Tukey's HSD and showed that RCT dyads ($M = 0.85$, $SD = 0.12$) performed significantly better than SCT ($M = 0.56$, $SD = 0.47$) and sct ($M = 0.57$, $SD = 0.23$) dyads. At 32.50 s, RCT dyads ($M = 0.86$, $SD = 0.25$) were recognized significantly more often ($p < 0.05$) than SCT ($M = 0.61$, $SD = 0.19$) and sct ($M = 0.66$, $SD = 0.16$) dyads. Due to a sphericity violation, comparisons at 48.00 s used a Bonferroni adjustment to control for Type I error instead of HSD, where $p = .05/6 = .008$ was required for significance. Recognition accuracy among classes of Sensory C/D was not significantly different at 48.00 s by the more conservative test.

Insert Figure 4 About Here.

The inferential tests indicated that the effect of retention period on STM for dyads is not systematic by C/D class or type. In this sense, the finding replicated the pure-tone dyad result. The present result also contrasts with the previous finding. Cognitively dissonant *pure-tone* dyads were poorly recognized at moderate retention periods (see Chapter 3, Figs. 1, 2) but cognitively dissonant *complex-tone* dyads were well recognized under the same experimental paradigm (Figs. 1, 2).

The interdependence of the C/D classifications complicated these analyses. Departures from either normal distributions or homogeneity of variance can affect the power of ANOVA tests (Stevens, 2002, pp. 256-267). Covariance issues were addressed by adjusting the degrees of freedom using either a Greenhouse-Geisser or Huynh-Feldt correction (depending on the severity of the violation) to provide a more accurate significance test. To address issues of normal distribution, Shapiro-Wilk tests were performed. A grand total of 80 tests of normality were performed — one for each of the eight C/D types at the five retention periods of interest for the two participant groups. The result showed that approximately one quarter of all score distributions were significantly skewed from normal, under a Bonferroni corrected alpha of $p = .05/20 = .0025$. (Note that the N is 20 and not 80 because the analysis segregated participant groups and C/D types. For example, one test looked at musicians' four classes of sensory C/D data at the five retention periods.)

Because the parametric ANOVAs found results that were skewed from the normal distribution, nonparametric repeated-measures ANOVAs were performed on these data to verify the parametric results. Twenty-one Friedman tests examined differences among the C/D classifications and retention periods by participant group. As before, the dependent variable used the average p(c)* for each participant. Five Friedman tests were significant at $p < .05$. The Wilcoxon test was chosen as a post hoc to the significant Friedman tests. A single significant difference under a conservative alpha of $p = .05/6 = .008$ was found at 24.25 s, where nonmusicians recognized RCT dyads significantly more often than sct dyads (see Table 2). Several other comparisons approached significance and are included in the table.

Insert Table 2 About Here.

The effect of secondary variables.

Root pitch, stimulus block, and presentation order had a smaller impact on recognition memory than did retention period, although the effect of root pitch was not trivial. Cell sizes for these variables were equally distributed across the design; therefore the unadjusted raw scores (correct/incorrect) from both novel and familiar trials ($N = 30$) were used as the dependent variables in repeated-measures ANOVAs. As described in the Methods section, listeners heard each of the 24 root notes six times in three novel and three familiar trials, thus a score of six was the maximum possible correct at each note. Root pitch had a significant effect on recognition memory, $F(23, 667) = 5.10$, $\varepsilon = .92$, $p < .001$. Post hoc paired comparisons used a Bonferroni corrected alpha of $p = .05/276 = .0002$ to control for Type I error. Twenty-nine of the 276 comparisons were significantly different. Mean correct scores by root note are listed in Table 3 and ranged from the most recognized note C4 ($M = 5.6$, $SD = 0.6$) to the least recognized note D#4 ($M = 4.0$, $SD = 1.2$).

The six most-recognized root notes were in the fourth octave, echoing the pure-tone dyad result. A paired-samples *t* test was conducted to examine the effect of root note octave on dyad recognition. The null hypothesis was rejected; dyads with root notes in the fourth octave were recognized significantly more often than dyads rooted in the third octave, $t(29) = 3.33$, $p < .01$.

Insert Table 3 About Here.

The three stimulus blocks were equally challenging; differences among them were not significant, $F(2, 58) = 1.11$, $p = .37$. The means and standard deviations out of the maximum possible score of 48 correct were $M_{Blk1} = 37$, $SD = 4.1$; $M_{Blk2} = 36$, $SD = 3.0$; $M_{Blk3} = 37$, $SD = 3.8$. The order of presentation (first, second, or last) of the stimulus blocks was randomized for each participant and showed no primacy, practice, or other effect of order, $F(2,58) = 0.12$, $p = .890$. The means and standard deviations for first, second, and last blocks were: $M_{First} = 37$, $SD = 3.4$; $M_{Second} = 37$, $SD = 3.6$; $M_{Last} = 37$, $SD = 4.1$.

Tests of auditory memory duration.

The pure-tone dyad experiment found that STM for many dyad classifications was robust and accurate at 48.00 s of retention — an especially heavy cognitive load. Roughly half of the complex-tone dyad classes were also recognized statistically better than chance at the same retention period. Paired-sample *t*-tests compared the recognition scores for all classes of dyads against chance (0.50) performance; eight tests were conducted for each participant group. The dependent variable was participants' p(c)* scores for each class at 48.00 s. It was not assumed that the eight tests were independent, so a *p* value of $.05/8 = .006$ was chosen for significance to avoid Type I error. Nine tests had performance that was significantly better than chance (see Table 4). Musicians' accuracy for just-tuned complex-tone dyads classified as MC, MD, SCT, and sct was especially robust after long retention. For nonmusicians under the same conditions, high retention accuracy was observed for MC, mc, SCT, sct, and rct dyads.

Insert Table 4 About Here.

Discussion

A few general trends reported in the pure-tone dyad experiment (see Chapter 3) were reproduced with complex-tone dyads but within specific dyad classifications the results were inconsistent. Consistent results include the observation that recognition accuracy decreased with time, although performance for more than half of the dyads was above chance at even the longest retention period, again significantly extending previous estimates of the duration of auditory short-term memory. Musicians consistently displayed a narrower range of recognition scores (i.e., less differential memory) among classes of dyads compared to the range of scores for nonmusicians (see Table 1). As they did with pure-tone dyads, nonmusicians' recognition scores for some complex-tone dyads dipped at 24.25 s and showed subsequent improvement over time. In contrast with the pure-tone results however, this pattern was shown for cognitively consonant (MC, mc) rather than cognitively dissonant (MD) complex-tone dyads. Regardless of dyad class, the replication of the phenomenon suggests that improved recognition accuracy with increasing delay for some classes of sounds is not merely an artifact of the experimental design.

The current experiment aimed to extend the findings of the pure-tone dyad data by adding harmonics to the stimuli. This additional information present in these dyads did not have the anticipated effect of making the task easier for nonmusicians. The musicians' mean score was identical for pure- and complex-tone just-tuned dyads and as in the earlier experiment, musicians showed higher d' indices than nonmusicians. The complex-tone data only partially supported the hypothesis that recognition accuracy would be higher for dissonant than for consonant dyads. Cognitive and sensory dissonance assisted dyad recognition at moderate retention periods only; at the longest retention period recognition of MD and RCT dyads was not particularly accurate. As seen with pure-tone dyads, nonmusicians were more sensitive (higher d') to small-integer than to large-integer ratio complex-tone dyads. Greater familiarity with consonant over dissonant intervals is presumed to account for nonmusicians' higher acuity for consonant dyads. Similarly, musicians' greater exposure to musical intervals in general probably contributed to the higher d' indices that this group displayed in both experiments. Nonmusicians are known to be slower than musicians at determining the C/D status of chords (Regnault et al., 2001; Zatorre & Halpern, 1979). Decreased sensitivity or d' is linked to slower response times and diminished confidence in decision-making (Yonelinas, 2001). It appears that any additional time nonmusicians spent processing dissonant dyads was only detrimental to STM when recognition occurred after very long delays.

Under the assumption that STM for dyads is influenced by explicit knowledge and musical experience, we presented unfamiliar musical intervals to listeners to examine the role of musical training. For the third experiment in this series (the second of two reported here), a set of microtuned, complex-tone dyads based on the quartertone scale was used under the same STM paradigm used for just-tuned dyads. Crowder (1993) noted that a pre-categorical memory store should reflect sensitivity to the sensory attributes of objects and exhibit insensitivity to their conceptual attributes. Microtuned dyads could be perceived as being more distinct sensorily than cognitively, due to listeners' relative lack of exposure to them. Conceptual knowledge of microtuned dyads will be absent or at least reduced, so STM for classes of cognitive dissonance could be less distinct. Any processing advantages linked to dyad classifications, such as additional neural resources for musical dissonance or heightened sensitivity to "natural intervals," should be absent. Thus STM for microtuned dyads should show a narrower range of recognition scores among classes (i.e., less differential memory) than seen in the common dyad experiments. If the pattern of recognition accuracy for microtuned dyads is similar to that reported for common dyads, where certain retention periods (e.g., 24.25 s) elicit differential accuracy in recognition, it would suggest that retrieval of stored categorical exemplars was not critical to the recognition pattern observed so far.

Experiment 2: Microtuned dyads: Method

Participants

Participants (*N* = 30; 10 men and 20 women; 19-55 years; *M* = 26, *SD* = 8.8) were recruited from a classified ad, or were volunteers from the Schulich School of Music and Psychology Department at McGill University. Recruits were paid $10 for their time, and volunteers served without pay. Three of the participants (2 musicians, 1 nonmusician) served in Experiment 1, but the two experiments were conducted three months apart. Fifteen participants (musician group) had nine or more years of formal musical training (*M* = 16, *SD* = 5.3); the remaining 15 (nonmusician group) had three or fewer years of music training (*M* = 1.6, *SD* = 1.3). All had normal hearing and none had absolute pitch perception by self-report. Musical training and listening habits were assessed by the modified Queen's University Music Questionnaire (Cuddy et al*.,* 2005). Persons accustomed to listening to music that included microtuned intervals, such as Indian and Arabic music, were excluded from the study.

After completing the task one person (a nonmusician) reported that he used a number counting technique rather than auditory memory to determine his responses. His final score was greater than two standard deviations above the mean for nonmusicians and so his data are not reported here.

Stimuli

The same sample of an alto saxophone playing the note D#4 used in Experiment 1 was used to create 72 microtuned dyads. Analog to digital conversion, pitch shifting, and the combining of notes were accomplished using the same equipment and procedure. As in Experiment 1, root notes (lower frequencies — $f1$) ranged from C3 (130.8 Hz) to B4 (493.8 Hz). Intervals were assigned to the root notes by random assignment. The top notes (upper frequencies — $f2$) ranged from D3+ (151.1 Hz) to A5+ (905.8 Hz). (A note or interval augmented by a quarter-tone — 2.9% in frequency — is indicated with a "+".) Dyads were 500 ms in duration, including 10-ms raised cosine onset and offset ramps and were presented to listeners at 57 $\pm$ 0.75 dBA SPL at the same headphone as measured with the same equipment described in Experiment 1.

The 72 microtuned dyads were composed of altered versions of 12 musical intervals of the Western chromatic scale: unison (uni+), minor 2nd (m2+), Major 2nd (M2+), minor 3rd (m3+), Major 3rd (M3+), perfect 4th (p4+), tritone (tt+), perfect 5th (p5+), minor 6th (m6+), Major 6th (M6+), minor 7th (m7+), and Major 7th (M7+), where the upper notes of the intervals were augmented by a single quartertone.

The frequency-ratio relationships were calculated by multiplying $2^{n/24}$ (where n represents an odd integer from 1 to 23) by whole numbers until the closest integer-ratio relationship was derived. For example, the lower frequency of the quartertone dyad uni+ is related to its upper frequency by a number that corresponds to 2 raised to 1/24, or 1.029 (a semitone dyad has a ratio of 1.059, the 12th root of 2). A distance of 21 quartertones or 1.834, for example, equals the ratio of the m7+ interval. By multiplying 1.834 by successive whole numbers, a whole number product is eventually derived describing the integer-ratio relationship of the m7+ in whole numbers. (For example, $1.834 \times 6 = 11.004$, so a ratio of 11:6 was assigned to the m7+. Each interval's ratio was determined once the closest whole integer, after rounding off at the hundredths place, was found.) When considered in terms of frequency-ratio complexity, all of the microtuned dyads used here were notably dissonant; however, some had frequency-ratio relationships that were less complex than others (see Appendix B for a complete description).

Although the distinctions between sensory and cognitive C/D — both within and across types — were less defined for microtuned than for just-tuned dyads, classifications along frequency-ratio and roughness ratings were assigned for continuity in the experimental protocol and to assist with data analysis. The 72 dyads were each assigned to four levels of frequency ratio complexity from the simplest ratio (the least dissonant of this set) to the most complex as follows: "D1" (m7+, p4+, m3+), "D2" (m2+, tri+, M2+), "D3" (p5+, M3+, uni+), and "D4" (M6+, m6+, M7+). In terms of frequency-ratio complexity, the majority, but not all, of the microtuned intervals were more dissonant than the just-tuned ratios (see Table B3). The 72 dyads were also assigned sensory C/D labels based on each dyad's mean roughness rating obtained from 30 participants (15 musicians, 15 nonmusicians) in a separate experiment (see Chapter 2). The four classes were labeled from the "smoothest" to the "roughest" as follows: smooth microtuned or "SMT," somewhat smooth microtuned or "smt," somewhat rough microtuned or "rmt," and rough microtuned or "RMT." It must be noted that the roughness ratings reported in Chapter 2 were assigned exclusively within each stimulus set. Therefore, the relative roughness of micro- and just-tuned dyads cannot be compared across stimulus sets.

Apparatus, procedure, and data analysis

Testing used the same apparatus and procedure, and data analysis followed the same methods described in Experiment 1. Appendix C lists the arrangement of dyads in blocks for Experiment 2.

Results

Overall performance and comparison of musicians and nonmusicians.

Short-term memory for microtuned dyads was essentially the same as STM for just-tuned complex-tone dyads in terms of overall accuracy, $t(29) = 0.31$, $p = .76$. The average score for microtuned dyads was nearly identical to that for common dyads. Participants averaged 76.7% correct in the task overall and as with just-tuned dyads the mean percent correct in familiar trials was significantly higher than in novel trials ($M_{familiar}$ = 0.82, *SD* = 0.07; M_{novel} = 0.72, *SD* = 0.09), $t(29) = 4.84$, $p < .001$. The hypothesis that microtuned dyads would show a narrower distribution of recognition scores across classifications was not supported. The grand means for musicians and nonmusicians over five retention periods, corrected for guessing, were within 1% of the just-tuned, complex-tone result (see Table 5). At some retention periods musicians' scores covered a wider range than nonmusicians' scores. An unexpected finding was that the average *d'* indices —indicators of task difficulty — dropped only slightly compared to the average *d'* for just-tuned intervals (musicians only). Nonmusicians' *d'* indices were midway between their values for just-tuned pure- and complex-tone dyads (cf. Table 1; Table 2, Chapter 3). The response criteria *c* were not different for microtuned dyads but this was unremarkable because biases toward responding "yes" or "no" should not have changed, given that test conditions were the same for all experiments. For microtuned dyads, *c* ranged from 0.09 (musicians, smt) to –0.27 (nonmusicians, SMT) indicating that nonmusicians had a moderately low propensity for regarding the perceptually "smoothest" class of dyads as familiar. Musicians, too, displayed this familiar bias toward smoother dyads, as shown in Table 5. As in the previous experiments, the Mann-Whitney *U* test indicated that participants with musical training were significantly more sensitive to novel/familiar status than were those with the least training, $U = 45.00$, $Z = 2.80$, $p < .01$.

Insert Table 5 About Here.

The effect of retention period.

Recognition period duration had the most predictive power on microtuned dyad recognition score, followed by degree of roughness (4 levels: SMT, smt, rmt, RMT) and block order (first, second, last). The variables of frequency-ratio relationship (4 levels: D1, D2, D3, D4), root note, and top note were not significant predictors of correct responding as shown in a binary logistic regression procedure (forward, stepwise). The dependent variable was the response score (correct/incorrect) on familiar trials. Retention period emerged as a significant predictor with an unstandardized logistic coefficient (*B*) of –0.30 ($p < .001$). Degree of roughness ($B = –0.11$, $p = .03$) and block order ($B = –0.15$, $p = .03$) were also significant. A model with these variables, however, did not significantly improve the null (intercept-only) model. A Hosmer-Lemeshow goodness-of-fit test indicated that the data did not fit the three-variable model well, $\chi^2(8, N = 2160) = 83.08$, $p < .001$. The Nagelkerke R^2 showed that only 8% of the variance in the data was explained by the full model with retention period, degree of roughness, and block order as predictors.

The effects of consonance and dissonance.

Microtuned dyad recognition over time showed some of the behavior observed with just-tuned dyads. Musicians outperformed nonmusicians in the eight dyad classifications after averaging the corrected scores across the five retention periods of interest, as seen in Table 5. Two three-way (2 × 5 x 4) repeated-measures mixed-design ANOVAs evaluated the between-subjects factor of Expertise (2 levels: musician and nonmusician) and two within-subject factors, Retention (5 levels: 16.00, 24.25, 32.50, 40.75, and 48.00 s) and C/D (4 levels in each test: D1, D2, D3, and D4, *or* SMT, smt, rmt, and RMT). The two ANOVAs each revealed a significant effect of Expertise: ANOVA with Frequency-Ratio C/D, $F(1, 28) = 5.24$, $p = .03$; ANOVA with Sensory C/D, $F(1, 28) = 6.54$, $p = .02$. There were no significant interactions of Expertise with either Retention and/or C/D. Subsequent tests of the effects of C/D examined musician and nonmusician data separately. The following tests used two-way (5 × 4) repeated-measures ANOVAs with Retention and C/D as within-subject independent variables and p(c)* as the dependent variable.

Frequency-ratio C/D. Musicians' STM for microtuned dyads classified by frequency-ratio relationship was differential at two critical retention periods — 24.25 and 48.00 s (see Fig. 5). Musicians recognized microtuned, complex-tone dyads similarly to how nonmusicians recognized just-tuned, complex-tone dyads, as distinguished by frequency-ratio complexity (cf Fig. 2, Fig. 5). Recognition accuracy for dyads in the simplest frequency-ratio relationships (D1, D2) approached chance at 24.25 s but improved at longer retention periods (see Fig. 5). The repeated-measures ANOVA for musicians and frequency-ratio C/D reported a significant effect of Retention, $F(4, 56) = 5.15$, $p = .001$, and Frequency-ratio C/D, $F(3, 42) = 4.63$, $p < .01$. The Frequency-ratio C/D × Retention interaction effect was significant, $F(12, 168) = 3.03$, $p = .001$. Following the significant interaction, the simple main effect of Frequency-ratio C/D was explored at each of the five retention periods. Musicians showed significant differential scores at 24.25 s, $F(3, 42) = 5.75$, $\varepsilon = .75$, $p < .01$, and 48.00 s, $F(3, 42) = 6.41$, $\varepsilon = .58$, $p < .01$. Post hoc comparisons of the means at these retention periods used a Bonferroni-corrected alpha of $p = .05/6 = .008$ for significance. At 24.00 s, dyads classified as D3 were recognized significantly more accurately than D2 dyads ($p = .001$) and were marginally better recognized than the D1 class ($p = .011$) under the conservative alpha. At 48.00 s, the D3 dyad class was again recognized significantly more often than the D2 dyad class ($p = .001$).

For the first time in these experiments, nonmusicians did not show differential STM at 24.25 s for dyads classified by frequency-ratio complexity. Differential memory appeared at long delay; nonmusicians displayed accurate STM for microtuned dyads with the most complex frequency-ratios (see Fig. 6). The repeated-measures ANOVA for nonmusicians and frequency-ratio C/D found a significant effect of Retention, $F(4, 56) = 7.40$, $p < .001$, and a significant Frequency-ratio C/D × Retention interaction, $F(12, 168) = 2.14$, $p = .02$. The effect of Frequency-ratio was not significant, $F(3, 42) = 1.54$, $p = .22$. Follow-up tests looked at the simple main effect of frequency-ratio C/D at each of five retention periods. Significant effects were observed at 40.75 s, $F(3, 42) = 2.99$, $p < .05$, and 48.00 s, $F(3, 42) = 3.55$, $p = .02$. Pairwise comparisons of the means at 40.75 s used a Tukey HSD and showed that dyads classified as D2 were recognized significantly more often than D1 dyads. At 48.00 s using a Tukey HSD, nonmusicians recognized dyads classified as D3 and D4 significantly more often D2 dyads.

Insert Figure 5 About Here.

Insert Figure 6 About Here.

Sensory C/D. Microtuned dyads distinguished by sensory C/D showed differential STM recognition by musicians, but only after moderate retention durations (see Fig. 7). The somewhat smooth dyads (smt) performed nearly at chance at 24.25 s retention, but were the most accurately recognized class at 48.00 s. The ANOVA for musicians and Sensory C/D found a significant effect of Retention, $F(4, 56) = 5.79$, $p = .001$, and a significant C/D × Retention interaction, $F(12, 168) = 2.68$, $p < .01$. The main effect of Sensory C/D approached significance, $F(3, 42) = 2.40$, $p = 0.08$. Simple main effects of Sensory C/D at five retention periods found a significant difference at 24.25 s, $F(3, 42) = 6.44$, $p = .001$. Pairwise comparisons of the means at this period used Tukey's HSD and showed that both SMT and rmt dyad classes were recognized significantly better than smt dyads by musicians at 24.25 s.

Nonmusicians also showed poor recognition for the smt dyad class at 24.25 s, followed by improved recognition as retention periods increased (see Fig. 8). "Somewhat smooth" dyads (sct, smt) dipped and improved for nonmusicians in both experiments reported here, and in the earlier experiment with pure-tone (sc) dyads (cf Fig. 4, Fig. 8; see also Fig. 4, Chapter 3). The ANOVA results for nonmusicians and Sensory C/D showed a significant effect of Retention, $F(4, 56) = 6.60$, $p < .001$, and Sensory C/D, $F(3, 42) = 4.04$, $p = .01$. The C/D × Retention interaction was significant, $F(12, 168) = 3.28$, $p = .001$. The simple main effects of Sensory C/D at five retention periods were tested following the significant interaction. Two tests reached the 0.05 significance level: 24.25 s, $F(3, 42) = 3.62$, $p < .05$, and 48.00 s, $F(3, 42) = 6.33$, $p = .001$. Pairwise comparisons of the means for each significant test were conducted using Tukey's HSD. Significant pairwise differences at 24.25 s were found for SMT and smt dyads. At 48.00 s, both smt and rmt dyads were recognized significantly more often than RMT dyads.

Insert Figure 7 About Here.

Insert Figure 8 About Here.

Shapiro-Wilk tests of normality on the microtuned dyad data revealed that as with the Experiment 1 data, roughly one quarter of the score distributions at each retention period deviated significantly from normality. Because of these nonnormal distributions, and to supplement the parametric ANOVA findings, conservative Friedman nonparametric ANOVAs were performed as in Experiment 1. The Friedman tests showed agreement with the parametric ANOVAs, revealing differential STM at 24.25 and 48.00 s. Six tests were significant at $p \leq .05$ (see Table 6). Follow-up pairwise comparisons on these six used Wilcoxon tests and controlled for Type I errors with a Bonferroni correction; $p = .05/6 = .008$ was required for significance. Significant pairwise differences were observed at 24.25 and 48.00 s in both participant groups.

Insert Table 6 About Here.

The effects of secondary variables.

The effect of root pitch on STM was more dramatic for microtuned than for just-tuned dyads, meaning that the range of scores was broader (see Table 7). Sixty-two out of 276 root-note pairs showed significant differences between their mean recognition scores. A one-way repeated-measures ANOVA was performed on root pitch (24 levels) using the number of correct recognitions as the dependent variable ($N = 30$). The effect of root pitch was significant, $F(23, 667) = 9.68$, $\varepsilon = .88$, $p < .001$. Follow-up pairwise comparisons used a Bonferroni corrected alpha of $p = .05/276 = .0002$ to control for Type I error. Mean correct scores by root note showed a broad range, from note D3 ($M = 5.4$, $SD = 0.62$) to note A3 ($M = 3.4$, $SD = 1.1$). As seen with just-tuned intervals, higher-pitched dyads (C4 to B4) were recognized significantly more often than those with lower pitches (C3 to B3), $t(29) = 4.59$, $p < .001$.

The three stimulus blocks were equally challenging, $F(2, 58) = 1.80$, $p = .18$. The means and standard deviations out of a maximum score of 48 correct were as follows: $M_{Blk1} = 37$, $SD = 3.0$; $M_{Blk2} = 37$, $SD = 3.2$; $M_{Blk3} = 36$, $SD = 3.2$. The order of presentation (first, second, or last) of the stimulus blocks did not have a significant effect on recognition accuracy, $F(2, 58) = 0.66$, $p = .936$. The means and standard deviations for first, second, and last blocks were: $M_{First} = 37$, $SD = 3.0$; $M_{Second} = 37$, $SD = 3.1$; $M_{Last} = 37$, $SD = 3.5$.

Insert Table 7 About Here.

Tests of auditory memory duration.

Familiar recognition of many microtuned dyads was accurate above chance performance at 48.00 s of retention, despite the heavy cognitive load. Nine of 16 paired-sample *t* tests showed better-than-chance (0.50) microtuned dyad recognition accuracy. Musicians accounted for the majority of the high scores at the longest retention period, recognizing all but D2 and RMT dyad classes better than chance at $p = .05/8 = .006$ (see Table 8). Nonmusicians recognized the smt and rmt dyads significantly better than chance at 48.00 s, but the D3 class approached significance under the conservative alpha.

Insert Table 8 About Here.

Discussion

Short-term memory for unfamiliar, microtuned dyads is as robust and accurate as STM for common, just-tuned dyads. This suggests that familiar dyad recognition after a single, novel presentation is not necessarily dependent on access to categorized exemplars in long-term storage. The results support the suggestion that STM for auditory events is automatic (Demany & Semal, 2007) and does not depend on conscious "knowing." Although some level of musical expertise correlates with improved dyad recognition memory, those without musical training can accurately recognize dyads on the basis of familiarity for items in the present context. In this way the novel/familiar paradigm is shown to tap a signal detection process and therefore the SDT analyses can be considered good indicators of participant confidence and sensitivity (Yonelinas, 2001). Short-term memory processing of unusual dyads is determined to be no more taxing than it is for typical dyads, when all other conditions are equal.

Neither the frequency-ratio complexity nor the degree of roughness of microtuned dyads mediated STM processes in a systematic manner but there were some observations that were consistent with just-tuned dyad STM findings. Microtuned dyads labeled along indices of C/D (more precisely indices of relative dissonance) were differentially recognized after moderate retention periods but similarly recognized at longer retention periods. This observation over the course of three experiments with three unique stimulus sets and populations argues in favor of a phenomenological rather than a methodological explanation. The strategy or mechanism that serves STM recognition under heavy cognitive load (long delay plus relatively many interfering items) may be ineffective or under-utilized under less demanding conditions, for certain stimuli.

Whether it is a cognitive or sensory property of a dyad that promotes best recognition at optimal retention periods awaits further testing using finer class distinctions among stimuli. For example, musicians showed a similar temporal recognition pattern for the least frequency-ratio complex microtuned dyad class (D1— m7+, p4+, m3+ dyads) and the “somewhat smooth” (smt) class (see Figs. 5 and 7). With only two exceptions, there was no overlap between intervals in the D1 class and those in the “smt” class, which consisted of mostly m6+, M6+, and p5+ dyads. The similar recognition pattern for these two (more or less) nonoverlapping C/D categories may underscore the dual nature of consonance and dissonance, complicating easy conclusions about the percept. *Based on these categorizations*, we cannot determine whether it was a microtuned dyad’s particular frequency-ratio relationship or its unique sensory traits that drove musicians’ poor recognition of it at 24.25 s. We can tentatively conclude that performance was driven by a property of the *stimulus set*, because musicians did not have difficulty recognizing just-tuned complex-tone dyads at 24.25 s (see Fig. 3), although nonmusicians did (see Fig. 4).

The likelihood that participants processed dyads as known intervals and recognized them from experience was no doubt lessened in the microtuned dyad experiment. Nevertheless, microtuned dyads did not reduce the difference between musicians' and nonmusicians' recognition performance; musicians' STM advantage over nonmusicians was virtually identical to that seen in the just-tuned dyad experiments. This suggests that the reported higher acuity of musical experts for processing auditory signals (Brattico et al*.,* 2009; Kreiman, Gerratt, & Berke, 1994; Lee et al., 2009; Magne et al., 2006; Neuhaus et al., 2006; Schön, Regnault, Ystad, & Besson, 2005; Strait et al., 2010; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005; Zatorre & Halpern, 1979) supports accurate auditory STM. During informal, post-experiment discussions musicians frequently reported that they tried to identify the microtuned dyads in comparison with intervals they knew. It was not uncommon to hear, "I knew I heard something between a tritone and a perfect 5th, but wasn't sure when it came around again." Presumably nonmusicians had a diminished capacity to compare microtuned dyads' active traces in STM against stored exemplars but this relative incapacity was not severely detrimental to their recognition scores, suggesting that a pre-categorical memory store mediated recognition. Such a memory store should be more sensitive to the sensory rather than to the conceptual attributes of objects (Goldstone & Barsalou, 1998; Jacoby, 1983) and so sensory dissonance was predicted to account for accurate STM where it was observed. The roughest microtuned dyads were not recognized any more accurately than their smoother counterparts, however. Nonmusicians showed that the roughest dyads in this stimulus set were their least well recognized (Table 5, Fig. 8). This result suggests that if a pre-categorical memory store was used in the microtuned dyad recognition task, some property other than roughness mediated the processing.

General Discussion

In summary, neither sensory nor cognitive (knowledge-based) distinctions along the continua of consonance versus dissonance mediate STM for dyads in a clear or systematic way. To a certain extent however, listeners in this series of experiments responded similarly to a few distinctions among dyads and these are summarized here:

> Differential STM for classes of dyads at 24.25 s retention favors dissonance — both frequency-ratio and frequency-separation — over consonance.
>
> Cognitive dissonance (MD) is the best-recognized category by musicians for just-tuned complex-tone dyads. Sensory dissonance (SD) is the best-recognized category by musicians for pure-tone dyads.
>
> Some (relative) categories of pure-tone and microtuned dyads perform similarly to each other over time, but differently than their counterparts in just-tuned complex-tone dyads. Examples include moderate frequency-ratio dissonance (md, D3) for musicians and moderate frequency-ratio consonance (mc, D2), somewhat sensory smooth (sc, smt), and sensory rough (SD, RMT) dyads for nonmusicians. This observation gives modest support to evidence from evaluated auditory roughness studies (see Chapter 2) suggesting that nonmusicians regard pure-tone and microtuned dyads as similarly "foreign," compared to just-tuned complex-tone dyads.
>
> Retention period duration has less of a mediating influence on STM for dyads in musicians than it does for nonmusicians.

The aim of these experiments was to determine the attributes of dyads that contribute to robust STM but the results defied a simple description of the relationship among the effects of retention period duration, the number of intervening items, and musical interval properties. It was assumed that either sensory or cognitive distinctions among sounds would drive differential cognitive processing, which would in turn be reflected in STM performance. Differential STM for dyads was observed but only under certain conditions and in a nonsystematic manner. The cumulative result of three experiments in this chapter and the previous chapter suggests that auditory STM for dyads is more robust than expected — resilient to at least five interfering items during 48.00 s of retention. Auditory STM as a function of auditory features is more complex than anticipated; multiple recognition strategies may be used, each with an optimal time course for maximum efficiency, and stimulus distinctiveness may influence the strategies.

If not attributes of consonance or dissonance, what accounted for the observed results? As noted, in the absence of conscious "knowing," when perceptual fluency is the best option for processing, the similarity of sequential items is known to constrain memory capacity and influence the hit and miss rates (Deutsch, 1972a; Goldstone, 1994). After accounting for number of years of musical training and thus the likelihood of having explicit categorical knowledge of intervals, it would be difficult to predict how a group of participants would gauge the similarity among dyads. Determining how perceptual similarities influenced recognition thus presented a challenge to interpreting the present results.

This challenge is illustrated by the difference in the correct response rates for two just-tuned complex-tone perfect 4ths — dyads 'Ap' (D#4 p4) and 'NN' (C#4 p4). Consider the following two subsequences from Experiment 1, arranged such that the oldest item (the first presented) is on the far left:

Seq. 1	D#4	D#3	B4	F#4	B4	D#4
	p4	oct.	p5	M6	p5	p4
Seq. 2	C#4	F3	B4	C#3	B3	C#4
	p4	m7	m7	M2	M6	p4

The familiar presentation of both p4s occurred 40.75 s after their novel presentations. Only 16 of 30 participants, or 53%, correctly recognized 'Ap' on its familiar trial. In contrast, 25 of 30 participants, or 83%, correctly recognized 'NN.' A close examination of each subsequence does not reveal an exclusive explanation for the wide discrepancy between recognition scores for the two familiar p4s. Was the D#3 octave *immediately following* the novel 'Ap' perceptually similar to the D#4 p4, thus weakening or disrupting an ongoing memory consolidation and increasing the rate of forgetting? The identical pitch chroma of the two D#s may have been responsible for memory disruption. Another possibility is that the familiar B4 p5 *immediately preceding* the familiar 'Ap' (plus the novel B4 p5 also falling in the intervening retention interval for Ap) had the biggest influence on the miss rate for the familiar dyad. The p4 and p5 might have caused confusion based on interval similarity (both are cognitively consonant) and led to recognition error. Or the recognition of the B4 p5 from a shorter retention interval may have perturbed that of the D#4 p4 over a longer interval.

Competing models of the effects of stimulus similarity present an opportunity for auditory STM research to make a contribution to this field of study (Creel, Newport, & Aslin, 2004; Goldstone, 1994; Hintzman, 1988; Johns & Mewhort, 2002; Kishiyama & Yonelinas, 2003; Nosofsky & Zaki, 2003; Stewart & Brown, 2004, 2005). The dyads used in the present study were rich in terms of stimulus attributes and separating the influence of each sensory and cognitive distinction will take dedicated work. What can be concluded is that the present result is the product of the temporal rate of forgetting (possibly not equivalent depending upon stimulus features), interference from previous items (where stimulus distinctiveness plays a role) and the error component from internal neural noise. The organizing principles underlying auditory STM for musical intervals must be examined from the perspective of each of these influences.

Conclusion

This series of experiments used an STM task to explore those attributes of dyads that contribute positively or negatively to their robustness against interference and decay. We used the concept of "natural intervals" or consonance versus dissonance to define and distinguish the stimuli. Auditory STM for dyads was found to persist longer than previously reported for single tones (Winkler et al., 2002), regardless of dyad classification. Robust STM was observed even when categorized exemplars from LTM were presumed absent, as was the case for microtuned dyads. Although differential memory for some classes of dyads was observed at moderate retention periods, the effect disappeared at longer retention periods. This phenomenon was observed in three separate experiments but was inconsistent with regard to dyad classification and type. The study's limitation was that the distribution of the sensory and cognitive properties of dyads was too broad to allow induction of a general principle of strong or weak STM performance. We concur with Burns and Houtsma (1999) that small-integer ratio dyads (or "natural intervals") convey no innate processing advantage because musicians' and nonmusicians' STM for these dyads was no better or worse than for other dyad classes.

These experiments also asked if and how musical training mediates auditory STM for musical intervals. Although musicians made fewer recognition errors and displayed higher confidence than nonmusicians in distinguishing novel from familiar trials, the difference found in these experiments was not so large as to declare that musical training develops a kind of superior auditory STM that might confer a practical advantage.

These findings contribute to the study of auditory signal processing by mapping the fate of musical interval memory traces over time. Future work should focus on whether encoding errors or signal degradation due to time and interference accounts for the observed pattern of recognition failures. If dyads in the stimulus set are encoded with the same fidelity and decay at the same rate, then accurate memory depends on interference from similar items. It cannot be assumed, however, that the rate of forgetting is the same for all stimuli and participants (Tierney & Pisoni, 2004; Wicklegren, 1977). Visual STM was recently shown to exhibit a deterministic rather than random decay pattern (Gold, Murray, Sekuler, Bennett, & Sekuler, 2005). Future work in audition could reveal a similar deterministic decay mechanism in auditory STM. Certain features of an auditory signal may be more prone to rapid degradation than others. Discovering these memory mechanisms and patterns will surely satisfy

Crowder's (1993) prediction that a variety of experimental techniques must be used to thoroughly explore the topic of auditory STM.

Acknowledgements

The work reported herein was submitted in partial fulfillment of the requirements for the Ph.D. in Psychology at McGill University by SER. We gratefully acknowledge the technical assistance of Karle-Philip Zamor and Bennett Smith, the support of grants from NSERC and Bell Canada to DJL, grants from NSERC and the Canada Research Chair program to SMc, and a CIRMMT Student Award and Canadian Acoustic Association Award to SER. Al Bregman, Norman Cook, David Huron, and Eugene Narmour provided valuable comments on this work and/or on earlier drafts of the paper. Richard Pastore offered helpful insights on applying detection theory methods to the analysis. Portions of this work were presented at the Neurosciences and Music III conference, Montréal, QC, June 2008.

Appendix A: Just-tuned Intervals

Assigning sensory consonance/dissonance classes

Sensory C/D levels were assigned to dyads on the basis of listener evaluations of their psychoacoustic roughness in a separate experiment (see Chapter 2). Thirty participants (15 musicians; 15 nonmusicians) positioned a sliding scale at a point between 0 and 1 that represented each dyad's perceived quality — "smooth" on one extreme and "rough" on the other. Before the task participants were familiarized with exemplars of dyads with generally accepted rough and smooth envelopes (e.g., m2, M2, octave, p5). Listeners were encouraged to attend to just the sensation produced by the signal's envelope and ignore musical qualities (euphoniousness, harmony) and pitch height. There were some differences between musician and nonmusician evaluations. Musicians tended to rate small-integer ratio dyads as much smoother than large-integer ratio dyads, whereas nonmusicians rated these more similarly. The ratings were combined, averaged, and used to assign a single roughness value to each dyad for purposes of the current experiment.

The evaluated roughness rating is listed; higher values represent greater roughness (Table A1, col. 8, "eval. rough."). It is noted that these sensory C/D ratings are both subjective and relative to the dyads in this stimulus set. More objective sensory C/D ratings of these dyads were obtained from two models of psychoacoustic roughness. These results are described in Chapter 2.

Assigning cognitive consonance/dissonance classes

Schellenberg and Trehub (1994b) adopted a previously used technique of quantifying the simplicity of dyad frequency ratios to derive a cognitive C/D rating (Table A.1., cols. 10, 11, 12). The reciprocal of the natural logarithm of the sum of a dyad's integers, X and Y, is taken: $[\ln(X + Y)]^{-1}$. The frequency ratios of dyads used here correspond to the *just tuned* scale, where instruments are tuned so that all notes of the scale are related by whole integers. This index was used to assign four levels of cognitive C/D, 18 dyads each, to the 72 dyads in the present experiment (Table B3).

Table A1

Experiment 1: Stimulus blocks for just-tuned complex-tone dyads.

Block 1.

dyad	mus. intvl.	lower note	upper note	f1 Hz	f2 Hz	mean freq.	eval. rough.	sens. C/D	freq. ratio	$\ln(X+Y)^{-1}$	cogn. C/D
AA	m6	C3	G#3	130.8	207.6	169.2	0.567	rct	5:8	0.390	md
BB	M2	C#3	D3	138.6	155.6	147.1	0.711	RCT	8:9	0.353	md
CC	oct	D3	D4	146.8	293.6	220.2	0.300	SCT	1:2	0.910	MC
DD	tri	D#3	A3	155.6	220.0	187.8	0.588	rct	32:45	0.230	MD
EE	M3	E3	G#3	164.8	207.6	186.2	0.445	sct	4:5	0.455	mc
FF	m7	F3	D#4	174.6	311.2	242.9	0.531	rct	9:16	0.311	MD
GG	p5	F#3	C#4	185.0	277.2	231.1	0.429	SCT	2:3	0.621	MC
HH	p4	G3	C4	196.0	261.6	228.8	0.520	rct	3:4	0.514	MC
II	m3	G#3	B3	207.6	246.9	227.3	0.446	sct	5:6	0.417	mc
JJ	m2	A3	A#3	220.0	233.1	226.6	0.720	RCT	15:16	0.291	MD
KK	M7	A#3	A4	233.1	440.0	336.6	0.604	rct	8:15	0.319	md
LL	M6	B3	G#4	246.9	415.2	331.1	0.499	sct	3:5	0.481	mc
MM	oct	C4	C5	261.6	523.5	392.6	0.303	SCT	1:2	0.910	MC
NN	p4	C#4	F#4	277.2	370.0	323.6	0.501	rct	3:4	0.514	MC
OO	M7	D4	C#5	293.6	554.4	424.0	0.708	RCT	8:15	0.319	md
PP	m6	D#4	B4	311.2	493.8	402.5	0.471	sct	5:8	0.390	md
QQ	m2	E4	F4	329.6	349.2	339.4	0.741	RCT	15:16	0.291	MD
RR	tri	F4	B4	349.2	493.8	421.5	0.491	sct	32:45	0.230	MD
SS	m3	F#4	A4	370.0	440.0	405.0	0.437	SCT	5:6	0.417	mc
TT	M3	G4	B4	392.0	493.8	442.9	0.456	sct	4:5	0.455	mc
UU	M2	G#4	A#4	415.2	466.2	440.7	0.686	RCT	8:9	0.353	md
VV	p5	A4	E5	440.0	659.2	549.6	0.400	SCT	2:3	0.621	MC
WW	M6	A#4	G5	466.2	784.0	625.1	0.441	SCT	3:5	0.481	mc
XX	m7	B4	A5	493.8	880.0	686.9	0.598	rct	9:16	0.311	MD

Block 2.

dyad	mus. intvl.	lower note	upper note	f1 Hz	f2 Hz	mean freq.	eval. rough.	sens. C/D	freq. ratio	$\ln(X+Y)^{-1}$	cogn. C/D
Aa	m3	C3	D#3	130.8	155.6	143.2	0.580	rct	5:6	0.417	mc
Ab	tri	C#3	G3	138.6	196.0	167.3	0.611	rct	32:45	0.230	MD
Ac	p5	D3	A3	146.8	220.0	183.4	0.483	sct	2:3	0.621	MC
Ad	oct	D#3	D#4	155.6	311.2	233.4	0.329	SCT	1:2	0.910	MC
Ae	M2	E3	F#3	164.8	185.0	174.9	0.690	RCT	8:9	0.353	md
Af	M7	F3	E4	174.6	329.6	252.1	0.654	RCT	8:15	0.319	md
Ag	p4	F#3	B3	185.0	246.9	216.0	0.415	SCT	3:4	0.514	MC
Ah	m7	G3	F4	196.0	349.2	272.6	0.560	rct	9:16	0.311	MD
Ai	m2	G#3	A3	207.6	220.0	213.8	0.758	RCT	15:16	0.291	MD
Aj	M6	A3	F#4	220.0	370.0	295.0	0.510	rct	3:5	0.481	mc
Ak	M3	A#3	D4	233.1	293.6	263.4	0.457	sct	4:5	0.455	mc
Al	m6	B3	G4	246.9	392.0	319.5	0.395	SCT	5:8	0.390	md
Am	M7	C4	B4	261.6	493.8	377.7	0.651	RCT	8:15	0.319	md
An	M3	C#4	F4	277.2	349.2	313.2	0.413	SCT	4:5	0.455	mc
Ao	M2	D4	E4	293.6	329.6	311.6	0.626	RCT	8:9	0.353	md
Ap	p4	D#4	G#4	311.2	415.2	363.2	0.493	sct	3:4	0.514	MC
Aq	tri	E4	A#4	329.6	466.2	397.9	0.531	rct	32:45	0.230	MD
Ar	m6	F4	C#5	349.2	554.4	451.8	0.434	SCT	5:8	0.390	md
As	M6	F#4	D#5	370.0	622.3	496.2	0.440	SCT	3:5	0.481	mc
At	m2	G4	G#4	392.0	415.2	403.6	0.831	RCT	15:16	0.291	MD
Au	oct	G#4	G#5	415.2	830.6	622.9	0.272	SCT	1:2	0.910	MC
Av	m3	A4	C5	440.0	523.5	481.8	0.477	sct	5:6	0.417	mc
Aw	m7	A#4	G#5	466.2	830.6	648.4	0.599	rct	9:16	0.311	MD
Ax	p5	B4	F#5	493.8	740.0	616.9	0.389	SCT	2:3	0.621	MC

Block 3.

dyad	mus. intvl.	lower note	upper note	f1 Hz	f2 Hz	mean freq.	eval. rough.	sens. C/D	freq. ratio	$\ln(X+Y)^{-1}$	cogn. C/D
ba	m7	C3	A#3	130.8	233.1	182.0	0.650	RCT	9:16	0.311	MD
bb	M7	C#3	C4	138.6	261.6	200.1	0.680	RCT	8:15	0.319	md
bc	p4	D3	G3	146.8	196.0	171.4	0.516	rct	3:4	0.514	MC
bd	M2	D#3	F3	155.6	174.6	165.1	0.725	RCT	8:9	0.353	md
be	m6	E3	C4	164.8	261.6	213.2	0.469	sct	5:8	0.390	md
bf	m2	F3	F#3	174.6	185.0	179.8	0.789	RCT	15:16	0.291	MD
bg	oct	F#3	F#4	185.0	370.0	277.5	0.271	SCT	1:2	0.910	MC
bh	M6	G3	E4	196.0	329.6	262.8	0.451	sct	3:5	0.481	mc
bi	tri	G#3	D4	207.6	293.6	250.6	0.501	rct	32:45	0.230	MD
bj	M3	A3	C#4	220.0	277.2	248.6	0.427	SCT	4:5	0.455	mc
bk	p5	A#3	F4	233.1	349.2	291.2	0.467	sct	2:3	0.621	MC
bl	m3	B3	D4	246.9	293.6	270.3	0.447	sct	5:6	0.417	mc
bm	m2	C4	C#4	261.6	277.2	269.4	0.790	RCT	15:16	0.291	MD
bn	m6	C#4	A4	277.2	440.0	358.6	0.484	sct	5:8	0.390	md
bo	m3	D4	F4	293.6	349.2	321.4	0.551	rct	5:6	0.417	mc
bp	m7	D#4	C#5	311.2	554.4	432.8	0.576	rct	9:16	0.311	MD
bq	p5	E4	B4	329.6	493.8	411.7	0.417	SCT	2:3	0.621	MC
br	M3	F4	A4	349.2	440.0	394.6	0.448	sct	4:5	0.455	mc
bs	M7	F#4	F5	370.0	698.4	534.2	0.717	RCT	8:15	0.319	md
bt	M6	G4	E5	392.0	659.2	525.6	0.456	sct	3:5	0.481	mc
bu	p4	G#4	C#5	415.2	554.4	484.8	0.552	rct	3:4	0.514	MC
bv	M2	A4	B4	440.0	493.9	467.0	0.713	RCT	8:9	0.353	md
bw	oct	A#4	A#5	466.2	932.3	699.3	0.235	SCT	1:2	0.910	MC
bx	tri	B4	F5	493.8	698.4	596.1	0.556	rct	32:45	0.230	MD

Table A2

Classification of sensory consonance/dissonance levels based on evaluated roughness (higher is rougher) — Just-tuned complex-tone dyads.

Class	Evaluated C/D	Number of dyads
SCT	< 0.442	18
sct	$0.442 < sc \leq 0.501$	18
rct	$0.500 < sc \leq 0.626$	18
RCT	> 0.625	18

Appendix B: Microtuned Intervals

Assigning sensory consonance/dissonance classes

Sensory C/D levels for microtuned dyads were obtained as described for the just-tuned dyads, using the same apparatus. Participants evaluated the 72 dyads of a single stimulus set per session: in other words just- and microtuned dyads were not evaluated together. The ratings were divided into 4 classes of sensory C/D (see Table B2).

Assigning frequency-ratio consonance/dissonance classes

The frequency-ratio complexity of microtuned dyads was derived using the same formula used for just-tuned dyads — $[\ln (X + Y)]^{-1}$ (see Table B3). Derivation of the frequency ratios for each microtuned interval is described in the text. The table shows that the simplest frequency ratio of the microtuned scale (m7+) has the same complexity as an M2 from the just-tuned scale. The most complex ratio in the just-tuned scale — the tritone — has approximately the equivalent complexity of the m6+ microtuned interval, second from the bottom of the microtuned scale. Therefore nearly all of the microtuned intervals are as musically dissonant as the five most dissonant of the twelve just-tuned intervals.

Table B1

Experiment 2: Stimulus blocks for microtuned complex-tone dyads.

Block 1.

dyad	mus. intvl.	lower note	upper note	f1 Hz	f2 Hz	mean freq.	eval. rough.	sens. C/D	freq. ratio	$\ln(X+Y)^{-1}$	ratio level
AAA	M6+	C3	A3+	130.8	226.4	178.6	0.619	RMT	45:26	0.235	D4
BBB	p4+	C#3	F#3+	138.6	190.4	164.5	0.602	RMT	11:8	0.340	D1
CCC	m3+	D3	F3+	146.8	179.7	163.3	0.577	rmt	11:9	0.334	D1
DDD	m6+	D#3	B3+	155.6	254.1	204.9	0.483	smt	49:30	0.229	D4
EEE	m2+	E3	F3+	164.8	179.7	172.3	0.755	RMT	12:11	0.319	D2
FFF	p5+	F3	C4+	174.6	269.3	221.9	0.495	smt	37:24	0.243	D3
GGG	tri+	F#3	C4+	185.0	269.3	227.1	0.490	smt	16:11	0.303	D2
HHH	M7+	G3	F#4+	196.0	380.8	288.4	0.613	RMT	68:35	0.216	D4
III	uni+	G#3	G#3+	207.6	213.7	210.6	0.587	rmt	35:34	0.236	D3
JJJ	M3+	A3	C#4+	220.0	285.3	252.7	0.439	SMT	35:27	0.242	D3
KKK	m7+	A#3	G#4+	233.1	427.4	330.2	0.659	RMT	11:6	0.353	D1
LLL	M2+	B3	C#4+	246.9	285.3	266.1	0.572	rmt	15:13	0.300	D2
MMM	tri+	C4	F#4+	261.6	380.8	321.2	0.547	rmt	16:11	0.303	D2
NNN	M7+	C#4	C5+	277.2	538.5	407.9	0.554	rmt	68:35	0.216	D4
OOO	p5+	D4	A4+	293.6	452.9	373.2	0.491	smt	37:24	0.243	D3
PPP	m2+	D#4	E4+	311.2	339.3	325.2	0.645	RMT	12:11	0.319	D2
QQQ	m6+	E4	C5+	329.6	538.5	434.1	0.350	SMT	49:30	0.229	D4
RRR	M6+	F4	D5+	349.2	604.4	476.8	0.482	smt	45:26	0.235	D4
SSS	p4+	F#4	B4+	370.0	508.3	439.1	0.420	SMT	11:8	0.340	D1
TTT	M2+	G4	A4+	392.0	452.9	422.4	0.500	smt	15:13	0.300	D2
UUU	M3+	G#4	C5+	415.2	538.5	476.9	0.420	SMT	35:27	0.242	D3
VVV	m3+	A4	C5+	440.0	538.5	489.3	0.386	SMT	11:9	0.334	D1
WWW	uni+	A#4	A#4+	466.2	479.9	473.0	0.511	smt	35:34	0.236	D3
XXX	m7+	B4	A5+	493.8	905.8	699.8	0.532	rmt	11:6	0.353	D1

Block 2.

dyad	mus. intvl.	lower note	upper note	f1 Hz	f2 Hz	mean freq.	eval. rough.	sens. C/D	freq. ratio	$\ln(X+Y)^{-1}$	ratio level
Ba	tri+	C3	F#3+	130.8	190.4	160.6	0.658	RMT	16:11	0.303	D2
Bb	M3+	C#3	F3+	138.6	179.7	159.2	0.554	rmt	35:27	0.242	D3
Bc	uni+	D3	D3+	146.8	151.1	149.0	0.665	RMT	35:34	0.236	D3
Bd	M2+	D#3	F3+	155.6	179.7	167.7	0.669	RMT	15:13	0.300	D2
Be	M6+	E3	C#4+	164.8	285.3	225.1	0.472	smt	45:26	0.235	D4
Bf	m6+	F3	C#4+	174.6	285.3	230.0	0.434	SMT	49:30	0.229	D4
Bg	M7+	F#3	F4+	185.0	359.4	272.2	0.561	rmt	68:35	0.216	D4
Bh	p4+	G3	C4+	196.0	269.3	232.6	0.533	rmt	11:8	0.340	D1
Bi	m7+	G#3	F#4+	207.6	380.8	294.2	0.552	rmt	11:6	0.353	D1
Bj	m2+	A3	A#3+	220.0	239.9	230.0	0.643	RMT	12:11	0.319	D2
Bk	p5+	A#3	F4+	233.1	359.4	296.3	0.496	smt	37:24	0.243	D3
Bl	m3+	B3	D4+	246.9	302.2	274.6	0.471	SMT	11:9	0.334	D1
Bm	M2+	C4	D4+	261.6	302.2	281.9	0.521	rmt	15:13	0.300	D2
Bn	p5+	C#4	G4+	277.2	403.5	352.3	0.476	smt	37:24	0.243	D3
Bo	m7+	D4	C5+	293.6	538.5	416.1	0.645	RMT	11:6	0.353	D1
Bp	M7+	D#4	D5+	311.2	604.4	457.8	0.526	rmt	68:35	0.216	D4
Bq	p4+	E4	A4+	329.6	452.9	391.2	0.510	smt	11:8	0.340	D1
Br	m3+	F4	G#4+	349.2	427.4	388.3	0.463	SMT	11:9	0.334	D1
Bs	M6+	F#4	D#5+	370.0	640.6	505.3	0.467	SMT	45:26	0.235	D4
Bt	M3+	G4	B4+	392.0	508.3	450.1	0.477	smt	35:27	0.242	D3
Bu	m2+	G#4	A4+	415.2	452.9	434.0	0.679	RMT	12:11	0.319	D2
Bv	uni+	A4	A4+	440.0	452.9	446.4	0.591	RMT	35:34	0.236	D3
Bw	m6+	A#4	F#5+	466.2	761.7	613.9	0.418	SMT	49:30	0.229	D4
Bx	tri+	B4	F5+	493.8	718.9	606.3	0.460	SMT	16:11	0.303	D2

Block 3.

dyad	mus. intvl.	lower note	upper note	f1 Hz	f2 Hz	mean freq.	eval. rough.	sens. C/D	freq. ratio	$\ln(X+Y)^{-1}$	ratio level
ca	M7+	C3	B3+	130.8	254.1	192.5	0.669	RMT	68:35	0.216	D4
cb	m2+	C#3	D3+	138.6	151.1	144.9	0.717	RMT	12:11	0.319	D2
cc	p5+	D3	A3+	146.8	226.4	186.6	0.633	RMT	37:24	0.243	D3
cd	m7+	D#3	C#4+	155.6	285.3	220.5	0.561	rmt	11:6	0.353	D1
ce	M3+	E3	G#3+	164.8	213.7	189.2	0.515	smt	35:27	0.242	D3
cf	p4+	F3	A#3+	174.6	239.9	207.3	0.569	rmt	11:8	0.340	D1
cg	m3+	F#3	A3+	185.0	226.4	205.7	0.539	rmt	11:9	0.334	D1
ch	m6+	G3	D#4+	196.0	320.3	258.2	0.451	SMT	49:30	0.229	D4
ci	M2+	G#3	A#3+	207.6	239.9	223.8	0.614	RMT	15:13	0.300	D2
cj	tri+	A3	D#4+	220.0	320.3	270.2	0.502	smt	16:11	0.303	D2
ck	M6+	A#3	G4+	233.1	403.5	318.3	0.473	smt	45:26	0.235	D4
cl	uni+	B3	B3+	246.9	254.1	250.5	0.589	rmt	35:34	0.236	D3
cm	m3+	C4	D#4+	261.6	320.3	291.0	0.432	SMT	11:9	0.334	D1
cn	tri+	C#4	G4+	277.2	403.5	340.3	0.459	SMT	16:11	0.303	D2
co	M7+	D4	C#5+	293.6	570.6	432.1	0.585	rmt	68:35	0.216	D4
cp	M3+	D#4	G4+	311.2	403.5	357.3	0.409	SMT	35:27	0.242	D3
cq	uni+	E4	E4+	329.6	339.3	334.4	0.537	rmt	35:34	0.236	D3
cr	m2+	F4	F#4+	349.2	380.8	365.0	0.680	RMT	12:11	0.319	D2
cs	m6+	F#4	D5+	370.0	604.4	487.2	0.477	smt	49:30	0.229	D4
ct	M6+	G4	E5+	392.0	678.5	535.3	0.446	SMT	45:26	0.235	D4
cu	p5+	G#4	D#5+	415.2	640.6	527.9	0.452	SMT	37:24	0.243	D3
cv	m7+	A4	G5+	440.0	807.0	623.5	0.514	smt	11:6	0.353	D1
cw	M2+	A#4	C5+	466.2	538.5	502.4	0.507	smt	15:13	0.300	D2
cx	p4+	B4	E5+	493.8	678.5	586.2	0.469	SMT	11:8	0.340	D1

Table B2

Classification of sensory consonance/dissonance levels based on evaluated roughness (higher is rougher) — Microtuned dyads.

Class	Evaluated C/D	Number of dyads
SMT	< 0.472	18
smt	0.471 < sc ≤ 0.516	18
rmt	0.515 < sc ≤ 0.590	18
RMT	> 0.589	18

Table B3
Classification of frequency-ratio complexity (lower is more dissonant).

Musical interval	Integer ratio (X:Y)	$\ln(X+Y)^{-1}$	Class	Musical interval	Integer ratio (X:Y)	$\ln(X+Y)^{-1}$	Class
				oct.	1:2	0.910	MC
				p5	2:3	0.621	MC
				p4	3:4	0.514	MC
				M6	3:5	0.481	mc
				M3	4:5	0.455	mc
				m3	5:6	0.417	mc
				m6	5:8	0.390	md
m7+	11:6	0.353	D1	**M2**	8:9	0.353	md
p4+	11:8	0.340	D1				
m3+	11:9	0.334	D1				
m2+	12:11	0.319	D2	**M7**	8:15	0.319	md
				m7	9:16	0.311	MD
tri+	16:11	0.303	D2				
M2+	15:13	0.300	D2				
				m2	15:16	0.291	MD
p5+	37:24	0.243	D3				
M3+	35:27	0.242	D3				
uni+	35:34	0.236	D3				
M6+	45:26	0.235	D4				
				tri.	32:45	0.230	MD
m6+	49:30	0.229	D4				
M7+	68:35	0.216	D4				

Appendix C: Block assignment

Table C1

Experiment 1: Just-tuned complex-tone dyads. The arrangement of dyads in three blocks. The symbol "–" stands for a novel presentation and "+" stands for familiar.

Block 1.

Trial #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
dyad	NN	FF	XX	BB	LL	NN	BB	FF	XX	LL	MM	MM	TT	VV	II	TT	II	VV	RR	CC	JJ	JJ	RR	CC
nov./fam.	–	–	–	–	–	+	+	+	+	+	–	+	–	–	–	+	+	+	–	–	–	+	+	+
# intvn.	*	*	*	*	*	4	2	5	5	4	*	0	*	*	*	2	1	3	*	*	*	0	3	3

Trial #	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
dyad	HH	HH	PP	UU	PP	UU	KK	KK	EE	WW	OO	GG	QQ	OO	EE	WW	GG	QQ	DD	AA	SS	SS	DD	AA
nov./fam.	–	+	–	–	+	+	–	+	–	–	–	–	–	+	+	+	+	+	–	–	–	+	+	+
# intvn.	*	0	*	*	1	1	*	0	*	*	*	*	*	2	5	5	4	4	*	*	*	0	3	3

Block 2.

Trial #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
dyad	Ae	Av	Am	Am	Av	Ar	Ae	An	Al	An	Au	Ar	Al	Au	Ao	Ao	Ai	Aw	Ah	Aa	Aq	Ab	Aq	Ai
nov./fam.	–	–	–	+	+	–	+	–	–	+	–	+	+	+	–	+	–	–	–	–	–	–	+	+
# intvn.	*	*	*	0	2	*	5	*	*	1	*	5	3	2	*	0	*	*	*	*	*	*	1	6

Trial #	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
dyad	Aw	Ah	Aa	Ab	Ag	Ag	Ap	Ad	Ax	As	Ax	Ap	Ad	As	Aj	Aj	Ak	Ac	Af	At	At	Ac	Ak	Af
nov./fam.	+	+	+	+	–	+	–	–	–	–	+	+	+	+	–	+	–	–	–	–	+	+	+	+
# intvn.	6	6	6	5	*	0	*	*	*	*	1	4	4	3	*	0	*	*	*	*	0	3	5	5

Block 3.

Trial #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
dyad	bl	bb	be	bi	bl	bi	bb	bx	be	bn	bx	bo	bo	bn	bq	bv	br	bm	bm	bv	bq	br	bs	bh
nov./fam.	–	–	–	–	+	+	+	–	+	–	+	–	+	+	–	–	–	–	+	+	+	+	–	–
# intvn.	*	*	*	*	3	1	4	*	5	*	2	*	0	3	*	*	*	*	0	3	5	4	*	*

Trial #	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
dyad	bs	bg	bh	bc	ba	bf	bg	ba	bd	bc	bg	bd	bp	bp	bt	bu	bt	bk	bj	bu	bj	bw	bw	bk
nov./fam.	+	–	+	–	–	–	+	+	–	+	+	+	–	+	–	–	+	–	–	+	+	–	+	+
# intvn.	1	*	2	*	*	*	4	2	*	5	4	2	*	0	*	*	1	*	*	3	1	*	0	5

Table C2

Experiment 2: Microtuned dyads.

Block 1.

Trial#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
dyad	DDD	PPP	SSS	GGG	CCC	DDD	GGG	PPP	SSS	CCC	AAA	AAA	III	XXX	QQQ	III	QQQ	XXX	RRR	JJJ	VVV	VVV	RRR	JJJ
nov./fam.	–	–	–	–	–	+	+	+	+	+	–	+	–	–	–	+	+	+	–	–	–	+	+	+
# intvn.	*	*	*	*	*	4	2	5	5	4	*	0	*	*	*	2	1	3	*	*	*	0	3	3

Trial#	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
dyad	LLL	LLL	MMM	BBB	MMM	BBB	OOO	OOO	FFF	NNN	EEE	KKK	UUU	EEE	FFF	NNN	KKK	UUU	HHH	TTT	WWW	WWW	HHH	TTT
nov./fam.	–	+	–	–	+	+	–	+	–	–	–	–	–	+	+	+	+	+	–	–	–	+	+	+
# intvn.	*	0	*	*	1	1	*	0	*	*	*	*	*	2	5	5	4	4	*	*	*	0	3	3

Block 2.

Trial#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
dyad	Bs	Bw	Bv	Bv	Bw	Bn	Bs	Br	Bi	Br	Bo	Bn	Bi	Bo	Bq	Bq	Bt	Bp	Bf	Bc	Bu	Ba	Bu	Bt
nov./fam.	–	–	–	+	+	–	+	–	–	+	–	+	+	+	–	+	–	–	–	–	–	–	+	+
# intvn.	*	*	*	0	2	*	5	*	*	1	*	5	3	2	*	0	*	*	*	*	*	*	1	6

Trial#	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
dyad	Bp	Bf	Bc	Ba	Bl	Bl	Bm	Be	Bb	Bj	Bb	Bm	Be	Bj	Bg	Bg	Bh	Bx	Bd	Bk	Bk	Ba	Bu	Bt
nov./fam.	+	+	+	+	–	+	–	–	–	–	+	+	+	+	–	+	–	–	–	–	+	+	+	+
# intvn.	6	6	6	5	*	0	*	*	*	*	1	4	4	3	*	0	*	*	*	*	0	3	5	4

Block 3.

Trial #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
dyad	cl	ce	cc	cr	cl	cr	ce	cw	cc	cx	cw	cm	cm	cx	cf	ck	cb	cd	cd	ck	cf	cb	cp	cu
nov./fam.	–	–	–	–	+	+	+	–	+	–	+	–	+	+	–	–	–	–	+	+	+	+	–	–
# intvn.	*	*	*	*	3	1	4	*	5	*	2	*	0	3	*	*	*	*	0	3	5	4	*	*

Trial #	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
dyad	cp	ch	cu	cv	co	cq	ch	co	cj	cv	cq	cj	cn	cn	cg	ca	cg	ct	cs	ca	cs	ci	ci	ct
nov./fam.	+	–	+	–	–	–	+	+	–	+	+	+	–	+	–	–	+	–	–	+	+	–	+	+
# intvn.	1	*	2	*	*	*	4	2	*	5	4	2	*	0	*	*	1	*	*	3	1	*	0	5

Table 1

Discriminability index, response bias, and mean adjusted proportion correct (p(c)), averaged over five retention intervals for just-tuned complex tone dyads.*

Musicians

	MC	mc	md	MD	SCT	sct	rct	RCT	Grand mean	Std. dev.
d'	1.47	1.61	1.68	1.61	1.77	1.37	1.52	1.71	1.60[b]	0.15
c	0.01	–0.09	0.17	–0.20	–0.01	–0.14	–0.06	0.24	–0.01	0.15
*mean p(c)**	0.78	0.78	0.73[a]	0.84[a]	0.80	0.78	0.81	0.68	0.77	0.05
std. dev.	0.03	0.03	0.03	0.03	0.04	0.03	0.04	0.04		

Nonmusicians

	MC	mc	md	MD	SCT	sct	rct	RCT	Grand mean	Std. dev.
d'	1 .30	1.39	1.01	1.03	1.23	1.02	1.34	1.29	1.20[b]	0.16
c	–0.02	–0.13	–0.13	–0.15	–0.01	–0.17	–0.15	–0.10	–0.11	0.06
*mean p(c)**	0.69	0.72	0.68	0.69	0.66	0.69	0.76	0.70	0.70	0.03
std. dev.	0.04	0.02	0.03	0.04	0.03	0.02	0.03	0.03		

Note. Discriminability index (d') is the inverse of the normal distribution function (z) for the hit rate (H = familiar correct/number of familiar trials) minus the false alarm rate (F = novel incorrect/number of novel trials), i.e., $d' = z(H) - z(F)$. Criterion (c) is the response bias where $c = -0.5*(z(H) + z(F))$. A positive value of c indicates a tendency to respond "no" (i.e., "I have not heard that before"). A negative value of c indicates a tendency to respond "yes." When there is no response bias, the value is 0.0. The proportions correct adjusted for guessing (p(c)*) is calculated from each participants' proportion of hits and false alarms by the formula $p(c)^* = 0.5*[P(H) + (1 - P(F))]$ (Macmillan & Creelman, 2005, pp 3-31).

[a] Significantly different at $p \leq 0.008$. Other means were compared at specific retention periods due to a significant Retention × C/D interaction.

[b] Significantly different at $p = 0.03$.

Table 2
Significant nonparametric test results for classes of dyads at specific retentions periods — Just-tuned complex-tone dyads.

			Friedman test			Wilcoxon test		
Group	C/D type	retention period	χ^2	*sig.*	*W*	pair	*z*	*sig.*
Mus	cogn.	32.50s	8.01	.05	0.18	MD-mc	2.03	.042
						MD-md	2.56	.010[a]
Nmus	cogn.	24.25s	9.74	.02	0.22	MD-mc	2.59	.010[a]
						md-mc	2.59	.010[a]
	sens.	24.50s	10.88	.01	0.24	RCT-sct	3.04	.002[a]*
	sens.	32.50s	13.02	.00	0.29	RCT-SCT	2.12	.034[a]
						RCT-sct	2.34	.019[a]
	sens.	40.75s	8.30	.04	0.18	rct-RCT	2.07	.038

For brevity only the Friedman tests with a significance value of $p < .05$ are shown. Wilcoxon signed-rank comparisons with pairwise differences having $p > .05$ are not shown.
[a] Significantly different in post hoc tests following parametric ANOVA.
* Significantly different Wilcoxon test following nonparametric ANOVA at $p \leq .008$.

Table 3
Mean recognition scores (out of 6) for 24 root notes and F(23, 667) test results — Just-tuned complex-tone dyads.

Root Note	*M, SD*	Root Note	*M, SD*	Root Note	*M, SD*
C4	5.6 ± 0.62	A3	4.8 ± 1.06 *	G#3	4.4 ± 1.00 *
B4	5.2 ± 0.97	C3	4.7 ± 0.94 *	D#3	4.4 ± 1.16 * **
G4	5.2 ± 0.83	F3	4.7 ± 0.95 *	E3	4.3 ± 0.84 * ***
F#4	4.9 ± 0.86	G3	4.6 ± 0.93 *	B3	4.2 ± 1.19 * ***
D4	4.9 ± 0.90 *	A#4	4.6 ± 0.93 *	C#3	4.2 ± 1.01 * ***
A4	4.9 ± 0.97 *	G#4	4.5 ± 0.97 *	E4	4.2 ± 1.00 * ***
D3	4.8 ± 0.90 *	A#3	4.5 ± 1.04 *	F4	4.2 ± 1.03 * ** ***
C#4	4.8 ± 0.94	F#3	4.4 ± 0.82 *	D#4	4.0 ± 1.19 * ** *** ****

* Significantly worse than C4, $p < .0002$.
** Significantly worse than B4, $p < .0002$.
*** Significantly worse than G4, $p < .0002$.
**** Significantly worse than F#4, $p < .0002$.

Table 4
Mean recognition scores at 48.00 s in comparison to chance (0.50) performance — Just-tuned complex-tone dyads.

	Musicians				Nonmusicians			
class	*M*	*SD*	*t*(14)	*sig.*	*M*	*SD*	*t*(14)	*sig.*
MC	0.71	0.24	3.44	.004	0.72	0.14	6.23	.000
mc	0.67	0.23	2.93	*ns*	0.72	0.19	4.48	.001
md	0.70	0.24	3.14	*ns*	0.65	0.20	2.83	*ns*
MD	0.70	0.21	3.64	.003	0.63	0.24	2.10	*ns*
SCT	0.78	0.23	4.62	.000	0.63	0.15	3.32	.005
sct	0.77	0.19	5.42	.000	0.75	0.18	5.61	.000
rct	0.68	0.22	3.10	*ns*	0.68	0.21	3.27	.006
RCT	0.32	0.47	–1.46	*ns*	0.42	0.47	–0.64	*ns*

Note: $p \leq .05/8 = .006$ is required for significance.

Table 5
Discriminability index, response bias, and mean adjusted percent correct and standard deviations, averaged over five retention intervals for microtuned dyads.

Musicians	D1	D2	D3	D4	SMT	smt	rmt	RMT	Grand mean	Std. dev.
d′	1.53	1.25	1.67	1.69	1.45	1.33	1.73	1.57	1.53[b]	0.17
c	0.08	–0.01	–0.18	–0.05	–0.25	0.09	–0.01	0.06	–0.03	0.12
*mean p(c)**	0.69	0.70	0.81	0.79	0.80	0.73	0.78	0.73	0.76	0.05
std. dev.	0.04	0.03	0.03	0.03	0.02	0.03	0.03	0.02		

Nonmusicians	D1	D2	D3	D4	SMT	smt	rmt	RMT	Grand mean	Std. dev.
d′	1.25	1.10	1.33	1.32	1.18	1.18	1.46	1.19	1.25[b]	0.11
c	–0.11	–0.08	–0.16	–0.05	–0.27	–0.11	–0.08	0.06	–0.10	0.09
*mean p(c)**	0.68	0.66	0.73	0.72	0.74	0.72	0.72	0.65	0.70	0.04
std. dev.	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.03		

Note. See Table 1.
[b] Significantly different at p <.01.

Table 6
Significant nonparametric test results for classes of dyads at specific retention periods — Microtuned dyads.

			Friedman test			Wilcoxon test		
Group	C/D type	retention period	χ^2	*sig.*	*W*	pair	*z*	*sig.*
Mus.	cogn.	24.25s	15.34	.002	0.34	D3-D1	2.20	.028[a]
						D3-D2	3.18	.001[a]*
						D3-D4	2.00	.046
						D4-D2	2.12	.034
	cogn.	40.75s	8.33	.040	0.18	D2-D1	2.28	.023
	cogn.	48.00s	14.54	.002	0.32	D3-D2	2.90	.004[a]*
						D4-D2	2.07	.038
	sens.	24.25s	16.28	.001	0.36	SMT-smt	2.47	.014[a]
						rmt-smt	2.97	.003[a]*
						SMT-RMT	2.12	.034
						rmt-RMT	2.17	.030
Nmus.	cogn.	40.75s	9.42	.024	0.21	D2-D1	2.50	.012[a]
						D3-D1	2.16	.031
	sens.	48.00s	15.26	.002	0.34	smt-SMT	1.99	.046
						smt-RMT	2.79	.005*
						rmt-RMT	2.94	.003*

[a] Significantly different in post hoc tests following parametric ANOVA.
* Significantly different Wilcoxon test following nonparametric ANOVA at $p \leq .008$.

Table 7

Mean recognition scores (out of 6) for 24 root notes and F(23, 667) test results — Microtuned dyads.

Root Note	*M, SD*	Root Note	*M, SD*	Root Note	*M, SD*
D3	5.4 ± 0.61	E4	4.8 ± 0.95	A#3	4.3 ± 0.84 * **
A#4	5.2 ± 0.71	G#3	4.8 ± 0.94	F#3	4.3 ± 1.08 * **
B4	5.2 ± 0.91	C3	4.7 ± 0.91	D#4	4.2 ± 0.95 * **
C#3	5.2 ± 0.70	C4	4.7 ± 0.74	G4	4.1 ± 1.28 * **
C#4	5.2 ± 0.65	D4	4.7 ± 1.08	F3	4.1 ± 0.98 * ** ***
F#4	4.9 ± 0.91	E3	4.7 ± 1.14	G#4	4.0 ± 1.00 * **
A4	4.9 ± 0.78	B3	4.7 ± 0.88	G3	3.7 ± 1.03 * ** *** ****
D#3	4.8 ± 0.91	F4	4.5 ± 0.86 *	A3	3.4 ± 1.19 * ** *** **** *****

* Significantly worse than D3, $p < .0002$.
** Significantly worse than A#4, B4, C#3, and C#4, $p < .0002$.
*** Significantly worse than F#4 and A4, $p < .0002$.
**** Significantly worse than D#3, E4, G#3, C3, C4, D4, and B3, $p < .0002$.
***** Significantly worse than E3, $p < .0002$.

Table 8
Mean recognition scores at 48.00 s in comparison to chance (0.50) performance — Microtuned dyads.

	Musicians				Nonmusicians			
class	*M*	*SD*	*t*(14)	*sig.*	*M*	*SD*	*t*(14)	*sig.*
D1	0.76	0.21	4.23	.000	0.61	0.20	2.10	*ns*
D2	0.52	0.29	0.22	*ns*	0.41	0.34	–1.06	*ns*
D3	0.88	0.14	10.59	.000	0.69	0.23	3.16	.007[a]
D4	0.77	0.23	4.54	.000	0.67	0.25	2.71	*ns*
SMT	0.76	0.23	4.46	.001	0.57	0.29	0.91	*ns*
smt	0.83	0.22	5.71	.000	0.77	0.22	4.74	.000
rmt	0.76	0.22	4.55	.000	0.78	0.21	5.24	.000
RMT	0.64	0.22	2.54	*ns*	0.47	0.19	–0.58	*ns*

Note: p = .05/8 = .006 is required for significance.
[a] marginally significant.

Figure Captions

Figure 1. Musicians' mean proportion of correct recognition of just-tuned, complex-tone dyads across retention periods for four classes of cognitive consonance/dissonance. Dyads in MD classes were recognized significantly more often than md classes after scores were corrected for guessing. At 48.00 s, musicians recognized MD and MC dyads significantly better than chance.

Figure 2. Nonmusicians' mean proportion of correct recognition of just-tuned, complex-tone dyads across retention periods for four classes of cognitive C/D. Dyads classified as MD and md were recognized significantly more often than mc dyads at 24.25 s retention after scores were corrected for guessing. At 48.00 s, nonmusicians recognized MC and mc dyads significantly better than chance.

Figure 3. Musicians' mean proportion of correct recognition of just-tuned complex-tone dyads across retention periods for four classes of sensory C/D. At 48.00 s, musicians recognized the "smoothest" dyads, SCT and sct, significantly better than chance after scores were corrected for guessing.

Figure 4. Nonmusicians' mean proportion of correct recognition of just-tuned complex-tone dyads across retention periods for four classes of sensory C/D. Dyads classified as RCT or most "rough" were recognized significantly more often than other dyads at 24.25 and 32.50 s of retention after scores were corrected for guessing. At 40.75 s, RCT dyads were poorly recognized compared to less rough rct dyads. At 48.00 s, RCT dyads were the only sensory C/D class that nonmusicians did not recognize significantly better than chance.

Figure 5. Musicians' mean proportion of correct recognition of microtuned dyads across retention periods for four classes of frequency-ratio C/D. D3 dyads were recognized significantly better than other frequency-ratio classes at 24.25 s after scores were corrected for guessing. At 40.75 s, D2 dyads significantly outperformed D1 dyads, but were significantly worse compared to D3 and D4 at 48.00 s. All but D2 dyads were recognized significantly better than chance at 48.00 s.

Figure 6. Nonmusicians' mean proportion of correct recognition of microtuned dyads across retention periods for four classes of frequency-ratio C/D. Dyads classified as D1 were poorly recognized compared to D2 dyads at 40.75 s after scores were corrected for guessing. At 48.00 s, nonmusicians recognized D3 dyads significantly better than chance.

Figure 7. Musicians' mean proportion of correct recognition of microtuned dyads across retention periods for four classes of sensory C/D. At 24.25 s, musicians recognized the smoothest (SMT) and moderately rough (rmt) dyads significantly better than moderately smooth (smt) dyads, after scores were corrected for guessing. At 48.00 s, the roughest (RMT) dyads were the only sensory C/D class that musicians did not recognize significantly better than chance.

Figure 8. Nonmusicians' mean proportion of correct recognition of microtuned dyads across retention periods for four classes of sensory C/D. At 48.00 s, smt and rmt dyads were recognized significantly better than RMT dyads after scores were corrected for guessing. Both smt and rmt dyad classes were recognized by nonmusicians significantly better than chance at 48.00 s.

Figure 9. Mean number of dyad recognitions (novel and familiar) by root note for (a) just-tuned, complex-tone dyads and (b) microtuned dyads.

Figure 1

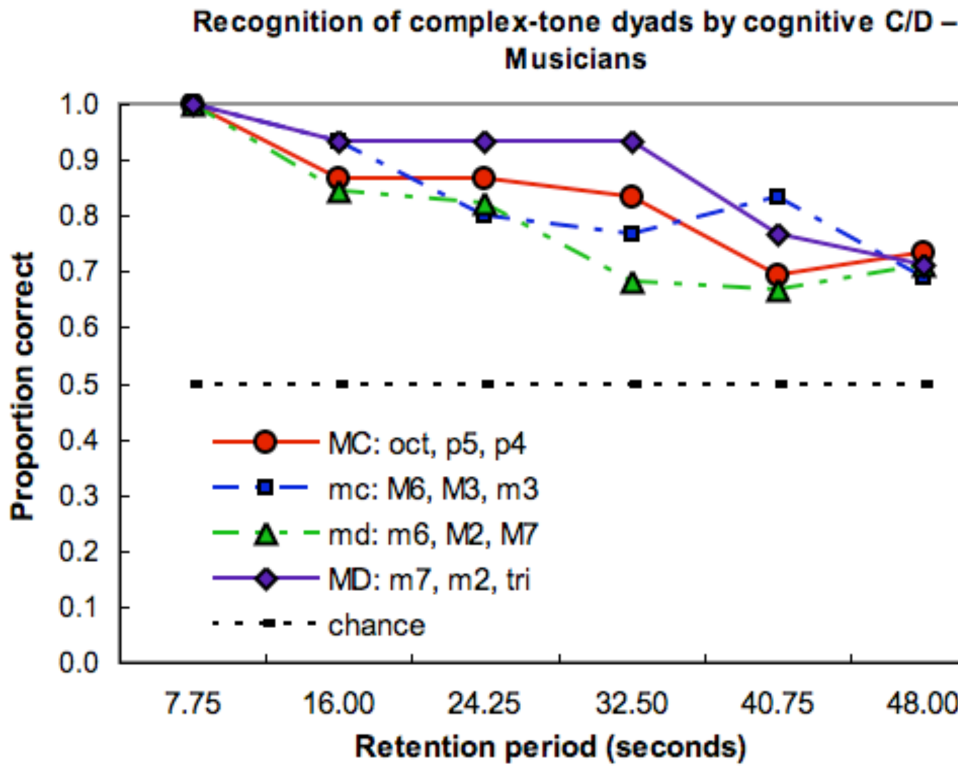

Recognition of complex-tone dyads by cognitive C/D –
Musicians
Proportion correct
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
7.75
16.00
24.25
32.50
40.75
48.00
Retention period (seconds)
MC: oct, p5, p4
mc: M6, M3, m3
md: m6, M2, M7
MD: m7, m2, tri
chance

Figure 2

Recognition of complex-tone dyads by cognitive C/D – Nonmusicians

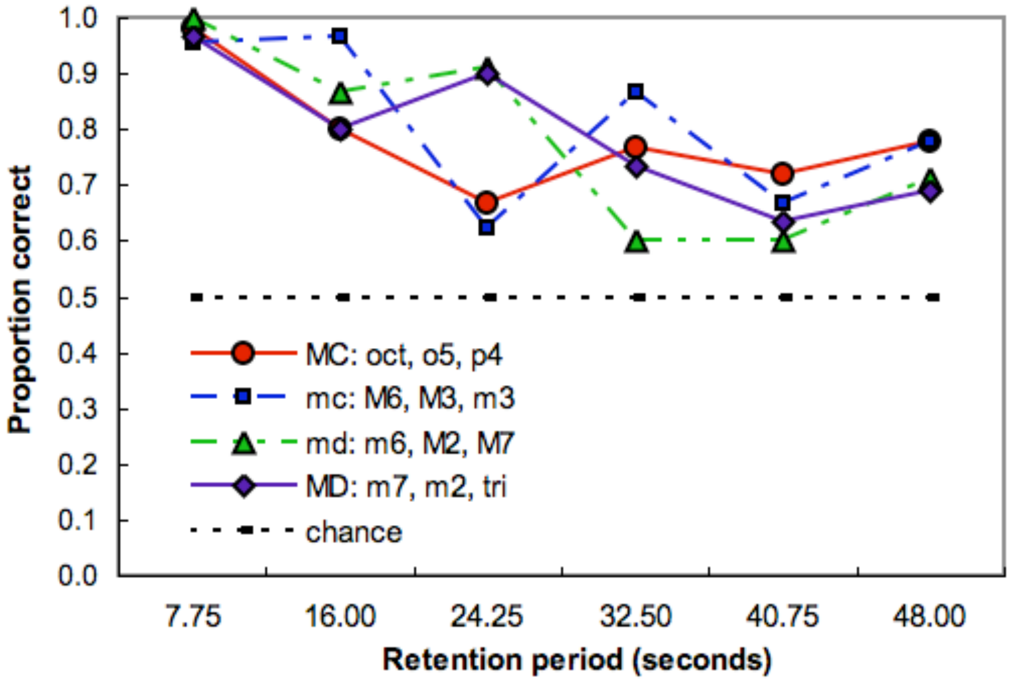


Figure 3

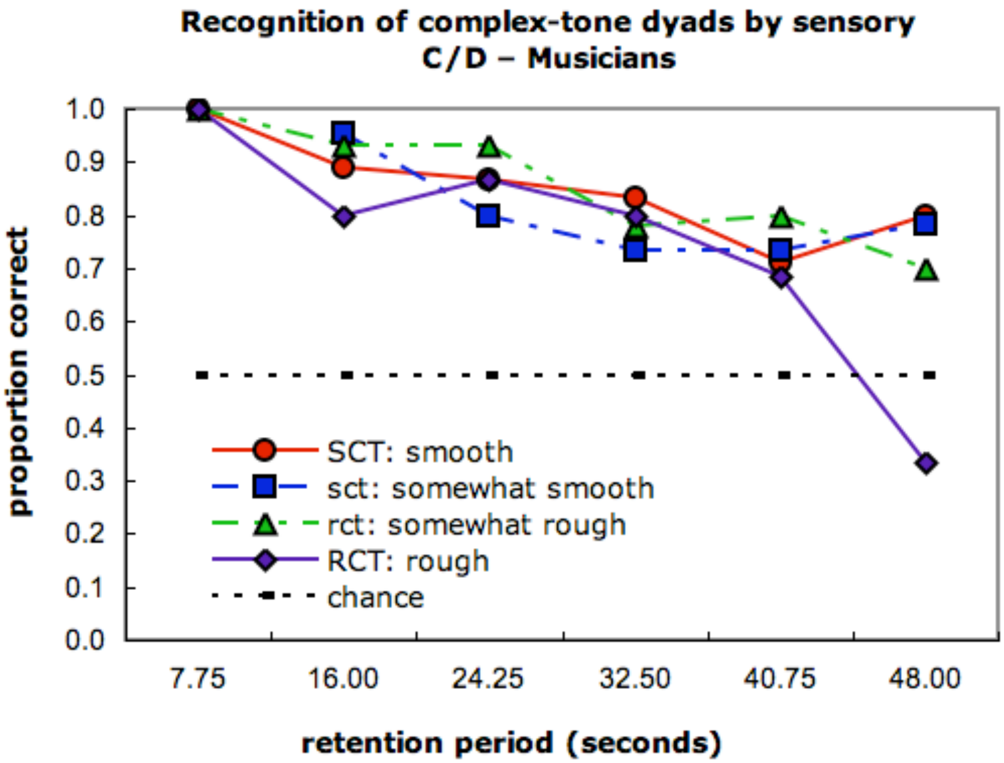

Recognition of complex-tone dyads by sensory
C/D – Musicians
proportion correct
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
SCT: smooth
sct: somewhat smooth
rct: somewhat rough
RCT: rough
chance
7.75
16.00
24.25
32.50
40.75
48.00
retention period (seconds)

Figure 4

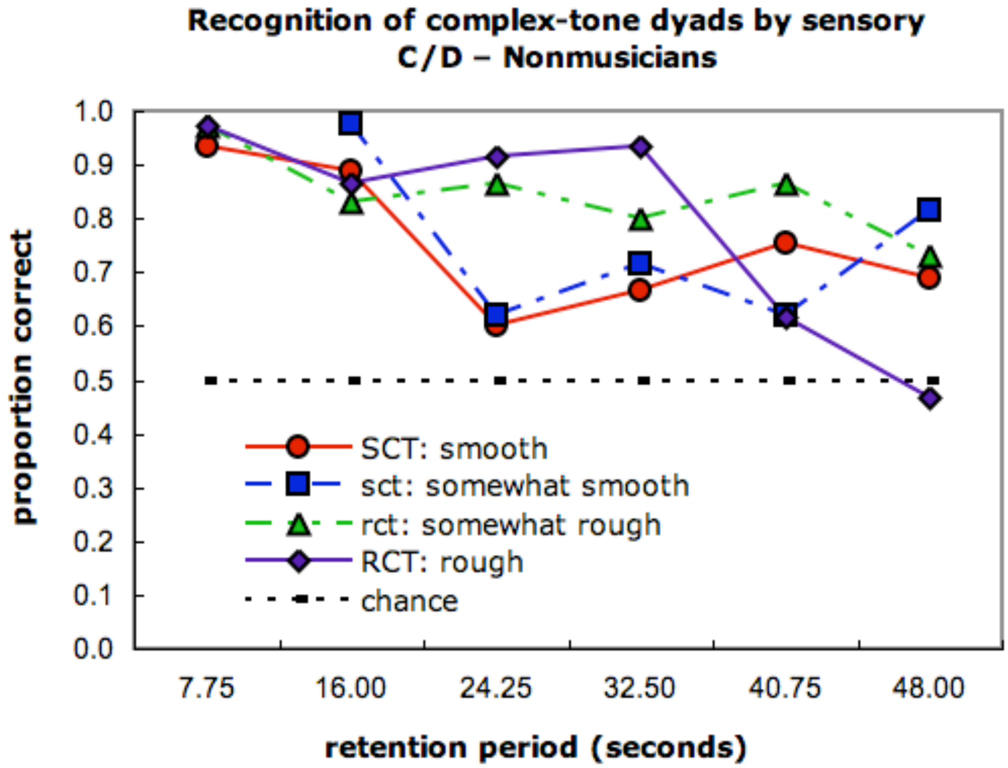

Recognition of complex-tone dyads by sensory
C/D – Nonmusicians
proportion correct
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
SCT: smooth
sct: somewhat smooth
rct: somewhat rough
RCT: rough
chance
7.75
16.00
24.25
32.50
40.75
48.00
retention period (seconds)

Figure 5

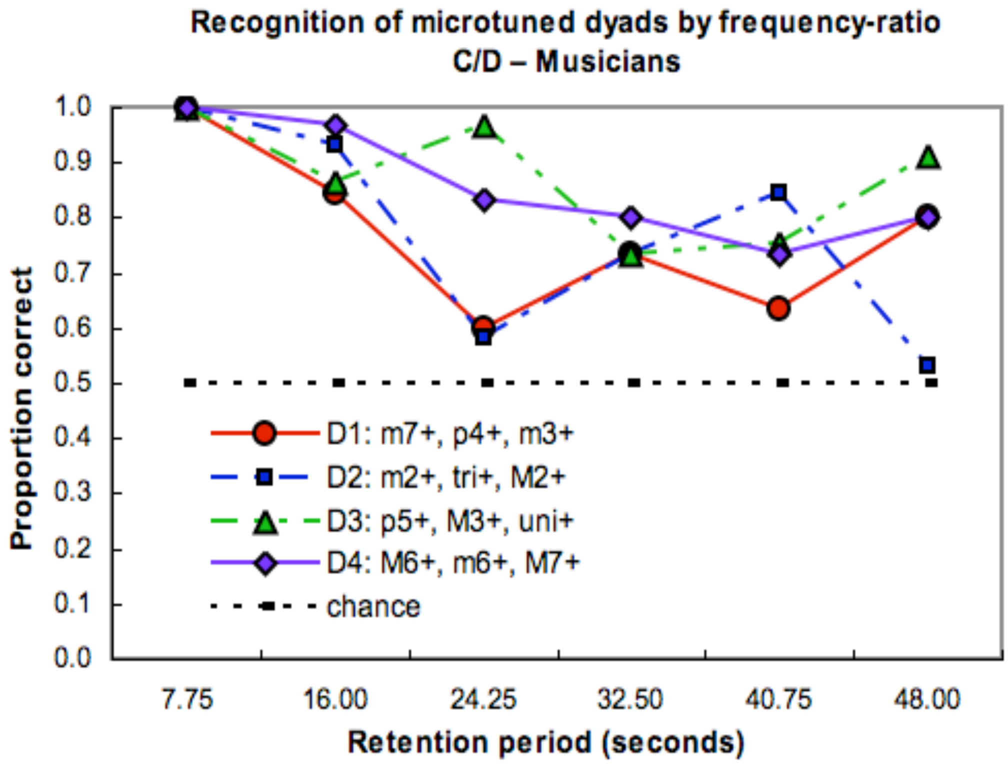

Recognition of microtuned dyads by frequency-ratio
C/D – Musicians
Proportion correct
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
D1: m7+, p4+, m3+
D2: m2+, tri+, M2+
D3: p5+, M3+, uni+
D4: M6+, m6+, M7+
chance
7.75
16.00
24.25
32.50
40.75
48.00
Retention period (seconds)

Figure 6

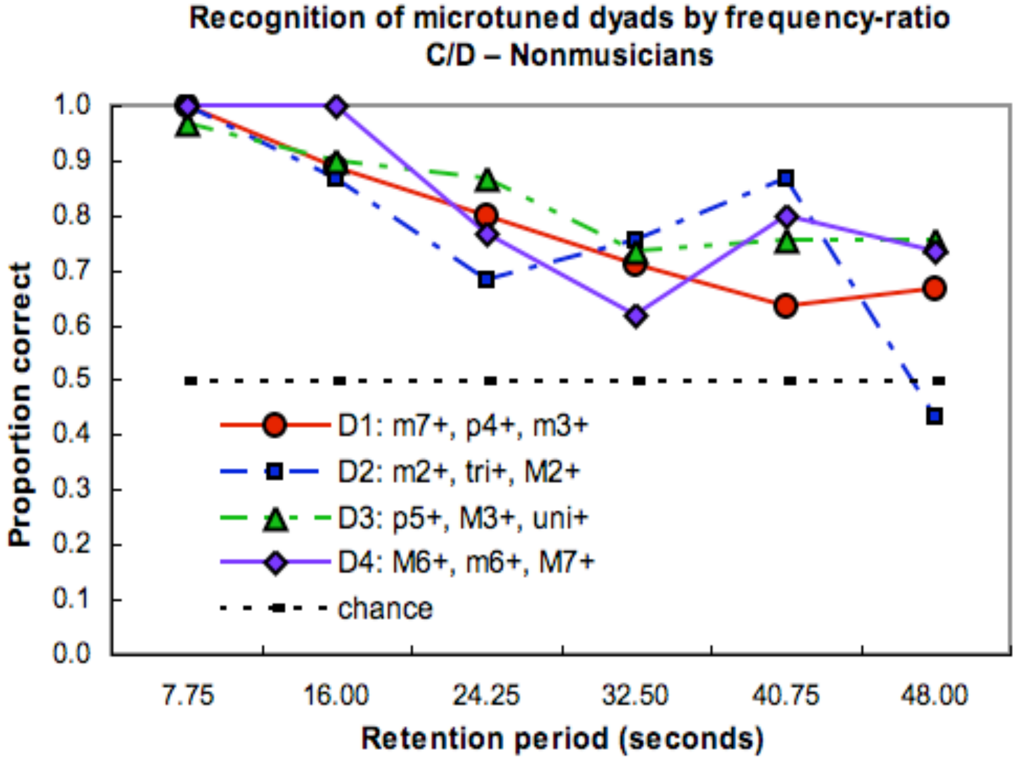

Recognition of microtuned dyads by frequency-ratio
C/D – Nonmusicians
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Proportion correct
D1: m7+, p4+, m3+
D2: m2+, tri+, M2+
D3: p5+, M3+, uni+
D4: M6+, m6+, M7+
chance
7.75
16.00
24.25
32.50
40.75
48.00
Retention period (seconds)

Figure 7

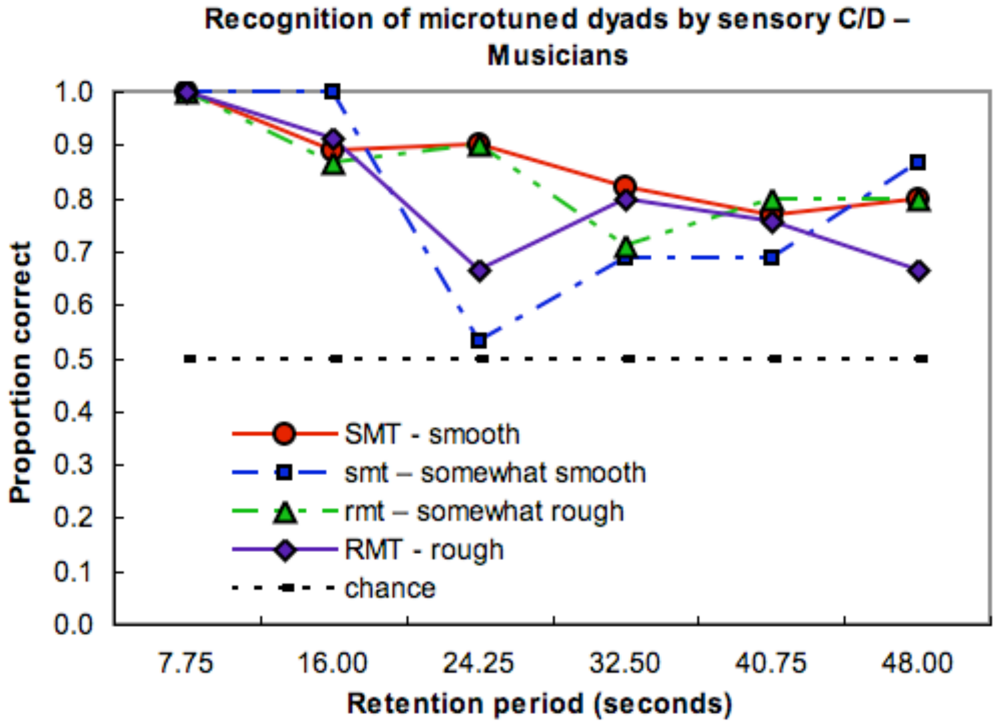

Recognition of microtuned dyads by sensory C/D – Musicians
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Proportion correct
SMT - smooth
smt – somewhat smooth
rmt – somewhat rough
RMT - rough
chance
7.75
16.00
24.25
32.50
40.75
48.00
Retention period (seconds)

Figure 8

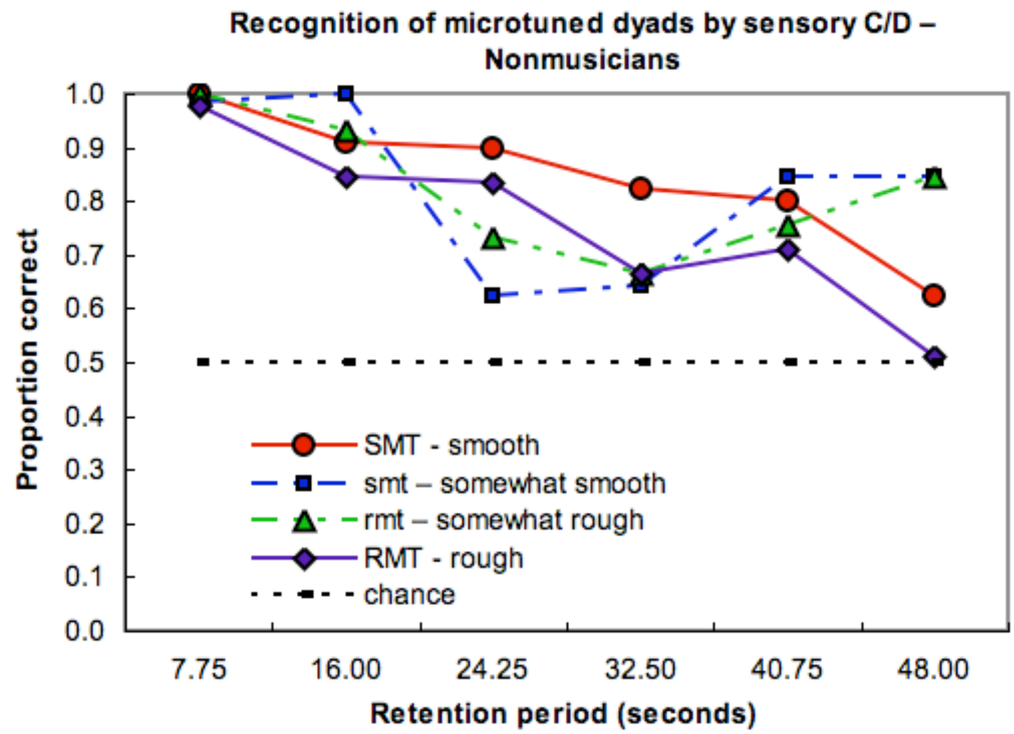
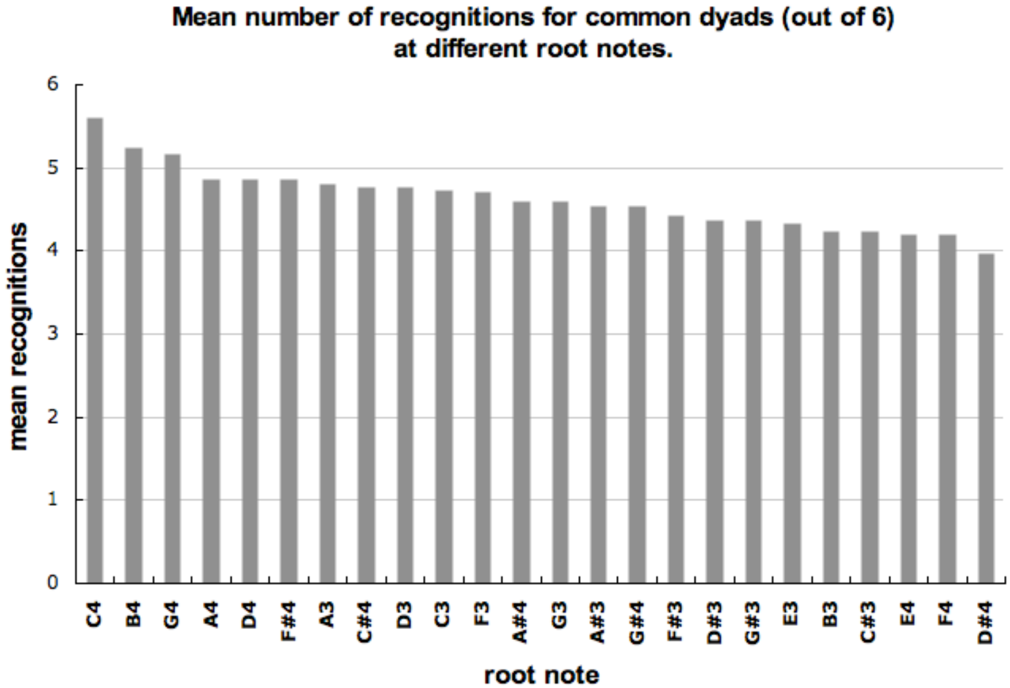
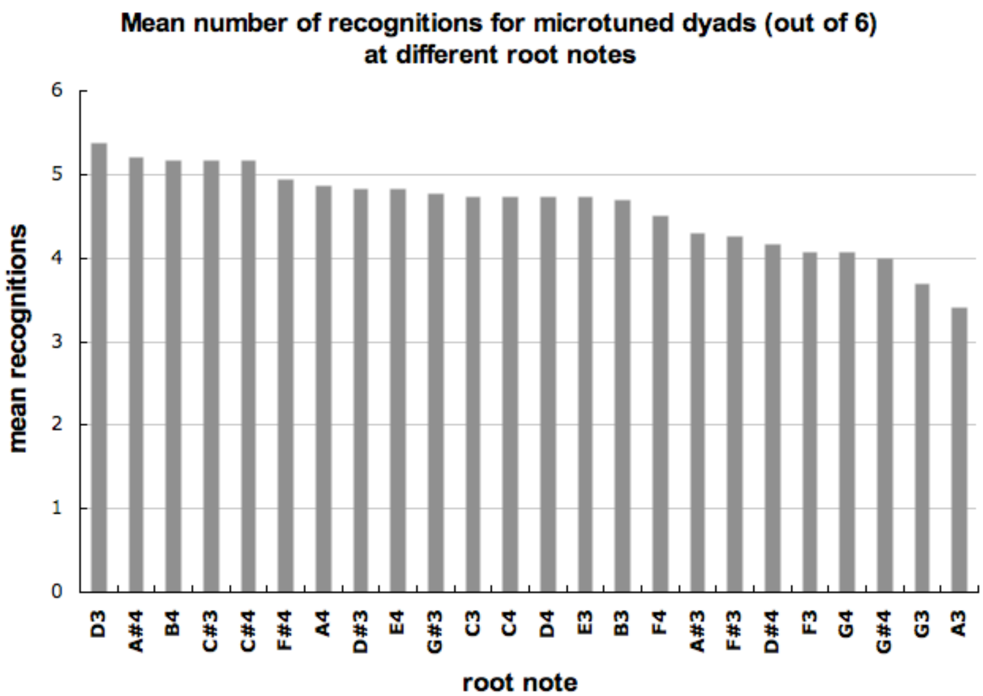

Recognition of microtuned dyads by sensory C/D – Nonmusicians
Proportion correct
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
7.75
16.00
24.25
32.50
40.75
48.00
Retention period (seconds)
SMT - smooth
smt – somewhat smooth
rmt – somewhat rough
RMT - rough
chance

Figure 9

a.


Mean number of recognitions for common dyads (out of 6) at different root notes.
mean recognitions
0
1
2
3
4
5
6
C4 B4 G4 A4 D4 F#4 A3 C#4 D3 C3 F3 A#4 G3 A#3 G#4 F#3 D#3 G#3 E3 B3 C#3 E4 F4 D#4
root note

b.


Mean number of recognitions for microtuned dyads (out of 6) at different root notes
mean recognitions
0
1
2
3
4
5
6
D3 A#4 B4 C#3 C#4 F#4 A4 D#3 E4 G#3 C3 C4 D4 E3 B3 F4 A#3 F#3 D#4 F3 G4 G#4 G3 A3
root note

CHAPTER 5

Chapter 5: Summary

Summary

Psychologists describe a sound's degree of consonance (C) or dissonance (D) from two main perspectives — one related to its physical dimensions as encoded by the sensory systems in the ear and the other related to the aesthetic degree of pleasantness as regarded by the cognitive systems in the brain. *Sensory C/D* is subsumed under the broader phenomenon of *cognitive C/D* (also termed *musical C/D*; Terhardt, 1984). For sounds embedded in musical compositions, cognitive C/D is influenced by the tonal context and the listener's musical enculturation (Cazden, 1945; Terhardt, 1984). Historically, the distinction was considered separable enough that adult listeners with normal hearing could consistently and accurately respond to the sensory properties of two simultaneous tones (a *dyad*), or the dyad's relative harmoniousness, in the absence of musical context (Ayers, Aeschbach, & Walker, 1980; Butler & Daston, 1968; Guernsey, 1928; Helmholtz, 1885/1954; Kameoka & Kuriyagawa, 1969a, 1969b; Malmberg, 1918; Plomp & Levelt, 1965; Regnault, Bigand, & Besson, 2001; Schellenberg & Trainor, 1996; Terhardt, 1974b, 1984; Van de Geer, Levelt, & Plomp, 1962). This thesis investigates those assumptions using methods that engage sensory or bottom-up, and cognitive or top-down, processing.

The principal findings of this thesis were as follows. Evaluated sensory C/D was found to be dependent on the life experiences of the perceiver. In particular, musical training was found to shape a listener's perceptual focus towards a musical interval's conceptual meaning and to heighten sensitivity to harmonic partials. I showed that cognitive consonances — the so-called "natural musical intervals" — impart no special advantage to auditory short-term memory (STM), which is shown to be more robust and accurate for longer durations than previously described. These findings have implications for theories of music cognition, auditory STM, and the psychophysical scaling of auditory event properties.

Chapter 2 systematically examined the influence of formal musical training on listeners' assessments of dyad *roughness* — a primary component of sensory dissonance. Chapter 3 showed the extent to which musical training and properties of sensory C/D and cognitive C/D affect auditory short-term memory; Chapter 4 supported and gave perspective to these findings by including more complex, ecologically valid and varied stimuli. The experiments reported in Chapters 2 and 4 asked Western listeners to respond to dyads in traditional Western semitone tuning and in rare (at least to the Western music listeners used here) quartertone tuning to study how implicit and explicit knowledge of tonality affects perceptual processing.

Here I review the thesis's overarching themes followed by a review of the main findings, contributions to knowledge, and an outline of the thesis's novelty from theoretical and methodological perspectives. Future directions are discussed in the closing pages.

Overarching Themes

Consonance and dissonance and the origins of the distinction.

Biologist Lewis Wolpert (2006) wrote that the ultimate purpose of our sensory organs is to tell us how to move, i.e., to help us decide whether to approach or flee. The origin of the C/D distinction might be rooted in auditory events having certain properties that conveyed an evolutionary advantage to early humans. Is the C/D distinction acquired through exposure or does it reflect an innate property of the system? Human preference and heightened sensitivity for consonance over dissonance is well documented in infants and adults (Blood, Zatorre, Bermudez, & Evans, 1999; Schellenberg & Trehub, 1994a, 1994b; Trainor & Heinmiller, 1998; Trainor & Trehub, 1993; Tramo, Cariani, Delgutte, & Braida, 2003; Tufts, Molis, & Leek, 2005; Zentner & Kagan, 1998). Some have proposed that consonance acuity arose as a by-product of the way that spectral relationships in speech sounds (formants) allow humans to differentiate among phonemes (Ross, Choi, & Purves, 2007; Schwartz, Howe, & Purves, 2003). Although the link between consonance acuity and aspects of speech seems likely, organisms that lack speech such as birds, rats, and some primates have also shown a preference for consonance over dissonance (Fannin & Braud, 1971; Hulse, Bernard, & Braaten, 1995; Izumi, 2000; Sugimoto et al., 2010). Consonance preference in nonhuman animals suggests that the C/D distinction evolved to serve neither a speech-specific nor a music-specific aesthetic response system. Dissonance aversion on the other hand, a behavior that could be indistinguishable from consonance preference under some conditions, may be found to have a biological basis rooted in the quality or timbre of distress calls. Much more work is required before it can be established that consonance preference and its corollary, the human music-making capacity, reflect music-specific biological adaptations that promoted human survival. Contrary evidence has shown that at least one closely related species of new world monkey does not regard consonance as qualitatively different from dissonance (McDermott & Hauser, 2004).

A fundamental theme of the present research concerns whether a processing advantage for consonance might appear in a nonmusical cognitive task, outside of a musical context. If so, and if this experimental approach were extended to a nonhuman animal (e.g., the Mongolian gerbil, *meriones unguiculatus*), these data would inform the nature/nurture question at the heart of the C/D distinction.

The question of whether there is a cognitive processing advantage for consonant over dissonant sounds as found in this thesis is a qualified "no". It appears that dissonance is no more cognitively taxing than consonance in a running memory task. In fact, the most consonant musical intervals are the least recognized under heavy cognitive loads. This finding is what might be expected of a processing system specialized to detect and maintain “deviant,” unexpected, or biologically salient properties such as roughness in distress calls (Bigand & Tillmann, 2005; Fabiani & Donchin, 1995). In discussing musical intervals, what constitutes deviance is unique to the cumulative experience of the perceiver and his or her musical culture. What appears to be an innate consonance predilection might be part of a broader cognitive advantage that includes the relative ease with which common sounds in our environment can be ignored.

A secondary theme of this thesis concerns individual differences so that a finer point can be put on the influence of nurture or life experiences on C/D perception. Perceptual differences among humans have recently been described using neurophysiological tools to objectively measure similarities and differences among those with various levels of musical training (Brattico et al., 2009; Foss, Altschuler, & James, 2007; Minati et al., 2009; Regnault et al., 2001; Schön, Regnault, Ystad, & Besson, 2005). The research into neurobiological differences among populations has energized the centuries-long tradition of C/D research, although the full extent to which various musical environments and types of musical training might have an impact on C/D perception is unknown at this early juncture. This thesis provides behavioral data showing that formal musical training modifies the capacity to perceive and identify elements of dissonance. The data on subjective C/D values is thus updated here in order to inform and inspire future neurophysiological work describing auditory perceptual differences between those with and without formal musical training.

Review of the Main Findings

A sound's degree of consonance or dissonance is mediated by both bottom-up (feature-driven) and top-down (knowledge-driven) processing mechanisms. When discussing the distinction as it originates in the auditory periphery, the term *sensory C/D* is used. When the C/D distinction is based upon the rules and practice of a culture's music-theory, the term *cognitive* (or *musical*) *C/D* is used. The independence of these two processing streams was investigated in this thesis using musical sounds presented to listeners outside of a musical, tonal context.

Chapter 2: Experiments 1, 2, and 3.

Chapter 2 reported on how the perception of auditory roughness in dyads and the stability of their qualitative ratings differ when rated by persons with musical training compared to persons without. Rater consistency (i.e., *intra*-rater reliability, as measured by coding consistency) was higher for familiar dyads tuned to the Western tonal standard than for dyads in an unfamiliar tuning. Rater reliability (i.e., *inter*-rater reliability, as measured by group coherence) was greater for *pure-tone dyads* (two simultaneously sounded sine tones) than for complex-tone dyads, but only for the nonmusician group. Musicians, but not nonmusicians, gave higher roughness ratings to pure-tone dyads in frequency-ratio relationships associated with cognitive dissonance than to dyads in cognitive consonance relationships. These high roughness ratings by musicians for *theoretically* dissonant pure-tone dyads were noteworthy because these dyads lacked harmonic partials responsible for the roughness percept. Musicians' ratings for complex-tone, just-tuned dyads were closer than nonmusicians' to the ratings predicted by a theoretical model of Western consonance (Hutchinson & Knopoff, 1978). Taken together, these findings show the extent to which knowledge of tuning systems influences perceptual judgments of dyads. The process of musical sound-quality assessment is thus shown to be more cognitive in nature, i.e., subject to the life experiences of the perceiver, than has been established in earlier sensory C/D work (Ayers et al., 1980; Butler & Daston, 1968; Guernsey, 1928; Helmholtz, 1885/1954; Kameoka & Kuriyagawa, 1969a, 1969b; Malmberg, 1918; Plomp & Levelt, 1965; Terhardt, 1974b; Van de Geer et al., 1962).

Roughness models that take into account the relations among spectral components, temporal fluctuations, and phase should provide reasonably accurate and similar roughness estimates (Pressnitzer & McAdams, 1999; Terhardt, 1974a), although it cannot be assumed that they do so for musical intervals (D. Cabrera, personal communication, May, 2008). The experiment reported in Chapter 2 compared the variability of listeners' ratings to those produced by two software-based auditory roughness analyzers. The analyzers' ratings correlated more strongly with listeners' ratings of just-tuned than of microtuned dyads, although this was most likely due to decreased rater reliability for microtuned dyads. This observation supports the notion that an internalized system of tonality influences the judgment of dyad roughness. Chapter 2 aimed to assist psychoacousticians in developing auditory models, analyzers, and devices (e.g., hearing aids, cochlear implants) by identifying those signal elements that account for the greatest variability among raters.

Chapter 3.

Chapter 3 reported on the use of a novel/familiar recognition method to explore the cognitive underpinnings of the distinction between consonance and dissonance. I manipulated the sensory and cognitive (music-theoretical as reflected in frequency-ratio relationship) properties of pure-tone, just-tuned dyads to track their retention in auditory STM. The results revealed accurate STM for pure-tone dyads well beyond the previously described limit of 30 s for single tones (Winkler et al., 2002). Correct recognition at retention periods of 48 s was observed for many dyads, despite the memory-degrading effects of interference from incoming dyads, and in the absence of likely mnemonics such as the ability to label the items. Differential memory for dyads sorted by C/D classification was found, but the effects of sensory and cognitive properties on STM were not clear or systematic. Intervals such as Major 2^{nd}s and minor 2^{nd}s with salient sensory attributes, as well as presumably cognitively simple intervals such as octaves, showed no memory advantage over other intervals. I separately analyzed data from musicians and nonmusicians. Musicians showed better dyad recognition overall and smaller performance differences across interval classes than did nonmusicians. I observed that certain categorical subsets showed poor recognition accuracy at moderate retention periods but good accuracy at longer retention periods. This observation suggested that more than one strategy or mechanism might be employed in STM for dyads, mediated by the amount of time afforded for processing.

Conclusions drawn from the work in Chapter 3 were considered preliminary for two reasons. (1) Pure-tone dyads are perhaps more readily distinguished by their sensory attributes than by their cognitive relevance, given that typical musical instruments do not produce them. (2) The nonsystematic pattern of dyad recognition could have been an artifact of the experimental design. I conducted the experiments presented in Chapter 4 to address these concerns.

Chapter 4: Experiments 1 and 2.

The work presented in Chapter 4 used the same experimental design as reported in Chapter 3 but presented more ecologically valid stimuli — complex-tone dyads — in both familiar just-tuning and in unfamiliar micro- (quartertone) tuning. As seen with pure-tone dyads, auditory STM was robust and accurate beyond 30 s retention, persisting to 48 s for several interval classes. Musicians achieved higher recognition scores than nonmusicians for both just- and microtuned dyads. The observation that STM for some classes of dyads was somewhat poor at moderate retention periods and improved at longer retention periods was repeated. I again found differential memory for dyads assigned to sensory and cognitive C/D classes, but as with pure-tone dyads, no obvious signal properties dominated recognition accuracy as a function of time. Dyad retention in auditory STM was shown to degrade over time in a nonsystematic manner, when viewed from the aspect of obvious sensory or music-theoretic signal attributes. Short-term memory for microtuned dyads was essentially the same as that for just-tuned dyads in terms of overall accuracy. This latter finding suggested that categorized musical interval exemplars in long-term memory were not used to facilitate accurate STM recognition.

Chapters 3 and 4 inform and discuss auditory STM research by mapping the fate of dyad memory traces over time. Identifying those features that contribute to robust signal strength and advancing the notion that certain features have optimal periods for availability will assist in the development of improved models of auditory learning, memory, and processing stages.

Contributions to Knowledge from Chapter 2

This thesis contributes to theories of C/D perception by providing modern, comprehensive data on listener roughness evaluation, while systematically taking knowledge-based cognitive differences (due to number of years of formal musical training) into account. It advances understanding of the psychophysical scaling of auditory roughness by identifying and discussing sources of rater variance linked to musical training. It provides behavioral evidence to help define the features and qualities of musical intervals that represent points of divergence between musicians and nonmusicians in their perceptual judgments.

Raters' roughness judgments were more consistent and reliable for familiar stimuli (just-tuned dyads) than for unfamiliar stimuli (microtuned dyads). Tests of inter- and intra-rater reliability showed that the level of agreement was not uniform — for individuals or for group members (musician or nonmusician). Rating consistency depended upon systemically diverse factors such as the frequency separation between two tones and the musical tuning system. The contribution to sensory C/D perception from cognitive processes was evidenced by musicians' tendencies to skew roughness ratings for pure-tone dyads in the direction of the intervals' cognitive C/D values. Nonmusicians were more sensitive than musicians to differences between the fundamental frequencies of each dyad but, perhaps as a corollary, nonmusicians were less sensitive to the roughness contribution from harmonic partials. Thus this study contributed behavioral support to findings from neurobiology indicating that nonmusicians have lower acuity to roughness cues from partials than do musicians (Lee, Skoe, Kraus, & Ashley, 2009). This thesis contends that the distinction between sensory consonance and sensory dissonance may be comparatively reduced in nonmusicians, who comprise the majority of music listeners.

Auditory analyzers.

Chapter 2's report on auditory roughness scaling presents findings that promote the development of more accurate algorithms for describing roughness perception. This in turn assists in the development of audio analyzers, music perception simulators, and hearing devices.

These data quantified the reliability and consistency of listeners' ratings, informing their subsequent comparison with ratings provided by objective auditory analyzers. Consistent with Rabinov and Kreiman's (1995) with voice quality ratings, I observed higher inter-rater agreement for items at the extremes of the scale (i.e., the roughest and smoothest dyads).

The correlation between one analyzer used here (Roughness DW; see Chapter 2 for a complete description) and listeners' ratings was higher for pure-tone than for complex-tone dyads, suggesting that the analyzer's computations placed too much weight on the frequency separation between two tones. Compared to the correlations between listener ratings of just-tuned dyads and both analyzers used here, correlations from microtuned dyads were much lower, with one exception. There was a moderate correlation between musicians' microtuned dyad ratings and those of the Spectral Roughness Analyzer (SRA). Apparently both the SRA and musicians were sensitive to a roughness component of dyads that nonmusicians overlooked. In terms of advancing knowledge, these findings exposed differences between populations distinguished by life experiences, differences that can determine the degree of a roughness model's goodness-of-fit to behavioral data.

Contributions to Knowledge from Chapters 3 and 4

Sensory and cognitive consonance/dissonance processing.

This thesis contributes to the understanding of how bottom-up (feature-driven) and top-down (knowledge-driven) processes give rise to perceived consonance or dissonance of vertical (simultaneously sounded, as opposed to sequentially sounded) musical intervals. The findings presented in Chapters 3 and 4 did not reveal a linear, systematic picture of the time course of stimulus-feature availability in auditory STM. Nor did the findings reveal memory advantages for "natural intervals" (Burns & Ward, 1982; Schellenberg & Trehub, 1996) — tones whose frequencies are related by simple integer ratios. Yet the picture that emerged showed robust and accurate dyad STM well beyond the hypothesized retention duration of 30 s. Although recognition memory for dyads classified according to their C/D attributes was nonsystematic over time, one particular observation recurred. In all three experiments (with three sets each of stimuli and participants), a few C/D classifications showed poor recognition accuracy at moderate periods, yet good recognition of the same classes at longer periods. This repeated observation suggests a phenomenological, rather than a methodological, cause. Perhaps more than one recognition strategy or mechanism serves auditory STM for dyads, mediated by task difficulty and by event properties that are unclear from traditional C/D perspectives. The moderate retention period used in this thesis — 24 s — could mark a point where, for example, a recognition strategy driven by an event's sensory elements has timed out, while the capacity to access the event's conceptual "meaning" is still ongoing. Future work is needed to identify the time course of these processing strategies and those properties of musical sounds that distinguish among them.

Based on the dual nature of C/D processing — sensory and cognitive —, and considering two theories of memory processing, I proposed more than one hypothesis predicting possible outcomes of the STM experiments. Each hypothesis centered on those features of the stimulus that would best support accurate STM. The actual outcomes reported in this thesis were apparently driven by something subtler than the interaction between retention period and stimulus feature availability.

The salient sensory properties of dyads were expected to promote increased processing efficiency and subsequently higher memory accuracy by eliminating the need to access any stored knowledge base of musical intervals (Goldstone & Barsalou, 1998). The thesis work did not support this hypothesis. Relative sensory dissonance within stimulus sets did not appear to be used as a cue to accurate recognition. In the just-tuned dyad sets, sensory dissonance was recognized with the same degree of accuracy as sensory consonance. In the case of microtuned dyads where all of the stimuli were very similar in their sensory properties, the results looked no different than seen in just-tuned sets having broader ranges of sensory dissonance.

Winkler and Cowan (2005) proposed that auditory memory based on stimulus features is best served by long-term familiarity with the global auditory context. Their work showed that the formation of durable auditory memories is assisted by acoustic regularities or frequency-of-occurrence of events in the global auditory scheme. The memory experiments reported in this thesis balanced the frequency-of-occurrence of the musical intervals so that no global, tonal context could be established. In this sense, the memory work of this thesis had only a tangential relationship to the experiment of Winkler and Cowan. Nevertheless, the most cognitively consonant intervals used here are the most prevalent in Western tonal music (Krumhansl & Kessler, 1982); therefore I hypothesized that these could show robust persistence in STM. The data reported in this thesis did not support the hypothesis that interval prevalence outside of the experimental setting would assist recognition memory. Cognitively consonant dyads were recognized with no greater accuracy than cognitively dissonant intervals — those in the most complex integer-ratio relationships of the just-tuned dyads described in Chapters 3 and 4 (minor 7ths, minor 2nds, tritones). Neither were unfamiliar, microtuned intervals recognized with substantially less accuracy than familiar, just-tuned intervals.

Assuming that conceptual knowledge of how an item functions outside of its experimental context influences STM performance, I proposed an alternative hypothesis — cognitive dissonance would boost STM accuracy. Bigand and Tillmann (2005) showed increased neural activation in the inferior frontal areas for musically "deviant" targets that violated listeners' tonal expectancies. Their work showed that neural activation was particularly strong for the most cognitively dissonant targets in their stimulus sets. Taken together with evidence from language studies, their finding suggested that deviants, or less frequently encountered events, "require more neural resources than processing of more familiar or prototypical stimuli." I hypothesized that if this were true, additional neural resources devoted to dissonance processing could promote accurate STM for cognitively dissonant (i.e., large-integer ratio) dyads.

As noted above, the work of this thesis did not uncover a consistent STM advantage for stimulus features along axes of cognitive or sensory C/D that was independent of retention period, and therefore this hypothesis was not supported.

Auditory short-term memory duration.

This thesis contributes to memory research by demonstrating that (what we presume to be) pure auditory STM for dyads is robust and accurate beyond the previously reported duration of 30 s. (Some of the musicians whose data are reported in this thesis may have been able to rehearse, visualize or label some of the dyads they encountered. However, post-experimental self-reports indicated that they did not employ these cues. Based on the observed pattern of responding, any alternate mnemonics were not consistently reliable tools. Nonmusicians presumably lacked any ability to visualize or label dyads, by self-report and by the extremely low chance that they were musicians in disguise!)

In the current STM experiments of Chapters 3 and 4, multiple presentations of each dyad that could have promoted long-term storage were not used. Only the trace from a single, novel presentation was available to facilitate later recognition. It was unlikely that a single presentation or a short subsequence of dyads was stored in LTM. Had that been the case, dyad recognition at the longest retention period (56.25 s) would have been just as accurate as recognition at other periods. Aside from a few exceptions, performance at 56.25 s was at chance. (Data at this period were not analyzed due to too few exemplars and lack of statistical power.) This observation confirmed that decay over time and interference from incoming sounds were indeed detrimental to STM recognition accuracy. Although STM for a single pitch fades by 30 s, this thesis shows that despite decay and interference effects STM for musical dyads from both familiar and unfamiliar tuning systems is robust and accurate beyond this duration.

Musical expertise.

This thesis advances knowledge of perceptual and cognitive differences among listeners as determined by their degree of musical training. It supports a growing body of research suggesting that higher-level cognitive processes mediate auditory perception (Allen, Kraus, & Bradlow, 2000; Kauramaki, Jaaskelainen, & Sams, 2007; Strait, Kraus, Parbery-Clark, & Ashley, 2010). It presents behavioral data that inform recent neurophysiological findings on differences between musicians and nonmusicians.

Musicians are known to be faster and more sensitive than nonmusicians at discriminating among and categorizing the sounds of voice (Bergan & Titze, 2001; Kreiman et al., 1994), piano (Bigand, Madurell, Tillmann, & Pineau, 1999; Regnault et al., 2001; Schön et al., 2005) and harmonic tone complexes (Burns & Houtsma, 1999; Lee et al., 2009; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005; Zatorre & Halpern, 1979). These differences are attributed in part to musicians' higher acuity in analyzing and describing the physical properties of sounds. Musicians' increased abilities to isolate sound features were supported by evidence in the psychophysical scaling task of Chapter 2, which demonstrated their higher sensitivity to harmonic partials, especially when attending to musical intervals in familiar tunings. The advantage in auditory processing speed and sensitivity that musicians have displayed in other experimental settings (cited above) does not translate, however, into a remarkably greater advantage over nonmusicians for nonverbal auditory STM persistence, as shown in Chapters 3 and 4. Although musicians made fewer recognition errors and displayed higher confidence than nonmusicians in distinguishing novel from familiar trials, the difference was not so large as to declare that musical training develops a kind of superior auditory STM that might confer a practical advantage. Short-term memory performance for both groups in the current experiments supported the presumption that the procedure tapped Crowder's (1993) so-called "pure auditory STM" by not reliably involving long-term associations. Under conditions where alternate mnemonic encoding strategies such as verbal rehearsal can be employed (presumably not employed here), musicians are known to have an advantage over nonmusicians in STM capacity (Chan, Ho, & Cheung, 1998; Franklin et al., 2008; Ho, Cheung, & Chan, 2003; Tierney & Pisoni, 2004), but that advantage was not observed here.

Novel Contributions

The principal novel contribution presented in this thesis is the use of a cognitive process — auditory short-term memory — to explore underlying psychological distinctions among musical intervals presented outside of a musical context. A second novel contribution, specific to behavioral C/D studies, includes the systematic segregation of listeners into two groups according to level of formal musical training. Seminal early C/D studies (Ayers et al., 1980; Butler & Daston, 1968; Guernsey, 1928; Helmholtz, 1885; Kameoka & Kuriyagawa, 1969a, 1969b; Malmberg, 1918; Plomp & Levelt, 1965; Van de Geer et al., 1962; Terhardt, 1974a) were conducted well before neurophysiological evidence highlighted the nontrivial perceptual and auditory processing differences between musicians and nonmusicians (Brattico et al., 2009; Foss et al., 2007; Minati et al., 2009; Passynkova, Neubauer, & Scheich, 2007; Regnault et al., 2001; Schön et al., 2005). Earlier C/D work frequently omitted one group or another (Green et al. 2008; Itoh, Suwazono, & Nakada, 2003; Koelsch, Fritz, v. Cramon, Müller, & Friederici, 2006; Passynkova et al., 2007; Passynkova, Sander, & Scheich, 2005; Tillmann, Janata, & Bharucha, 2003; Tufts et al., 2005; Van de Geer et al., 1962; Vos, 1986) or indiscriminately analyzed data from both populations (Butler & Daston, 1968; DeWitt & Crowder, 1987; Geary, 1980; Malmberg, 1918; Schellenberg & Trainor, 1996). The careful segregation of populations in pursuit of accurate C/D processing data may thus be considered a timely novel contribution of this thesis to the field of music cognition.

A contribution of this thesis, taken together with neurophysiological work performed by others in the last decade, is to advance a better understanding of the extent to which C/D perception depends on the life experiences of the perceiver. This line of investigation culminates in the proposal that the more precise term "cognitive consonance" be adopted by psychoacousticians and music cognition researchers to replace Terhardt's (1984) term "musical consonance," which he defined as "sensory consonance plus harmony." The new term "cognitive consonance" will assist researchers in describing similar percepts from both musical and nonmusical sounds. The term is understood to be distinct from "sensory consonance" in that it refers to the contribution to C/D perception from higher-level cortical processing as shaped by an individual's life experiences with sounds — musical and otherwise. Cognitive consonance refers to sensory "pleasantness" as defined by Helmholtz (1885/1954) and Terhardt (1984), and the hierarchical factors of tone-affinity, root-note relationships, and frequency of occurrence and usage in the auditory environment.

An additional novel contribution includes the introduction of a new protocol for C/D assessment. Listener evaluations of musical intervals in earlier decades were confounded by inconsistent methodology and the likelihood of distortions in the signal path (Krumhansl, 1991). Chapter 2's experiment on perceptual roughness was motivated in part by the opportunity to collect evaluative C/D data under a stricter experimental protocol, using advanced audio technology not available in earlier decades. In each of this thesis's experiments, a linear digital audio signal path and high-quality headphones replaced the nonlinear devices used in the vast majority of the seminal C/D work such as: analog tape machines, tuning forks, resonators, and/or loudspeakers (Ayers et al., 1980; Butler & Daston, 1968; Guernsey, 1928; Helmholtz, 1885; Kameoka & Kuriyagawa, 1969a, 1969b; Malmberg, 1918; Plomp & Levelt, 1965; Van de Geer et al., 1962). These new upgrades circumvented the sound coloration that most likely contaminated (to a greater or lesser extent) data from prior decades. This included amplitude normalization of all stimuli to ensure that each dyad was perceived as equally loud. In addition, stimuli were presented at a listening level below that which induces distortion by-products in the inner ear (Clack & Bess, 1969; Gaskill & Brown, 1990; Plomp, 1965). (Please note that I am not assuming earlier work by others necessarily used stimuli that were not amplitude-normalized or level-controlled, but am instead noting that if they did so, it was not reported.)

The work in Chapter 2 used a finer degree experimental control than typically reported in earlier psychophysical scaling measures of C/D. Psychophysical measures must control for statistical learning so that participants cannot predict the quality and magnitude of an upcoming stimulus (Ward, 1987). The scaling experiment reported in Chapter 2 randomized stimulus order for each participant to reduce the overall error that accumulated from sequence effects. In addition, a continuous, high-resolution scale was used instead of a discrete low-resolution scale in order to increase inter-rater correlations (Kreiman, Gerratt, & Ito, 2007). I used these and the signal path upgrades cited above to help ensure that the present findings would be optimally useful to future modelers of dissonance.

Statistical tests of intra- and inter-rater reliability, familiar to voice quality researchers (Kreiman et al., 2007), were used as novel measures for C/D assessment of musical intervals. When raters fail to use scales consistently or fail to agree in their ratings, perceptual assessments lose their validity and usefulness outside of the laboratory. Chapter 2's report on rater consistency and agreement is a novel C/D approach that promotes a better understanding of the stability of listeners' internal standards and the relative ease with which they isolate the attribute under test — auditory roughness, in this case.

Chapter 2 compared listeners' perceptual evaluations of musical intervals with ratings provided by auditory analyzers and two theoretical models. The comparison between subjective and objective ratings has been made in voice quality work (Rabinov & Kreiman, 1995), but to my knowledge this is the first time it has been used in a musical interval study. Likewise, this study is not the first to use microtuned intervals in quality evaluations (Ayers et al., 1980; Bucht & Huovinen, 2004; Geary, 1980; Guthrie & Morrill, 1928), but it is the first to do so while also distinguishing between listeners with and without formal musical training and employing this level of methodological control.

Before presenting sequential stimuli to listeners in the auditory short-term memory studies reported in Chapters 3 and 4, I exercised care to control for repetition bias — the likelihood of the listener predicting whether an upcoming trial would be novel or familiar. Unlike a coin toss, where the odds of heads or tails remain an ideal 50% on every trial, presenting an equal number of novel and familiar trials within a finite-trial sequence meant that statistical learning could be used to predict the novel/familiar status of an upcoming stimulus. (With each successive novel trial, for example, the odds increased that a familiar trial would follow because the number of available novel trials was incrementally decreasing.) Even pigeons have demonstrated the automaticity of this bias (Todd & Mackintosh, 1990). The research reported in Chapters 3 and 4 followed their example by creating stimulus blocks with an arrangement of carefully distributed subsequences of novel and familiar trials. The stimuli were ultimately presented in blocks in which the change rate (the odds of an upcoming stimulus being the same status as the previous stimulus) ranged from a near-ideal probability of 53% to a high of 64% (see Chapter 3 for a detailed explanation). Controlling for repetition bias strengthened the outcome of the data analysis.

Future Directions

The experiments presented in this thesis advance theories of auditory perception and cognition. Implications of the findings apply to the fields of auditory learning and memory and psychoacoustic modeling. Future work will aim to address questions raised by these findings.

Auditory memory, decay, and feature extraction.

The finding presented in Chapters 3 and 4 that auditory STM accuracy for dyads is nonsystematic over time suggests ideas for future memory research through the lens of music cognition. Several outstanding questions arose from these data. To what extent does sequence (the temporal order of items or features) affect the encoding of specific stimuli? Recognition memory for a single pitch is strongly influenced by the similarity of intervening tones to the to-be-remembered pitch (Deutsch, 1972a,b). Are errors caused by failure to properly encode novel dyads, or is it unequal rates of forgetting that determine which dyads will be correctly recognized over time? As the STM neural trace of a dyad decays, are some dyad features more likely to persist than others? Is the feature decay random or can the decay time of certain auditory attributes be predicted, as suggested by a study of decay times for visual stimuli (Gold, Murray, Sekuler, Bennett, & Sekuler, 2005)? If all dyads decay at the same rate, differential persistence may reflect differences in the robustness with which dyads are encoded. On the other hand, if all dyads are encoded with more or less the same fidelity, differential persistence is most likely due to differential rates of forgetting. We cannot assume that the rate of forgetting is the same for all stimuli and for all participants (Tierney & Pisoni, 2004; Wickelgren, 1977), and this can be addressed in future work.

Follow-up studies could substitute new pitches (what would be the impact on recognition accuracy if the pitch range of the stimulus set were reduced by presenting intervals all having a single root pitch?). A variety of timbres might be substituted in place of a variety of intervals (what would happen if the interval were held constant to allow recognition processes to access timbral cues?). A factor analysis of the present data distinguishing dyads along an axis other than C/D may reveal individual differences driven by (as of today) an unknown factor. Future work could manipulate this factor in a recognition memory task.

Implicit vs explicit auditory learning.

Mental processes involved in *identification* and *recall*, not addressed in the experiments of this thesis, tap explicit memory for an event (Roediger, 1990; Wagner & Gabrieli, 1998). *Recognition* — the behavioral measure used here — taps priming from past experience; high accuracy typically reflects implicit memory of an item's physical details (Jacoby, 1983; Roediger, 1990). Accurate recognition memory is easier than accurate identification or recall because either explicit or implicit memories may be used (Petrides, 1995; Petrides & Milner, 1982; Yonelinas, 2001). Thus participants in the STM experiments of Chapters 3 and 4 could have used either knowledge of musical intervals, fluency at remembering perceptual features, or some combination of the two to accurately recognize familiar dyads (Yonelinas, 2001). As shown in Chapter 4, the high recognition accuracy for microtuned, unfamiliar dyads suggests that perceptual fluency was the primary recognition strategy used in these experiments. Future work will attempt to confirm and elaborate this idea.

Implicit learning is defined as "experiences [that] remain concealed from consciousness and yet produce an effect which is significant" (Ebbinghaus, 1885/1964 as cited in Roediger, 1990). Amnesic persons — those whose brain injuries impair their abilities to retain new experiences (Roediger, 1990) — display intact recognition accuracy for information acquired through implicit auditory learning (Schacter, Church, & Treadwell, 1994). (Verbal memory in long-term storage is an example of intact implicit memory in amnesiac persons; Church & Schacter, 1994.) Tests of implicit memory tap a level of retention that is below *conscious knowing*, of the type expressed in recall tasks (Petrides & Milner, 1982; Roediger, 1990; Yonelinas, 2001). The distinction between knowledge that is acquired implicitly versus explicitly would allow me to further explore the extent to which memories formed by these two acquisition processes permit correct dyad recognition in STM, using the experimental protocol described in this thesis. If future work showed that amnesic persons demonstrate intact dyad STM processing, it would further distinguish those aspects of auditory memory performance that depend on explicit musical learning.

Individual differences.

The study of individual differences in music perception and cognition will benefit from future in-depth examination of specific causes of heightened auditory acuity in musicians. The experimental paradigms reported in this thesis can be used to further examine perceptual and cognitive differences among subcategories of musicians. To what extent do players of slightly inharmonic instruments (e.g., vibraphones, marimbas) and nonpitched instruments (e.g., hand percussion, many drums) acquire the sensitivities displayed by the harmonic instrument player? Are players of fretted or fixed-pitch instruments (e.g., guitar, piano) less sensitive to dissonance than players of fretless instruments with continuous control over pitch (e.g., violin, cello)? Years of experience attending to the roughness components when tuning notes from bowed instruments, for example, may shape dissonance perception differently than in the listener whose experience of roughness comes from individual tunings he does not control.

A closer look at the participants whose data are reported in Chapter 3 revealed four subgroups based on the amount of musical training and practice: practicing musician, nonpracticing musician, practicing nonmusician, and nonpracticing nonmusician. *Nonpracticing musicians* were those who reported having had five or more years of formal musical training beginning in childhood, but who stopped playing at some point and never resumed. *Practicing nonmusicians* were those with little (< 2 years) or no formal musical training who described themselves as self-taught musicians. At the time of their participation, “practicing nonmusicians” reported playing an instrument, although in one case this was limited to musical sequence programming. With one exception these few participants’ corrected memory scores were in the range of the nonmusician group. (This observation strengthened the cogency of the Queen’s Musical Questionnaire for identifying subgroups based on musical experience.) There were not enough of these types to justify a separate statistical analysis of their data in this case. Although all participants were assigned to one of the two groups reported in the chapter, future work might consider examining differences among various samples in these two populations. A more in-depth assessment of participants’ musical training, practice, and exposure to music would help determine the life stage at which musical training is essential to the formation of auditory pathways supporting finer auditory acuity.

Individual differences based on the amount of musical training were revealed here in the assessment of dyads with an obvious *beating* quality (amplitude modulations produced by two simultaneous tones with a frequency difference of less than 15 Hz). Future studies of roughness or sensory dissonance will benefit from a closer examination of this observation. Beating was the prominent feature of the narrowest interval reported in this thesis (the uni+ — two pitches separated by a quartertone, or half of a semitone). Musicians and nonmusicians rate the roughness of the uni+ differently; nonmusicians rate it as much smoother on average than musicians do, although the difference reported in this thesis fell just short of statistical significance. This trend reproduced the finding of Ayers et al. (1980), who also compared musicians’ and nonmusicians’ sensory dissonance ratings of microtuned dyads. Nonmusicians might perceive such an extremely narrow interval as a mistuned unison (the smoothest dyad) while musicians might perceive it as a mistuned m2 (a very rough dyad) (Plomp, 1967). Cognitive differences among persons for such a salient sensory attribute as beating warrant future in-depth study.

Comparative psychology.

Studies on the origin of the C/D distinction will benefit greatly from involving nonhuman mammals. The comparative psychology literature has not concluded that nonhuman animals share the same preference or heightened sensitivity for consonance over dissonance (Fannin & Braud, 1971; Hulse et al., 1995; Izumi, 2000; McDermott & Hauser, 2004; Sugimoto et al., 2010). Typical of C/D research with humans, methodology is often cited in the debate over nonhuman perception of C/D (Lamont, 2005). An early study involving rats, for example, reported that the animals voluntarily "played" samples of consonant organ chords more often than they played the dissonant samples (Fannin & Braud, 1971). This finding was compromised by the authors' perhaps unwitting failure to report that the rat's incapacity for discriminating frequencies below 1 kHz no doubt affected its ability to detect C/D differences among the chords used in the study (Fay, 1974).

Species with mental capacities ranging from those of primates (Mishkin & Delacour, 1975) to pigeons (Todd & Mackintosh, 1990) can readily perform novel/familiar recognition memory tasks, as long as the stimuli are discriminable. The STM paradigm described in Chapters 3 and 4 of this thesis was designed such that only the apparatus need be modified to perform the same experiment with other mammals. The common Mongolian gerbil (*Meriones unguiculatus*) is an ideal species to use with the stimuli described in this thesis. Gerbils display a much lower frequency range of hearing than do rats (Syka, 1996) and presumably could distinguish among the current dyads. Future work comparing humans' STM recognition memory with that of nonhumans can assist in the interpretation of the present findings, as well as inform theories of consonance and dissonance.

Conclusions and Implications

The research reported in this thesis advances knowledge of the sensory and cognitive elements contributing to the distinction between consonance and dissonance in vertical musical harmony. These findings make a unique intellectual contribution in part by showing that nonverbal auditory STM is robust and accurate for substantially longer retention periods than previously reported. This thesis informs theories of music cognition by measuring the impact of formal musical training on listeners' responses to musical sounds. Behavioral data were added to the existing neurophysiological data that have shown enhanced capacities for auditory attention and processing in those listeners with musical training. This thesis refutes the notion that music-theoretical consonance, as expressed by small integer-ratio relationships between frequencies, conveys an innate signal processing advantage over large integer-ratio, dissonant intervals as measured with this STM paradigm.

Bibliography

Alain, C., Arnott, S.R., & Picton, T.W. (2001). Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Human Perception and Performance, 27*(5), 1072-1089.

Allen, J., Kraus, N., & Bradlow, A. (2000). Neural representation of consciously imperceptible speech sound differences. *Perception & Psychophysics, 62*(7), 1383-1393.

Aures, W. (1985). Ein berechnungsverfahren der rauhigkeit. *Acustica, 58*, 268-281.

Ayers, T., Aeschbach, S., & Walker, E.L. (1980). Psychoacoustic and experiential determinants of tonal consonance. *Journal of Auditory Research, 20*, 31-42.

Bacon, S.P. & Viemeister, N.F. (1985). Temporal modulation transfer functions in normal-hearing and hearing-impaired subjects. *Audiology, 24*, 117-134.

Baddeley, A.D. & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation.* (Vol. 8, pp. 47-89). New York: Academic Press.

Bahrick, H.P. & Bahrick, P.O. (1964). A re-examination of the interrelations among measures of retention. *Quarterly Journal of Experimental Psychology 16*(4), 318-324.

Balkwill, L.L. & Thompson, W.F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception, 17*(1), 43-64.

Bergan, C.C. & Titze, I.R. (2001). Perception of pitch and roughness in vocal signals with subharmonics. *Journal of Voice, 15*(2), 165-175.

Bidelman, G.M. & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience, 29*(42), 13165-13171.

Bigand, E. Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 184-197.

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception & Psychophysics, 59,* 1098-1107.

Bigand, E., & Tillmann, B. (2005). Effect of context on the perception of pitch structures. In C. J. Plack, A. J. Oxenham, R. R. Fay, & A. N. Popper (Eds.), *Pitch: Neural coding and perception,* (306-351). New York: Springer-Verlag.

Blood, A.J., Zatorre, R.J., Bermudez, P. & Evans, A.C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, *2*, 382-387.

Boring, E.G. (1942). *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century Company.

Brattico, E., Pallesen, K., Varyagina, O., Bailey, C., Anourova, I., Järvenpää, Eerola, T., & Tervaniemi, M. (2009). Neural discrimination of nonprototypical chords in music experts and laymen: An MEG study. *Journal of Cognitive Neuroscience, 21*(11), 2230-2244.

Bucht, S. & Huovinen, E. (2004). Perceived consonance of harmonic intervals in 19-tone equal temperament. *Proceedings of the Conference on Interdisciplinary Musicology (CIM04),* 1-10.

Bugg, E.G. (1963). A criticism of Leibniz's Theory of Consonance. *The Journal of Aesthetics and Art Criticism, 21*(4), 467-472.

Burdette, G. & Butler, D. (2002). From acoustics to tonpsychologie. In T.S. Christensen (Ed.), *The Cambridge History of Western Music Theory*. New York: Cambridge University Press.

Burns, E.M, & Houtsma, A.J.M. (1999). The influence of musical training on the perception of sequentially presented mistuned harmonics. *Journal of the Acoustical Society of America, 106*(6), 3564-3570.

Burns E.M (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music, 2nd edition* (pp. 215-264). New York: Academic Press.

Burns E.M, & Ward, W.D. (1982). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (pp. 241-269). New York: Academic Press.

Butler, J.W. & Daston, P.G. (1968). Musical consonance as musical preference: A cross-cultural study. *Journal of Experimental Psychology, 79*, 57-66.

Cambouropoulos, E. (1996). A general pitch interval representation: Theory and applications. *Journal of New Music Research, 25*, 231-251.

Cariani, P. (2004). A temporal model for pitch multiplicity and tonal consonance. *Proceedings of the 8th International Conference on Music Perception & Cognition.* 310-314.

Cazden, N. (1945). Musical consonance and dissonance: A cultural criterion. *Journal of Aesthetics and Art Criticism, 4*, 3-11.

Cazden, N. (1958). Pythagoras and Aristoxenos reconciled. *Journal of the American Musicological Society, 11*(2/3), 97-105.

Cazden, N. (1980). The definition of consonance and dissonance. *International Review of the Aesthetics and Sociology of Music, 11*(2), 123-168.

Chan, A.S., Ho, Y.C., Cheung, M.C. (1998). Music training improves verbal memory. *Nature, 396*, 128.

Clack, T.D. & Bess, F.H. (1969). Aural harmonics: The tone-on-tone masking vs. the best-beat method in normal and abnormal listeners. *Acta Oto-Laryngologica, 67*(2), 399-412.

Conover, W.J. (1971). *Practical nonparametric statistics.* New York: Wiley.

Cortina, J.M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104*(2)*,* 163-191.

Cowan, N. Saults, J., & Nugent, L. (1997). The role of absolute and relative amounts of time in forgetting within immediate memory: The case of tone pitch comparisons. *Psychonomic Bulletin & Review, 4*(3)*,* 393-397.

Creel, S.C., Newport, E.L., & Aslin, R.N. (2004). Distant melodies: Statistical learning on nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(5), 1119-1130.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297-334.

Crowder, R.G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 472-478.

Crowder, R.G. (1993). Auditory memory. In S. McAdams & E. Bigand (Eds.). *Thinking in sound: The cognitive psychology of human audition*, (113-145). London: Oxford University Press.

Cuddy, L.L., Balkwill, L.L., Peretz, I., & Holden, R.R. (2005) Musical difficulties are rare: A study of "Tone Deafness" among university students. *Annals of the New York Academy of Science, 1060*, 311-324.

Daniel, P. & Weber, R. (1997). Psychoacoustical roughness: Implementation of an optimized model. *Acustica, 83,* 113-123.

De Baene, W., Vandierendonck, A., Leman, M., Widmann, A., & Tervaniemi, M. (2004). Roughness perception in sounds: Behavioral and ERP evidence. *Biological Psychology, 67*, 319-330.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport, 8,* 919-924.

Demany, L. & Semal, C. (2007). The role of memory in auditory perception. In W.A. Yost, A.N. Popper, & R.R. Fay (Eds.), *Auditory perception of sound sources.* New York: Springer.

Deutsch, D. (1972a). Mapping of interactions in the pitch memory store. *Science, 175*, 1020-1022.

Deutsch, D. (1972b). Effect of repetition of standard and comparison tones on recognition memory for pitch. *Journal of Experimental Psychology, 93*(1), 156-162.

Deutsch, D., Henthorn, T., & Dolson, M. (2004). Speech patterns heard early in life influence later perception of the tritone paradox. *Music Perception, 21,* 357-372.

Dewar, K.M., Cuddy, L.L., & Mewhort, D.J.K., (1977). Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning and Memory, 3*(1), 60-67.

DeWitt, L.A. & Crowder, R.G. (1987). Tonal fusion of consonant musical intervals. *Perception & Psychophysics, 41,* 73-84.

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

Durlach, N., & Braida, L. (1969). Intensity perception: I. Preliminary theory of intensity resolution, *Journal of the Acoustical Society of America, 46*, 372-383.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology.* (H.A. Ruger and C.E. Bussenius, Trans.). New York: Dover. (Original work published 1885).

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309-331.

Fabiani, M. & Donchin, E. (1995). Encoding processes and memory organization: A model of the von Restorff effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 224-240.

Fannin, H.G. & Braud, W.G. (1971). Preference for consonant over dissonant tones in the albino rat. *Perceptual and Motor Skills 32*, 191-193.

Fay, R. R. (1974). Auditory frequency discrimination in vertebrates. *Journal of the Acoustical Society of America, 56*(1), 206-209.

Feldt, L.S., Woodruff, D.J., & Salih, F.A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement, 11*(1), 93-103.

Fishman, Y.I., Reser, D.H., Arezzo, J.C., & Steinschneider, M. (2000). Complex tone processing in primary auditory cortex of the awake monkey. I. Neural ensemble correlates of roughness. *Journal of the Acoustical Society of America, 108*(1), 235-246.

Fishman, Y.I., Volkov, I.O., Noh, M.D., Garell, P.C., Bakken, H., Arezzo, J.C., Howard, M.A., & Steinschneider, M. (2001). Consonance and dissonance of musical chords: Neural correlates in auditory cortex of monkeys and humans. *Journal of Neurophysiology 86,* 2761-2788.

Flax, M., & Ferguson, S. (2007). PsySound3 Roughness DW (version 1.0) [Computer software]. Sydney, Australia. http://psysound.wikidot.com/

Foss, A.H., Altschuler, E.L., & James, K.H. (2007). Neural correlates of the Pythagorean ratio rules. *NeuroReport, 18*(15), 1521-1525.

Franklin, M.S., Rattray, K., Moore, K.S., Moher, J., Yip, C.Y., & Jonides, J. (2008). The effects of musical training on verbal memory. *Psychology of Music, 36*, 353-365.

Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience, 16*(6), 1010-1021.

Fulop, S.A. & Fitz, K. (2006). Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications. *Journal of the Acoustical Society of America, 119*(1), 360-371.

Gaab, N., & Schlaug, G. (2003). The effect of musicianship on pitch memory in performance matched groups. *NeuroReport, 14*(18). 2291-2295.

Gaskill, S.A. & Brown, A.M. (1990). The behavior of the acoustic distortion product, $2f1—f2$, from the human ear and its relation to auditory sensitivity. *Journal of the Acoustical Society of America, 88* (2), 821-839.

Geary, J.M. (1980). Consonance and dissonance of pairs of inharmonic sounds. *Journal of the Acoustical Society of America, 67*, 1785-1789.

Glasberg, B.R. & Moore, B.C.J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research, 47*, 103-138.

Gold, J.M., Murray, R.F., Sekuler, A.B., Bennett, P.J., & Sekuler, R. (2005). Visual memory decay is deterministic. *Psychological Science, 16*(10), 769-774.

Goldinger, S.D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105,* 251-279.

Goldstone, R.L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition, 52,* 125-157.

Goldstone, R.L. & Barsalou, L.W. (1998). Reuniting perception and conception. *Cognition, 65,* 231-262.

Green, A.C., Bærentsen, K.B., Stødkilde-Jørgensen, H., Wallentin, M., Roepstorff, A., & Vuust, P. (2008). Music in minor activates limbic structures: A relationship with dissonance? *NeuroReport, 19*(7), 711-715.

Green, S.B., Lissitz, R.W., & Mulaik, S.A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement, 37*, 827-838.

Green, S.B. & Salkind, N.J. (2005). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (4th ed.). New Jersey: Pearson Prentice Hall.

Greenwood, D.D. (1961a). Auditory masking and the critical band. *Journal of the Acoustical Society of America, 33*(4), 484-502.

Greenwood, D.D. (1961b). Critical bandwidth and the frequency coordinates of the basilar membrane. *Journal of the Acoustical Society of America, 33*(10), 1344-1356.

Greenwood, D.D. (1991). Critical bandwidth and consonance in relation to cochlear frequency-position coordinates. *Hearing Research, 54,* 164-208.

Greenwood, D.D. "Plomp&Levelt vs. Zwicker&Fastl - Little evidence of conflict." Online posting. 5 Oct. 1999. Auditory List Archives. <http://www.auditory.org/postings/>.

Guernsey, M. (1928). The role of consonance and dissonance in music. *American Journal of Psychology, 40,* 173-204.

Guirao, M. & Garavilla, J.M. (1976). Perceived roughness of amplitude-modulated tones and noise. *Journal of the Acoustical Society of America, 60*(6), 1335-1338.

Guthrie, E.R. & Morrill, H. (1928). The fusion of non-musical intervals. *American Journal of Psychology, 40*(4), 624-625.

Habib, R., McIntosh, A.R., Wheeler, M.A., & Tulving, E. (2003). Memory encoding and hippocampally-based novelty/familiarity discrimination networks. *Neuropsychologia, 41,* 271-279.

Heathcote, A. (2003). Item recognition memory and the receiver-operating characteristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition 29*(6), 1210-1230.

Helmholtz, H.L.F. von (1954). *On the sensations of tone as a physiological basis for the theory of music* (2nd ed.). (A.J. Ellis, Trans.). New York: Dover. (Original work published 1885).

Helmholtz, H.L.F. von (1995). On the physiological causes of harmony in music. In D. Cahan (Ed.). *Science and culture: Popular and philosophical essays.* Chicago: University of Chicago Press. (Original work published 1873).

Helson, H. (1964). *Adaptation level theory: An experimental and systematic approach to behavior.* New York: Harper & Row.

Hintzman, D.L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*(4), 528-551.

Hintzman, D.L., Caulton, D.A. & Levitin, D.L. (1998). Retrieval dynamics in recognition and list discrimination: Further evidence of separate processes of familiarity and recall. *Memory & Cognition 26*(3), 449-462.

Ho, Y.C., Cheung, M.C., & Chan, A.S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology, 17*(3), 439-450..

Hoeldrich, R. (1999). An optimized model for the objective assessment of roughness sensations in vehicle noise. *Proceedings of the AVL Sound Engineering Conference, Institute of Electronic Music*. 2-9.

Hulse, S.H., Bernard, D.J. & Bratten, R.F. (1995). Auditory discrimination of chord-based spectral structures by European starlings (*Sturnus vulgaris*). *Journal of Comparative Psychology 111*, 3-13.

Hutchinson, W. & Knopoff, L. (1978). The acoustic component of Western consonance. *Interface, 7,* 1-29.

Iacobucci, D. & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence. *Journal of Consumer Psychology, 13*(4), 478-487.

Itoh, K., Suwazono, S., & Nakada, T. (2003). Cortical processing of musical consonance: An evoked potential study. *NeuroReport, 14*(18), 2303-2306.

Izumi, A. (2000). Japanese monkeys perceive sensory consonance of chords. *Journal of the Acoustical Society of America 108* (6), 3073-3078.

Jacoby, L.L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Behavior, 22,* 485-508.

Jacoby, L.L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*(3*),* 306-340.

Johns, E.E. & Mewhort, D.J.K. (2002). What information underlies correct rejections in short-term recognition memory? *Memory & Cognition, 30*(1), 46-59.

Kærnbach, C., & Schlemmer, K. (2008). The decay of pitch memory during rehearsal (L). *Journal of the Acoustical Society of America, 123*(4), 1846-1849.

Kameoka, A., & Kuroyagawa, M. (1969a). Consonance theory: Part I. Consonance of dyads. *Journal of the Acoustical Society of America, 45*(6), 1451-1459.

Kameoka, A., & Kuroyagawa, M. (1969b). Consonance theory: Part II. Consonance of complex-tones and its calculation method. *Journal of the Acoustical Society of America, 45*(6), 1460-1469.

Kauramäki, J., Jääskeläinen, I.P., & Sams, M. (2007). Selective attention increases both gain and feature selectivity of the human auditory cortex. PLoS One 2, e909.

Keller, T.A., Cowan, N., & Saults, J.S. (1995). Can auditory memory for tone pitch be rehearsed? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21,* 635-645.

Kishiyama, M.M., & Yonelinas, A.P. (2003). Novelty effects on recollection and familiarity in recognition memory. *Memory & Cognition, 31*(7), 1045-1051.

Koelsch, S., Fritz, T., v. Cramon, Y., Müller, K. & Friederici, A.D. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping, 27*, 239-250.

Kormi-Nouri, R., Nilsson, L.G., Ohta, N. (2000). The novelty effect: Support for novelty/encoding hypothesis. *International Journal of Psychology, 35*(3-4), 196-197.

Kreiman, J., Gerratt, B.R., & Berke, G.S. (1994). The multidimensional nature of pathologic vocal quality. *Journal of the Acoustical Society of America, 96*(3), 1291-1302.

Kreiman, J., Gerratt, B.R., & Ito, M. (2007). When and why listeners disagree in voice quality assessment. *Journal of the Acoustical Society of America, 122*(4), 2354-2364.

Kreiman, J., Gerratt, B.R., Kempster, G.B., Erman, A., & Berke, G.S. (1993). Perceptual evaluation of vocal quality: Review, tutorial, and a framework for future research. *Journal of Speech and Hearing Research, 36*(1), 21-40.

Krumhansl, C.L. (1990). *Cognitive foundations of musical pitch*. London: Oxford University Press.

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

Krumhansl, C.L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*(1), 159-179.

Krumhansl, C.L., & Kessler, E.J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*(4), 334-368.

Lamont, A.M. (2005). What do monkey's music choices mean? *Trends in Cognitive Sciences, 9* (8), 2359-361.

Lee, K.M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *Journal of Neuroscience, 29*(18), 5832-5840.

Leman, M. (2000). Visualization and calculation of the roughness of acoustical musical signals using the synchronization index model (SIM). *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00),* 1-6.

Link, S. (2005). The structure of context. In J.S. Monahan, S.M. Sheffert, & J.T. Townsend, (Eds.). *Fechner Day 2005: Proceedings of the Twenty-First Annual Meeting of the International Society for Psychophysics,* (169-174).

Lockhead, G.R. (2004). Absolute judgments are relative: A reinterpretation of some psychophysical ideas. *Review of General Psychology, 8*(4), 265-272.

Macmillan, N.A. & Creelman, C.D. (2005). *Detection theory: A user's guide* (2nd ed.). New York: Lawrence Erlbaum Associates.

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience, 18*(2), 199-211.

Malmberg, C.F. (1918). The perception of consonance and dissonance. *Psychology Monographs, 25*, 93-133.

Mashinter, K. (2006). Calculating sensory dissonance: Some discrepancies arising from the models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff. *Empirical Musicology Review, 1,* 65-84.

Massaro, D.W. (1970). Retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology, 83*(1), 32-39.
Mayer, A.M. (1894). Researches in acoustics. *Philosophical Magazine, 5*(11), 37-226, 259-288.
McDermott, J. & Hauser, M. (2004). Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition, 94,* B11-B21.
McGraw, K.O., & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30-46.
McKinney, M.F., Tramo, M.J., & Delgutte, B. (2001). Neural correlates of musical dissonance in the inferior colliculus. In A.J.M. Houtsma et al. (eds.) *Physiological and psychophysical bases of auditory function.* Maastricht: Shaker Publishing.
Medin, D.L. & Schaffer, M.M. (1978). Context theory of classification learning. *Psychological Review 85,* 207-238.
Minati, L., Rosazza, C., D'Incerti, L., Pietrocini, E., Valentini, L., Scaioli, V., Loveday, C., & Bruzzone, M.G. (2009). Functional MRI/Event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. *NeuroReport, 20,* 87-92.
Mondor, T.A., & Morin, S.R. (2004). Primacy, recency, and suffix effects in auditory short-term memory for pure tones: Evidence from a probe recognition paradigm. *Canadian Journal of Experimental Psychology, 58*(3), 206-219.
Moore, B.C.J. (2003). *An introduction to the psychology of hearing* (5th ed.). San Diego: Academic Press.
Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125*(6)*,* 826-859.
Nairne, J.S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*, 53-81.
Neuhaus, C., Knösche, T.R., & Friederici, A.D. (2006). Effects of musical expertise and boundary markers on phrase perception in music. *Journal of Cognitive Neuroscience, 18*(3), 472-493.
Nosofsky, R.M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition 14*(4), 700-708.
Nosofsky, R.M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance 17*(1), 3-27.
Nosofsky, R.M. & Zaki, S.R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition 28*(5), 924-940.
Nosofsky, R.M., & Zaki, S.R. (2003). A hybrid-similarity exemplar model for predicting distinctiveness effects in perceptual old-new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1194-1209.
Nunnally, J.C. & Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.). New York:

McGraw-Hill.
Parasuraman, R. & Beatty, J. (1980). Brain events underlying detection and recognition of weak sensory signals. *Science, 210,* 80-83.
Parasuraman, R., Richer, F., & Beatty, J. (1982). Detection and recognition: Concurrent processes in perception. *Perception & Psychophysics, 31*(1), 1-12.
Parncutt, R. (2006). Commentary on Keith Mashinter's "Calculating sensory dissonance: Some discrepancies arising from the models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff." *Empirical Musicology Review, 1*(4), 201-203.
Passynkova, N., Neubauer, H., & Scheich, H. (2007). Spatial organization of EEG coherence during listening to consonant and dissonant chords. *Neuroscience Letters, 412,* 6-11.
Passynkova, N., Sander, K., & Scheich, H. (2005). Left auditory cortex specialization for vertical harmonic structure of chords. *Annals of the New York Academy of Sciences, 1060,* 454-456.
Pastore, R.E., Crawley, E.J., Berens, M.S., & Skelly, M. (2003). "Nonparametric" *A*' and other modern misconceptions about signal detection theory. *Psychonomic Bulletin & Review, 10*(3), 556-569.
Patel, A.D. (2008). *Music, language, and the brain*. New York: Oxford University Press.
Peretz, I., Gaudreau, D., & Bonnel, A.M. (1998). Exposure effects on music preference and recognition. *Memory & Cognition, 26,* 884-902.
Peretz, I., & Zatorre, R.J. (2005). Brain organization for music processing. *Annual Review of Psychology, 56*, 89-114.
Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologica, 20*(3), 249-262.
Piston, W. (1978). *Harmony.* New York: W. W. Norton.
Plomp, R. (1965). Detectability threshold for combination tones. *Journal of the Acoustical Society of America, 37*(6), 1110-1123.
Plomp, R. (1967). Beats of mistuned consonances. *Journal of the Acoustical Society of America, 42*(2), 462-474.
Plomp, R., & Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America, 38,* 548-560.
Plomp, R. & Steeneken, H.J.M. (1968). Interference between two simple tones. *Journal of the Acoustical Society of America, 43,* 883-884.
Porres, A.T. & Jônatas, M. (2007). A roughness model in Pd for an adaptive tuning patch controlled by antennas. *Proceedings of the 7th International Conference on New Interfaces for Musical Expression.* 363-366.
Pressnitzer, D., McAdams, S. (1999). Two phase effects in roughness perception. *Journal of the Acoustical Society of America, 105*, 2773-2782.
Pressnitzer, D. & McAdams, S., Winsberg, S., & Fineberg, J. (2000). Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. *Perception & Psychophysics, 62*, 66-80.
Prünster, G., Fellner, M., Graf, F., & Mathelitsch, L. (2004). An empirical study on the sensation of roughness. In R. Parncutt, A. Kessler, & F. Zimmer (Eds.). *Proceedings of the Conference on Interdisciplinary Musicology (CIM04),* 1-3.

Rabinov, C.R. & Kreiman, J. (1995). Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *Journal of Speech, Language, and Hearing Research, 38*(1), 26-32.

Rameau, J.P. (1971). *Treatise on harmony.* (P. Gossett, Trans.). New York: Dover. (Original work published 1722).

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related potentials. *Journal of Cognitive Neuroscience, 13,* 241-255.

Roediger, H.L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043-1056.

Ross, D., Choi, J., & Purves, D. (2007). Musical intervals in speech. *Proceedings of the National Academy of Sciences, 104*(23), 9852-9857.

Schellenberg, E.G., & Trainor, L.J. (1996). Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: Tests of adult and infant listeners. *Journal of the Acoustical Society of America*, *100,* 3321-3328.

Schellenberg, E.G., & Trehub, S.E. (1994a). Frequency ratios and the discrimination of pure tone sequences. *Perception & Psychophysics, 56*, 472-478.

Schellenberg, E.G., &Trehub, S.E. (1994b). Frequency ratios and the perception of tone patterns. *Psychonomic Bulletin & Review, 1,* 191-201.

Schellenberg, E.G., & Trehub, S.E. (1996). Natural musical intervals: Evidence from infant listeners. *Psychological Science, 7*(5), 272-277.

Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences, 1060*. 219-230.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*(4), 350-353.

Schneider, A. (1997). "Verschmelzung", tonal fusion, and consonance: Carl Stumpf revisited. In M. Leman (Ed.), *Music, Gestalt, and Computing: Cognitive and Systematic Musicology,* (117-143). Berlin: Springer.

Schön, D., Regnault, P, Ystad, S., & Besson, M. (2005). Sensory consonance: An ERP study. *Music Perception, 23,* 105-117.

Schopenhauer, A. (1969) *The world as will and representation* (2nd ed.). (E.F.J. Payne, Trans.). New York: Dover. (Original work published 1819).

Schwartz, D.A., Howe, C.Q., & Purves, D. (2003). The statistical structure of human speech sounds predicts musical universals. *Journal of Neuroscience, 23,* 7160-7168.

Seashore, C.E. & Mount, G.H. (1918). Correlation of factors in musical talent and training. *Psychology Monographs, 25,* 83.

Seither-Preisler, A., Krumbholz, K., Patterson, R., Johnson, L., Nobbe, A., Seither, S., & Lükenhöner, B. (2007). Tone sequences with conflicting fundamental pitch and timbre changes are heard differently by musicians and nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance, 33*(3), 743-751.

Sethares, W.A. (1993). Local consonance and the relationship between timbre and scale. *Journal of the Acoustical Society of America, 94*, 1218-1228.

Sethares, W.A. (1998). *Tuning, timbre, spectrum, scale.* London: Springer-Verlag.
Shepard, R.N., & Jordan, D.S. (1984). Auditory illusions demonstrating that tones are assimilated to an internalized musical scale. *Science, 226*(4680). 1333-1334.
Shrout, P., & Fleiss, J. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420-428.
Skovenborg, E. & Nielsen, S.H. (2002). Measuring sensory consonance by auditory modelling. *Proceedings of the 5th International Conference on Digital Audio Effects (DAFX-02),* 251-256.
Squire, L.R. (1987). *Memory and brain*. New York: Oxford University Press.
Stevens, J.P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, New Jersey: Erlbaum.
Stewart, N. & Brown, G.D.A. (2004). Sequence effects in the categorization of tones varying in frequency. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*(2), 416-430.
Stewart, N. & Brown, G.D.A. (2005). Similarity and dissimilarity as evidence in perceptual categorization. *Journal of Mathematical Psychology, 49*, 403-409.
Strait, D.L., Kraus, N., Parbery-Clark, A., & Ashley, R., in press. Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*.
Stumpf, C. (1898). *Konsonanz und Dissonanz.* Leipzig: J.A.Barth (Beiträge zur Akustik und Musikwissenschaft, Heft 1).
Sugimoto, T., Kobayashi, H., Nobuyoshi, N., Kiriyama, Y., Takeshita, H., Nakamura, T., & Hashiya, K. (2010). Preference for consonant music over dissonant music by an infant chimpanzee. *Primates, 51,* 7-12.
Syka, J., Rybalko, N., Brozek, G., & Jilek, M. (1996). Auditory frequency and intensity discrimination in pigmented rats. *Hearing Research, 100*, 107-113.
Tenney, J. (1988). *A history of consonance and dissonance*. New York: Excelsior.
Terhardt, E. (1974a). On the perception of periodic sound fluctuations (roughness). *Acustica, 30*, 201-213.
Terhardt, E. (1974b). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America, 55*, 1061-1069.
Terhardt, E. (1984). The concept of musical consonance: A link between music and psychoacoustics, *Music Perception, 1,* 276-295.
Terhardt, E. (2000, February 15). Sensory consonance. Retrieved August 18, 2009, from http://www.mmk.ei.tum.de/persons/ter.html.
Terhardt, E. & Stoll, G. (1981). Skalierung des wohlklangs von 17 umweltschallen und untersuchung der beteiligten hörparameter. *Acustica, 48*, 247-253.
Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research, 161*, 1-10.
Tierney, A., & Pisoni, D.B. (2004). Some effects of early musical experience on sequence memory spans. *Research on Spoken Language Processing.* Bloomington, IN: Indiana University.
Tillmann, B., Janata, P., & Bharucha, J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research, 16,* 145-161.

Todd, I.A., & Mackintosh, N.J. (1990). Evidence for perceptual learning in pigeon's recognition memory for pictures. *Quarterly Journal of Experimental Psychology, 42B,* 385-400.

Trainor, L.J., & Heinmiller, B.M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development, 21*(1)*,* 77-88.

Trainor, L.J. & Trehub, S.E. (1993). What mediates infants' and adults' superior processing of the major over the augmented triad? *Music Perception 11* (2), 185-196.

Tramo, M.J., Cariani, P.A., Delgutte, B., & Braida, L.D. (2003). Neurobiology of harmony perception. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music,* (127-151). New York: Oxford University Press.

Tsai, C.G. (2004). Auditory grouping in the perception of roughness induced by subharmonics: Empirical findings and a qualitative model. *Proceedings of the International Symposium on Musical Acoustics, (ISMA 2004),* 257-260.

Tufts, J.B., Molis, M.R., & Leek, M.R. (2005). Perception of dissonance by people with normal hearing and sensorineural hearing loss. *Journal of the Acoustical Society of America, 118*, 955-967.

Tulving, E. & Kroll, N. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychonomic Bulletin and Review, 2*(3), 387-390.

Tulving, E., Markowitsch, H.J., Kapur, S., Habib, R., & Noule, S. (1994). Novelty encoding networks in the human brain: Positron emission tomography data. *NeuroReport, 5,* 2525-2528.

Van de Geer, J.P., Levelt, W.J.M., & Plomp, R. (1962). The connotation of musical consonance. *Acta Psychologica, 20,* 308-319.

Vassilakis, P.N. (2007). SRA: A web-based research tool for spectral and roughness analysis of sound signals. *Proceedings of the 4th Sound and Music Computing Conference,* 319-325.

Viemeister, N.F. & Fantini, D.A. (1987). Discrimination of frequency ratios. In W.A. Yost & C.S. Watson (Eds.), *Auditory Processing of Complex Sounds,* (47-56). Hillsdale, N.J.: Lawrence Erlbaum.

Vos, J. (1986). Purity ratings of tempered fifths and major thirds. *Music Perception, 3,* 221-258.

Wagner, A.D., & Gabrieli, J.D.E. (1998). On the relationship between recognition familiarity and perceptual fluency: Evidence for distinct mnemonic processes. *Acta Psychologica, 98*, 211-230.

Ward, L.M. (1987). Remembrance of sounds past: Memory and psychophysical scaling. *Journal of Experimental Psychology, 13*(2), 216-227.

Wickens, D.D., Born, D.G., & Allen, C.K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior 2*, 440-445.

Wickens, T.D. (2002). *Elementary signal detection theory.* New York: Oxford University Press.

Wicklegren, W.A. (1977). *Learning and memory.* New Jersey: Prentice Hall.

Wild, J. (2002). The computation behind consonance and dissonance. *Interdisciplinary Science Reviews, 27*(4), 299-302.

Winkler, I. & Cowan, N. (2005). From sensory to long-term memory: Evidence from auditory memory reactivation studies. *Experimental Psychology 52* (1), 3-20.
Winkler, I., Korzyukov, O., Gumenyuk, V., Cowan, N., Linkenkaer-Hansen, K., Ilmoniemi, R.J., Alho, K., & Näätänen, R. (2002). Temporary and longer term retention of acoustic information. *Psychophysiology, 39*, 530-534.
Wolpert, L. (2006). *Six impossible things before breakfast: The evolutionary origins of belief.* New York: Norton & Co.
Yonelinas, A.P. (2001). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General, 130*(3), 361-379.
Zaki, S.R. & Nosofsky, R.M. (2001). Exemplar accounts of blending and distinctiveness effects in perceptual old-new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition 27*(5), 1022-1041.
Zatorre, R.J., & Halpern, A.R. (1979). Identification, discrimination, and selective adaptation of simultaneous musical intervals. *Perception & Psychophysics, 26*(5), 384-395.
Zentner, M.R., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior and Development, 21*, 483-492.
Zmigrod, S. & Hommel, B. (2009). Auditory event files: Integrating auditory perception and action planning. *Attention, Perception, & Psychophysics, 71*(2), 352-362.
Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: Facts and models.* Berlin: Springer-Verlag.