Beyond the Emotional Impact of Dissonance: Inharmonic Music Elicits Greater Cognitive

Interference Than Does Harmonic Music

by

Tanor Bonin

A thesis

presented to the University of Waterloo

in fulfillment of the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2016

© Tanor Bonin 2016

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The present research evaluates whether task-irrelevant inharmonic music produces greater interference with cognitive performance than task-irrelevant harmonic music. Participants completed either an auditory (Experiments 1 and 2) or a visual (Experiment 3) version of the cognitively demanding 2-back task in which they were required to categorize each digit in a sequence of digits as either being a target (a digit also presented two positions earlier in the sequence) or a distractor (all other items). They were concurrently exposed to task-irrelevant harmonic music (judged to be consonant), task-irrelevant inharmonic music (judged to be dissonant), or no music at all as a distraction. The main finding across all three experiments was that performance on the 2-back task was worse when participants were exposed to inharmonic music than when they were exposed to harmonic music. Interestingly, performance on the 2-back task was generally the same regardless of whether harmonic music or no music was played. I suggest that inharmonic, dissonant music interferes with cognitive performance by requiring greater cognitive processing than harmonic, consonant music, and speculate about why this might be.

Acknowledgments

Thank you to Dr Daniel Smilek for his active, encouraging, and enthusiastic mentorship.

Thank you to my undergraduate research assistants Tracy Duncan, Jennifer Chandrabose, Laura Obdeyn, and Katherine Patricia Paus for their diligent work collecting the data for these experiments.

Table of Contents

List of Tables

List of Figures

Introduction

Despite its ubiquity in the musical environment, *dissonance* remains one of the most enigmatic phenomena in music cognition research. Dissonance is a phenomenology characterized by negative affect and a sense of structural instability within the music (Bonin & Smilek, 2015). It is integral to the musical expressions of tension, unrest, and a myriad of negative emotions. As such, it maintains a critical function within Western tonal theory and the creation of musical symbolism as the functional counterpart of musical *consonance*— a phenomenological appraisal of musical *stability, resolution* and *pleasantness* (Costa, Bitti & Bonfiglioi, 2000; Bigand, Parncutt & Lerdahl, 1996; Cook & Fujisawa, 2006; Malmberg, 1918; Zentner & Kagan, 1998; Blood, Zatorre, Bermudez & Evans, 1999; McDermott, Oxenham & Lehr, 2010). Expert composers are seemingly those who can strike the delicate balance between dissonance and consonance to guide the listener through the desired musical landscape (Bidelman & Heinz, 2011; Krumhansl, 1990). As the Grammy award-winning producer and composer Quincy Jones said: "Music in movies is all about tension and release, dissonance and consonance," (Farndale, 2010).

Psychoacousticians have long sought an acoustic signature of dissonant musical stimuli. But psychoacoustic theories of dissonance have focused exclusively on the acoustic *harmonicity* of musical sounds. The ancient Greek Pythagoras suggested that dissonance arises from musical sounds whose constituent frequency components lack divine (i.e., simple integer) relation to one another (Tenney, 1988). In the late 19th century, Hermann von Helmholtz extended this view by suggesting that dissonance arises from the destructive interference patterns of an inharmonic acoustic signal, and that these *beating* and *roughness* interference phenomena irritate the basilar membrane sense organ (Yost, 2008). Several decades later, Plomp and Levelt (1965) provided

empirical support for this hypothesis by outlining what they termed *critical bands* of the basilar membrane. Critical bands were defined as the lower bound for the acoustic frequency intervals that could be effectively transduced along the basilar membrane. The simultaneous presence of two frequencies within a critical band would produce the beating and roughness phenomena proposed by Helmholtz. Such frequency intervals are absent in the simple harmonic sounds we generally experience as "consonant" but are prevalent within the complex inharmonic spectra of dissonant sounds. Thus, researchers established a physiological basis (the *sensory dissonance hypothesis*) for the relation between acoustic inharmonicity and the phenomenology of dissonance.

This model was nevertheless stifled by its own limitations several years later. Ernst Terhardt (1978; 1984) noted that a sensory mechanism could not be used to explain melodic dissonance. Since the sonic constituents of a melody are temporally distinct, they cannot elicit simultaneous stimulation of the basilar membrane or the beating and roughness interference phenomena. So while dissonant melodies often contain inharmonic frequency spectra, this acoustic inharmonicity cannot be related to the phenomenology of dissonance by the physiological model championed by the sensory dissonance hypothesis. Aligned with this criticism, McDermott, Oxenham & Lehr (2010) have recently reported empirical evidence that acoustic inharmonicity predicts even *chordal* dissonance in the absence of beating and roughness. Perhaps disappointingly, the most resilient correlate of musical dissonance is acoustic inharmonicity, an observation first established mathematically by Pythagoras 2600 years ago. How and why acoustic inharmonicity relates to human emotion remains unknown.

I attempted to address this conundrum in culmination of my Bachelor's of Science, yielding what I termed the *source dilemma hypothesis* of dissonance perception (SDH; Bonin,

2

2014). This psychophysical framework predicts that a listener will experience dissonance when a musical stimulus exhibits psychoacoustic properties that produce multiple, incoherent inferences about the auditory environment (Bonin, 2014). Mechanistically, the hypothesis is based on the evolutionary basis for human emotion (e.g., Tooby & Cosmides, 1990; Frijda, 1993; Levenson, 1999) and the principles of auditory scene analysis (Bregman, 1990).

Evolutionary theory proposes that the neurophysiological basis of human emotions has evolved to enable adaptive problem solving within our environment (Tooby & Cosmides, 1990; Frijda, 1993; Levenson, 1999). Our emotive physiology serves not only to produce cognitive and affective assessments of the environment that emphasize its most informative features, but also to metabolically prioritize the behaviours that allows us to respond most effectively to this information (Frijda, 1993; Tooby & Cosmides, 1990). Generally speaking, "pleasant" emotive physiology precipitates rewarding thoughts and feelings following an adaptive response in order to direct our attention, motivation, and behaviours toward maintaining that response. Conversely, unpleasant emotions confront maladaptive situational responses, eliciting intrinsically painful and unsustainable cogitation and affect to direct attention, motivation and behaviours towards an alternative, adaptive response (Levenson, 1999).

If one assumes that musical emotions stem from the same neurophysiological substrates as those that produce emotion more generally (c.f., Blood & Zatorre, 2001), one might hypothesize that musical stimuli contain information content that triggers the organism's environmental problem solving apparatus. One might further suspect that acoustic inharmonicity reliably induces dissonance (a phenomenology characterized by "negative affect" and "a sense of structural instability") by way of representing a problem in the auditory environment. To

understand how such a relation between sound and emotion might form, one needs first to understand how the brain organizes sound.

Auditory scene analysis (ASA) describes the processes by which the brain represents sonic stimulation as auditory perception. These processes allow the brain to derive inferences about the sound sources in its current auditory environment (Bregman, 1990). Physically speaking, sound sources emit longitudinal compression waves of the surrounding air particles. These sound waves are each associated with a characteristic spectro-temporal signature of the sources that created them. Concurrent sound waves are summed and reach the ear as one complex wave. The auditory system is faced with the challenge of parsing this complex sound wave into a representation of auditory objects on the basis of the temporal and spectral signatures of the sound sources that created them.

The system first conducts a series of parametric analyses on the incoming complex sound wave to determine which sensory components most likely originated from the same sound source and should thus be *fused* as a single auditory object in perception, and which components should be *segregated* as perceptually distinct auditory objects because they most likely originated from different sound sources. This process involves the analysis of both simultaneous and sequential sensory components, as a single sound source such as a melody or speech signal naturally varies across time (Bregman, 1990). Several perceptually salient and experimentally verified parameters include: the temporal envelope (Bregman & Pinker, 1978; Dannenbring & Bregman, 1978), spatial position (Moore, 2013), timbre (Caclin, McAdams, Smith, & Winsberg, 2005; Siedenburg, Jones-Mollerup & McAdams, 2015), and harmonicity (DeWitt & Crowder, 1987). Each of these parameters produces a best-estimate of the number of sound sources in the environment. In most situations, the estimates of each parameter are compatible with those of

the others, creating a coherent perception of the auditory environment. In some cases, however, these parameters produce conflicting inferences about the number and type of sound sources in the outer world, creating an incoherent auditory percept. These latter cases provide the crux of the SDH: If the various psychoacoustic parameters of a musical stimulus generate conflicting inferences about the constituents of the natural world (a source dilemma), then the listener will experience dissonance (Bonin, 2014).

By this account, inharmonic music does not elicit dissonance because of its inharmonicity *per se*, but because the source inference this inharmonicity generates is incompatible with those of the other parametric analyses and the otherwise coherent percept of the musical stimulus. Such a conceptualization readily accounts for the persistent correlation between dissonance phenomenology and complex inharmonic frequency spectra (Pythagoras re: Tenney, 1988; Helmholtz, 1863; Kameoka & Kuriyagawa, 1969a, 1969b; Hutchinson & Knopoff, 1978; McDermott, Lehr & Oxenham, 2010) by describing the interrelatedness of several well-studied causal mechanisms in the extant psychological and physiological literatures. This mechanistic specificity gave rise to the counterintuitive prediction that inharmonic music needn't be experienced as dissonant, so long as the perceptual malfunction (source dilemma) it produced could be resolved through manipulations of the music's other psychoacoustic parameters.

To test this prediction for my undergraduate thesis, I designed two experiments in which I manipulated the harmonicity, spatial orientation, and timbres of twenty-four musical stimuli. Each of these manipulations reliably altered the listener's experience of dissonance. Critically, I was able to demonstrate that manipulations of the music's spatial or timbral parameters altered the listener's experience of dissonance without any ancillary changes to the harmonic content of a musical signal. The manipulations were also bi-directionally effective. Not only was it

possible to mitigate the dissonance elicited by inharmonic music through complementary manipulations of its timbral and spectral parameters, it was also possible to enhance the dissonance elicited by harmonic music, by segregating the timbral or spatial parameters of an otherwise perceptually fused musical composition. These results beckoned a conceptualization of dissonance within a multidimensional psychoacoustic space comprising, at minimum, the influences of spatial, timbral, and harmonic psychoacoustic parameters, and provided strong support for the SDH (Bonin, 2014; in Huron, in press, MIT Press).

The SDH has implications for theoretical accounts of the cognitive processing requirements of musical dissonance, as well. If dissonant music is characterized by a perceptual malfunction in the auditory system, then one might expect the perceptual system to redirect cognitive processing to the resolution of that auditory percept, creating a measurable load on cognitive machinery relative to a consonant counterpart assumedly devoid of such a malfunction. This line of reasoning led to my focus for the present thesis. Concretely, the question I posed was: Does dissonant music produce greater cognitive interference than consonant music? A review of the extant cognitive interference literature provided inconclusive evidence regarding whether or not dissonant music might produce more cognitive interference than consonant music.

Bodner, Gilboa and Amir (2007) expected dissonant music to induce greater cognitive interference than consonant music on the basis that the *tension* and *unfulfilled expectations* it creates would produce supra-optimal levels of arousal. This hypothesis assumes that the relation between arousal and performance on cognitive tasks can be represented with an inverted U-shaped curve (cf., Yerkes & Dodson, 1908), with the lowest and highest levels of arousal leading to poorer cognitive performance than the intermediate levels of arousal, which facilitates optimal

cognitive performance. Specifically, the authors suggested that, by violating musical expectations, dissonant music might push arousal levels to the extreme high end of the arousal curve where performance decrements are typically observed (Bodner, Gilboa & Amir, 2007).

Surprisingly, they found no evidence to support their expectations. In fact, under some conditions they found performance to be best while dissonant music was played. Participants in these studies performed *better* on simple cognitive tasks such as the Letter Cancellation Task (LCT) and the Adjective Recall From a Story (ARS) task when exposed to dissonant music compared to consonant music or no music. Additionally, when completing the hardest task (Adjective Recall From a List; ARL) participants performed worse while listening to either consonant or dissonant music compared to completing the task in silence, but exhibited no performance differences between the consonant and dissonant listening conditions. Though contrary to their predictions, the authors interpreted the performance benefits associated with exposure to dissonant music as a result of increased arousal and task engagement. The authors suggested that the dissonant music elicited enough arousal to promote optimal performance in the easier tasks (LCT & ARS), while the consonant music and no music conditions elicited insufficient arousal and suboptimal cognitive performance. Addressing the results of the most difficult task (ARL), they suggested that both consonant and dissonant music elicited too much arousal relative to no music, leading to equally poor performance between the consonant and dissonant conditions and relatively better performance in the no music condition (Bodner, Gilboa & Amir, 2007, pg. 300). A critical shortcoming of this study was that, while the melodic character was retained between the consonant and dissonant music segments, the dissonant excerpts contained greater chordal densities and different spectral ranges than their consonant

counterparts, thus making unclear whether and to what extent the observed results reflect these low-level acoustic disparities or the difference in the listener's phenomenological experience.

Some evidence consistent with the idea that dissonant music might negatively impact performance on specific cognitive tasks relative to consonant music comes from a recent study by Masataka and Perlovsky (2013). Participants in this study listened to consonant or dissonant music while at the same time while naming the colour of neutral (coloured strings of Xs) or incongruently coloured words (e.g., BLUE in red font) in a Stroop task. While musical dissonance did not influence performance on the neutral Stroop trials, participants responded more slowly and less accurately to incongruent Stroop trials when dissonant music was played than when consonant music was played. These findings led the authors to suggest that the interfering effect of musical dissonance manifests *only* when an individual is faced with a task that requires the resolution of incompatible cognitions, such as the incompatible response demands of the word-colour information of incongruent Stroop trials. In other words, according to Masataka and Perlovsky (2013), musical dissonance has a very specific and targeted impact, restrictively hindering performance on tasks that involve a specific type of incompatibility, which they refer to as 'cognitive dissonance' (Masataka & Perlovsky, pg. 5).

While Masataka and Perlovsky's (2013) conclusion that musical dissonance influences only tasks that involve incompatible cognitions is certainly consistent with their findings, there remains the alternative possibility that musical dissonance might have a more general effect on cognitive processing. Specifically, the findings are also consistent with the view that dissonant music has a more general effect on cognitive performance, either via its greater processing demands or its elicitation of supra-optimal arousal, and that this interference is simply more pronounced as the cognitive demands of any concurrent cognitive task increase. According to

this alternative view, musical dissonance should influence performance on any sufficiently demanding cognitive task, even if that task does not involve the specific sort of response selection conflict typified by incongruent trials on the Stroop task. Applying this more general view to the findings reported by Masataka and Perlovsky (2013), musical dissonance would have affected performance on incongruent Stroop trials and not neutral Stroop trials because incongruent trials are more cognitively demanding than neutral trails. It has yet to be shown, however, that dissonant music could impair performance to a greater extent than consonant music on a general cognitive task that does not involve response selection conflict, or, as Masataka and Perlovky (2013) put it, 'cognitive dissonance.'

Lastly, studies of the irrelevant sound effect (ISE; see Banbury, Macken, Tremblay & Jones, 2001; Hughes & Jones, 2001, and Ellermeier & Zimmer, 2014 for reviews) have examined the psychoacoustic properties of sounds that influence primary task completion. A seminal finding from this literature is that unattended steady-state stimuli are far less distracting than their changing-state counterparts. A particularly topical investigation of this phenomenon found that distracting musical stimuli generate larger ISE when performed with staccato articulation than with legato articulation (Schlittmeier, Hellbrück & Klatte, 2008). One related possibility is that dissonant melodic stimuli, by virtue of the more salient state changes among their melodic constituents, might produce greater cognitive interference than their consonant steady-state counterparts.

Building from this literature, the present research investigated whether task-irrelevant dissonant music produces greater interference with concurrent cognitive processing than task-irrelevant consonant music. This interference should be most strongly evident during a sufficiently demanding cognitive task, where the potential effects of source dilemma, arousal,

and-or sensory complexity might be most readily observed. In addition, as an attempt to generalize the findings of Masataka and Perlovsky (2013), the present methodology challenged the assertion that dissonance interferes only with tasks than entail response selection conflict by employing a primary task that required sustained cognitive processing but did not entail response selection conflict. Finally, to address the noted shortcomings of the Bodner, Gilboa and Amir (2007) study, careful consideration was given to control the spectral characteristics of the musical stimuli, manipulating their position on the continuum of consonance and dissonance solely on the basis of their *harmonicity* and leaving otherwise untouched their chordal densities and spectral ranges. Isolating this spectral component allowed for targeted interpretations of the results, and provided an acoustic basis for comparing these results with those of potential future investigations of the ISE and the cognitive effects of dissonant music. Participants' phenomenological appraisals of each stimulus were used to confirm that this acoustic manipulation produced the desired psychological effects.

Participants in these experiments were required to complete either an auditory (Experiments 1 and 2) or a visual (Experiment 3) version of the 2-back task—a sustained cognitively demanding task often used as an indicator of working memory capacity (Owen, McMillan, Laird & Bullmore, 2005). In the 2-back task, participants were presented with a stream of digits and were required to press one response key when the presented digit matched the digit presented two positions earlier in the sequence (i.e., the digit it is a target), and a different response key in all other cases (i.e., the digit is a distractor). While completing this primary task, participants were exposed either to no distractions (no music), task-irrelevant harmonic (consonant) music, or task-irrelevant inharmonic (dissonant) music. Performance on

the primary 2-back task was predicted to be worse when participants were simultaneously

presented with inharmonic music compared to when they were presented with harmonic music.

Experiment 1

Introduction

The purpose of Experiment 1 was to evaluate whether inharmonic music demands greater cognitive processing than harmonic music. Participants were presented a sequence of numbers for the 2-back task in one ear while simultaneously listening to music (either harmonic or inharmonic) in the other ear. Participants were instructed to attend to the numbers of the 2-back task and ignore the music. In the present version of the 2-back task, the sequence of numbers contained infrequent targets, which were defined as a number in the sequence that was also presented two trials earlier in the sequence. All of the remaining numbers in the sequence were distractors. Participants were required to respond to every number, pressing a specific key when a target number was presented and a different key when a distractor number was presented. This allowed measurements of performance accuracy (in terms of sensitivity derived from hits and false alarms), as well as response times to both target and distractor numbers. If inharmonic music demands greater cognitive processing than harmonic music, then performance (in terms of sensitivity and response time) on the 2-back task should be poorer when inharmonic music is simultaneously played than when consonant music is simultaneously played.

While the primary empirical focus was on the differential cognitive demands of harmonic and inharmonic music, I also decided to measure performance on the auditory 2-back task in the absence of any musical distraction. Collection of these data allowed for comparisons between performances on the 2-back task when no music was played and when either harmonic or inharmonic music was played. No a priori predictions were made with regard to these comparisons.

Method

Participants

Thirty undergraduate students (mean age = 20.03 years, SD = 1.87 years; 8 male) from the University of Waterloo were included in the final analysis. The students participated in a thirty-minute experimental session and were compensated with partial course credit. Participants were not selected on the basis of musical training, but the number of years of music lessons ranged from 0 to 20 years (mean = 4.2 years, SD = 4.60 years).

A sample size of thirty participants was predetermined for Experiment 1 before data collection began based on the results of a small pilot study (N = 11). After completing data collection for an initial sample of thirty participants, the data from three participants were excluded due to non-compliance (responding only to target trials, responding always with one key, or prematurely terminating the experiment) and data from three participants were excluded because their accuracy scores fell 2.5 standard deviations below the group mean. As a result, six additional participants were recruited to complete the full counterbalance and reach the predetermined sample size of thirty.

Apparatus

A Python (2.7.9; Van Rossum, 2007) script was written to create the auditory 2-back task, present all primary 2-back task stimuli and distracting musical stimuli, and record all measurement data, including the accuracy of the response (i.e., hits and false alarms) and the response time. Musical stimuli were recorded using Steinberg's Cubase 6 digital audio workstation, the Steinberg HalionSonic SE VST, a Samson Graphite 49 MIDI keyboard, and a Yorkville foot controller.

The experiment was conducted on an Apple Mac Mini with OS X 10.6.6 and a 2.6GHz Core i7 processor. On-screen instructions and prompts for the aesthetic appraisals of the harmonic and inharmonic musical stimuli were presented on a 24" Phillips 244E monitor at a resolution of 1920x1080. Auditory stimuli were delivered through circumaural closed-back headphones (Sony MDR-MA100). The attended number stream and the distracting music stream were quasi-controlled for loudness by equating RMS amplitudes across conditions. Participants listened to the stimuli at comfortable hearing levels and were reminded that they should notify the experimenter if their listening experience became uncomfortable at any time.

Stimuli

Two-back Task. The stimuli for the 2-back task were nine simulated female voice recordings of the spoken numbers 1 through 9 created using Apple's Text to Speech application. The Python program then generated a pseudo-random sequence of these numbers with two constraints: Frist, twenty percent of the numbers in the sequence were the same as the number that was presented two positions earlier in the sequence. These numbers served as the *targets* in the 2-back task. Second, each number was presented once, without repetition, before the first 2-back stimulus sequence occurred. Each participant received a different randomized sequence of the numbers, and it was this sequence that constituted the experiment's primary 2-back task.

Music. The harmonic and inharmonic musical distractors were derivatives of a novel 8'10" piano performance by one of the authors (TB). The performance was conducted to a constant metronome of 70bpm, with various triplet and straight rhythmic permutations of 3/4 and 4/4 time. Beginning in C major, the performance modulated directly to A natural minor at 3'46" and modulated back to C major from 5'36"—5'49". The piece consisted of 6 unique contrapuntal voices (designated by frequency range and harmonic function, see Appendix), and

the number of simultaneous voices varied from 1 to 5 throughout the duration of the piece.

Mindful that particular beat densities and tempos potentiate particular states of arousal or

emotional valence over others (Hevner, 1935; 1937; Peretz, Gagnon & Bouchard, 1998), the

performer varied the tactus of the performance from quarter note pulses at its slowest (857.14ms

SOA) to triplet sixteenth pulses at its fastest (142.86ms SOA). The performance was recorded as

MIDI data in Cubase 6. The original (recorded) MIDI data from this performance constituted the

harmonic stimulus. The MIDI data from the original performance were then copied (including

note velocities and pedal points) and pasted to separate tracks in Cubase 6 (one for each

contrapuntal voice), where systemic pitch shifts were applied to each voice in order to create the

inharmonic music. The Appendix provides a complete list of pitch shifts and interval changes.

Both the harmonic and inharmonic stimuli shared a total frequency range between F0

(21.83 Hz) and E6 (1318.51 Hz). Thus, the two pieces shared every sonic characteristic but their

respective tonalities, with octaves (unisons), major thirds, perfect fifths, major sixths, and major

sevenths of the harmonic performance being performed as minor ninths, minor thirds, tritones

(diminished fifths), minor sixths, and minor sevenths, respectively, in some voices of the

inharmonic version. [1]

The MIDI data for both the harmonic and inharmonic stimuli were then submitted as

triggers to the HalionSonic SE Yamaha S90ES piano sample bank. The HalionSonic SE VST

produces panned stereo output to create a realistic acoustic image of its virtual instruments. In

---

[1] These pitch manipulations resulted in virtually omnipresent chordal inharmonicity within the
Inharmonic stimulus. For example, in the first 1'56" of the piece (34 bars; 132 beats), there was one beat
containing a harmonic interval, and this happened to occur at a brief transition point in the piece where
only two voices sounded. Furthermore, with a harmonic interval prevalence of only ~0.7 %, there is
reason to suspect that these rare events were themselves experienced as "dissonant," as they exhibited low
pitch commonality with the surrounding tones of the continuous inharmonic musical stream in which they
are heard (cf., Bigand, Parncutt, & Lerdahl, 1996; Bigand & Parncutt, 1999).

the specific case of the Yamaha S90ES piano, the lower piano notes are panned to the left of stereo midline and the higher notes are panned to the right of midline. Because the intent was for the musical stimulus to be heard only from the participants' right auditory field, I exported the harmonic and inharmonic performances as mono wave files to ensure that they would retain their full spectral characteristics regardless of where they were panned along the auditory azimuth.

Procedure

After providing written consent, receiving a verbal briefing of the task instructions from the experimenter and reading the on-screen instructions, participants first completed a practice block consisting of 15 trials and 3 targets. The practice trials would present an error tone (Apple "blow.aiff") if they made a mistake during the practice trials; this error tone was not present during the actual experiment. After completing the practice trials the participants were prompted to ask the experimenter for clarification or to ask any remaining questions concerning the task before continuing to the experiment proper.

The experiment proper was divided into three blocks, with one block corresponding to each of the three critical within-participant conditions in the study: Harmonic Music, Inharmonic Music and No Music. The order of these blocks was counterbalanced across participants. Each block contained a to-be-attended auditory 2-back task with 39 targets among 196 spoken number stimulus trials (19.89%). In all three blocks, the stream of numbers constituting the primary 2-back task was panned 90 degrees left in stereo space and thus presented only to the participants' left ear. The musical stimuli in the Harmonic Music and Inharmonic Music blocks were panned 85 degrees right in stereo space, thus perceived to be coming from the participants' right ear. The slight bias towards midline for the musical distractors was chosen because it is known to reduce the strain on a playback single channel imposed by the low frequency audio content, thereby

reducing saturation (distortion) and resulting in an increased clarity of the signal compared to a full pan to the right channel, while imposing very little influence on the perceived location of the sound source when both channels are playing (White, 2000).

Before each block, participants were told whether or not they would hear music in the upcoming block. If music was to be presented, they were instructed to attend only to the number stream while ignoring the music. In the No Music condition, participants were simply instructed to attend to the number stream. In all blocks, participants were instructed to respond to target trials by pressing the "z" key, and to respond to non-target trials by pressing the "/" key.

After the Harmonic Music and Inharmonic Music blocks, participants were prompted to complete a series of four aesthetic appraisals on the dimension of "pleasantness", "unpleasantness", "consonance", and "dissonance". Specifically, participants were asked: "On a scale from 1 – 7, how [Pleasant, Unpleasant, Consonant, Dissonant] was the music you just listened to?" Beneath the questions, participants were informed: "1 represents 'not at all' and 7 represents 'very.'" Participants responded by pressing one of the corresponding numbers on the keyboard and were also given the option of pressing "x" if they were unsure.

Results

The primary analytic focus was participants' performance on the 2-back task as a function of the Music condition (Harmonic Music, Inharmonic Music and No Music). First described are the participants' phenomenological appraisals of the harmonic and inharmonic music. Next I report analyses of the accuracy of responses to the primary 2-back task, and finally the analyses of participants' response times to the primary 2-back task.

Phenomenological Appraisals

Nine participants opted not to provide aesthetic appraisals of the musical excerpts, leaving only twenty-one participants for the analyses of the aesthetic appraisals. Mean aesthetic appraisals (i.e. "Pleasant," "Unpleasant," "Consonant," and "Dissonant") of the harmonic and inharmonic music were each submitted as the dependent variable to separate repeated measures two-tailed t-tests. The mean ratings are reported in Figure 1 with standard deviations provided in brackets.
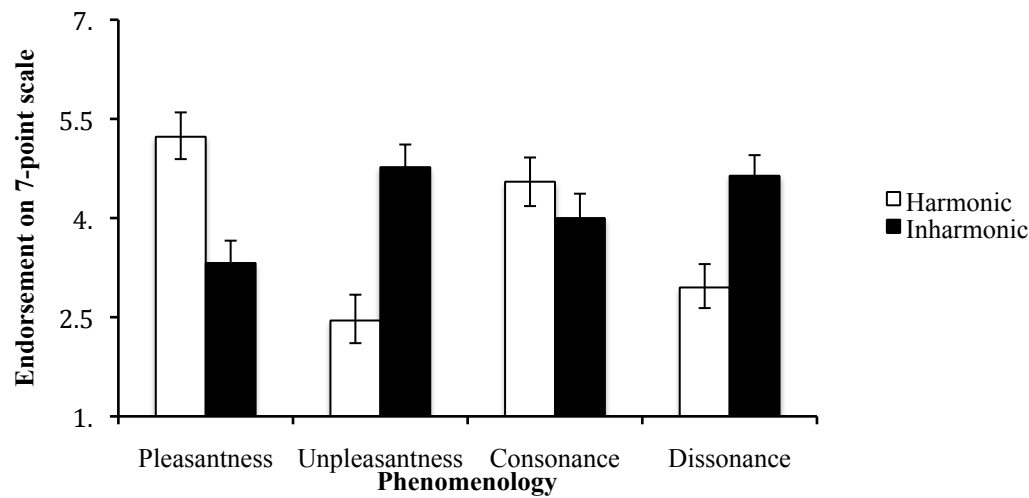
Figure 1. Mean phenomenological appraisals of the harmonic and inharmonic music in Experiment 1 (n=30). Larger numbers indicate greater experience of the rated dimension (1 = 'not at all,' 7 = 'very'). Error bars represent one standard error of the mean.

Repeated measures t-tests revealed statistically significant differences of the "Pleasant," $t(1,20) = 5.397$, $p < 0.0001$, "Unpleasant," $t(1,20) = 5.23$, $p < 0.0001$, and "Dissonant," $t(1,20) = 3.675$, $p = 0.001$ ratings of the musical pieces, with the inharmonic piece being rated less pleasant, more unpleasant, and more dissonant than the harmonic piece. There was no

statistically significant difference in the "Consonant" ratings of the two pieces, t(1,20) = 1.073, p = 0.296, though the trend was in the expected direction.

Accuracy

The means (and standard deviations) of the hit rates (proportion of targets correctly identified as targets), false alarm rates (proportion of distractors wrongly identified as targets) and sensitivity scores (a performance quotient relating participants hit rates and false alarm rates; A' as per Macmillan & Creelman, 2005) for the Harmonic Music, Inharmonic Music and No Music conditions are shown in Table 1. While the mean hits and false alarms are included in the table for completeness, analyses focused on the sensitivity scores (A'), a single performance accuracy measure combining hits and false alarms. Of primary interest was the difference in A' between the Harmonic and the Inharmonic conditions, which was assessed with a repeated measures two-tailed t-test. The analysis confirmed what can be seen in Figure 2, namely that participants performed more poorly when performing the 2-back task while listening to inharmonic music than while listening to harmonic music, t(1,29) = 2.305, p = 0.029 (mean A' difference = 0.021).

Accuracy Index	Music		
	Harmonic	Inharmonic	No Music
Hits	0.797 (0.150)	0.759 (0.170)	0.804 (0.175)
False Alarms	0.055 (0.104)	0.085 (0.081)	0.072 (0.175)

Table 1. Mean hit rates and false alarm rates (and standard deviations) for each condition in Experiment 1 (n = 30).
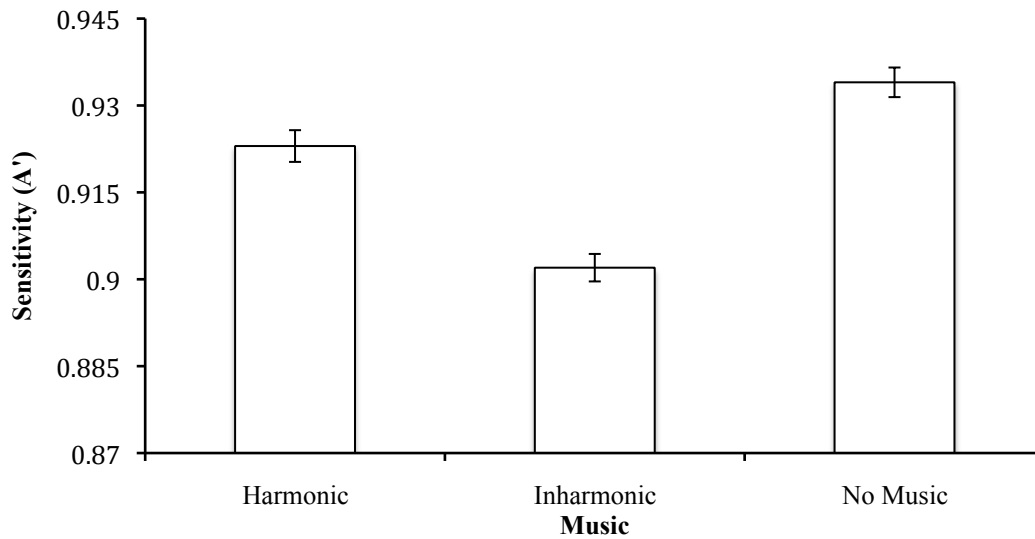
Figure 2. Mean sensitivity (A') for each condition in Experiment 1 (n=30). Error bars represent one standard error of the mean.

Two additional repeated measures t-tests compared mean A' scores in the No Music condition with those in each of the Harmonic and Inharmonic conditions. These analyses revealed that participants performed better in the No Music condition relative to the Inharmonic condition, $t(1,29) = 3.66$, $p = 0.001$ (mean difference = 0.032), but not the Harmonic condition, $t(1,29) = 1.253$, $p = 0.220$ (mean difference = 0.011).

Response Time

Mean response times (RTs) for all correct responses to Targets and Distractors of the 2-back task in the Harmonic Music, Inharmonic Music, and No Music conditions are reported in Figure 3. A test of the primary research question first compared the RTs in the Harmonic Music and Inharmonic Music conditions. The mean RTs were submitted to a 2 x 2 repeated measures factorial ANOVA with Music (Harmonic Music, Inharmonic Music) and Trial Type (Distractor, Target) serving as the within-participant factors. Critically, the analysis revealed a main effect of

Music, $F(1,29) = 25.71$, $p < 0.0001$, confirming that participants responded more slowly when inharmonic music was played than when harmonic music was played as the distracting stimulus (mean difference = 74 ms). There was also a main effect of Trial Type, $F(1,29) = 5.859$, $p = 0.022$, indicating that participants responded more slowly to Target trials than to Distractor trials (mean difference = 44 ms). There was no significant interaction between Music and Trial Type, $F(1,29) = 0.892$, $p = 0.353$.
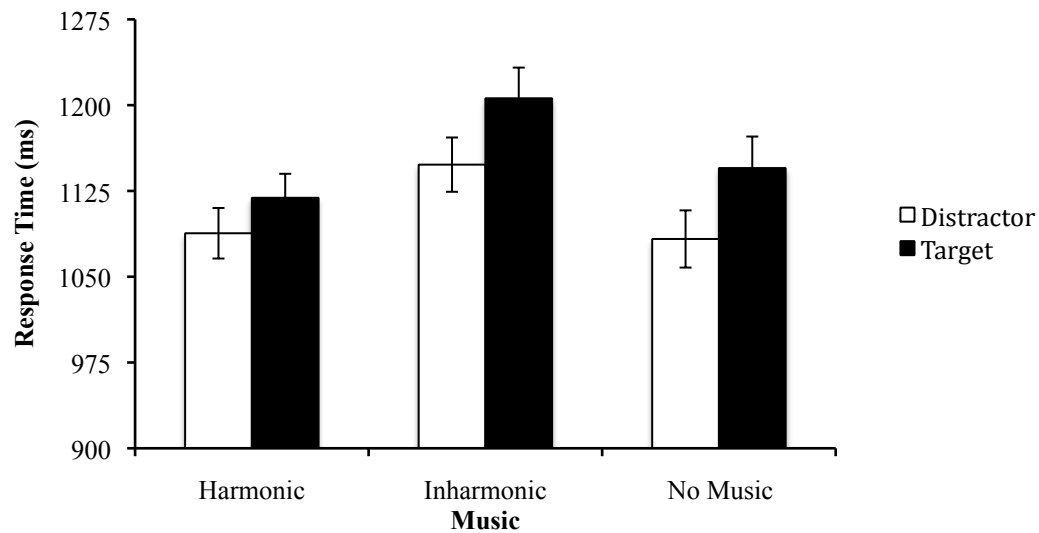
Figure 3. Mean correct response times in milliseconds for each condition and trial type in Experiment 1 (n=30). Error bars represent one standard error of the mean.

A subsequent analysis compared RTs in the Inharmonic Music and No Music conditions. The mean RTs for each of these Music conditions (Inharmonic Music, No Music) and each Trial Type (Distractor, Target) were analyzed using a 2 by 2 repeated-measures ANOVA. The analysis revealed that responses were slower in the Inharmonic Music condition than in the No Music condition, $F(1,29) = 10.533$, $p = 0.003$ (mean difference = 63 ms). In addition, responses were slower on Target trials than on Distractor trials, $F(1,29) = 7.264$, $p = 0.012$ (mean

difference = 60 ms). There was no statistically significant interaction between Music and Trial Type, F(1.29) = 0.034, p =0.855.

Finally, I compared the RTs in the Harmonic Music condition and the No Music condition as a function of Target and Distractor trials, again using a repeated-measures ANOVA. The analysis showed that the RTs for Harmonic Music and No Music conditions did not significantly differ from each other, F(1,29) = 0.296, p = 0.591 (mean difference = 11 ms). However, a main effect of Trial Type, F(1,29) = 6.653, p = 0.015 (mean difference = 47) was again observed, demonstrating slower responses to Target trials than to Distractor trials. There was no statistically significant interaction between Music and Trial Type, F(1,29) = 1.453, p = 0.238.

Summary and Discussion

Analyses of participants' phenomenological appraisals of the harmonic and inharmonic music confirmed that the inharmonic musical excerpt was indeed experienced as more dissonant than its harmonic counterpart. Both the accuracy and the reaction time data showed that performance on the 2-back task was poorer when inharmonic music was played relative to when harmonic music was played, suggesting that inharmonic music imposes greater cognitive processing demands than does harmonic music. Poorer performance on the 2-back task was also observed when participants listed to inharmonic music compared to when they listened to no music. There were no detectable differences in performance on the 2-back task when participants listened to harmonic music compared to when no music was presented, suggesting that harmonic music did not impose a measurable load on cognitive processing.

Experiment 2

Introduction

The purpose of Experiment 2 was three-fold. First, to replicate the findings from Experiment 1. Accordingly, in Experiment 2 participants were again required to respond to numbers presented in one ear (completing a 2-back task) while being presented with harmonic, inharmonic or no music in the other ear. Participants were also once again instructed to ignore the distracting music. Second, to examine whether the interfering effect of dissonant music remained even when participants were explicitly instructed to respond as quickly as possible (while maintaining high accuracy)— an instruction not provided in Experiment 1. If dissonance interferes with primary 2-back task performance under this constraint as it did in Experiment 1, the findings would suggest that the interfering effects of musical dissonance cannot be addressed with strategic control. Finally, the third goal of Experiment 2 was to replicate the participants' phenomenological appraisals of the harmonic and inharmonic music in a new sample, this time requiring all participants to provide such ratings of the music.

Method

Participants

Forty-eight undergraduate students (mean age = 19.51 years, SD = 1.82 years; 16 male) from the University of Waterloo were included in the final analysis. The students participated in a thirty-minute experiment and were compensated with partial course credit. Participants were not selected on the basis of musical training, but the number of years of music lessons ranged from 1 to 17 years (mean = 6.18= years, SD = 4.59 years).

After completing data collection for an initial sample of forty-eight participants, the data from 10 participants were excluded from the original data set for behavioural non-compliance

(responding only to target trials, prematurely terminating the experiment, and one case where two participants removed their headphones to instigate an unrelated conversation with one another as they continued the experiment). One additional participant was excluded from the original data set (n=48) because their accuracy scores fell 2.5 standard deviations below the mean (mean = 90.7%, SD = 14.4%). As a result, 11 additional participants were recruited to complete the full counterbalance and reach the predetermined sample size of forty-eight.

Apparatus and Stimuli

The apparatus and stimuli were identical to those used in Experiment 1.

Procedure

The procedure was identical to that used in Experiment 1 (section 2.2.4), except that participants were instructed in the verbal briefing and by the on-screen instructions that preceded each block to "respond as quickly and accurately as possible." In addition, all participants were required to provide aesthetic appraisals of each of the harmonic and inharmonic musical excerpts on the 1–7 Likert scales described in Experiment 1, so the option to press "x" to withhold aesthetic ratings was removed.

Results

Phenomenological Appraisals

Figure 4 presents the mean phenomenological appraisals of the harmonic and inharmonic music on each of the four dimensions (i.e. "Pleasant," "Unpleasant," "Consonant," and "Dissonant"). The mean appraisals for each dimension were submitted to a separate repeated measures two-tailed t-test. These tests revealed significant differences in ratings of the harmonic music and inharmonic music on all of the dimensions, with the inharmonic music being judged as less "pleasant," $t(1,47) = 7.816$, $p < 0.0001$, more "unpleasant," $t(1,47) = 6.239$, $p < 0.0001$,

less "consonant," $t(1,47) = 3.601$, $p = 0.001$, and more "dissonant," $t(1,20) = 5.190$, $p < 0.0001$ than the harmonic music.
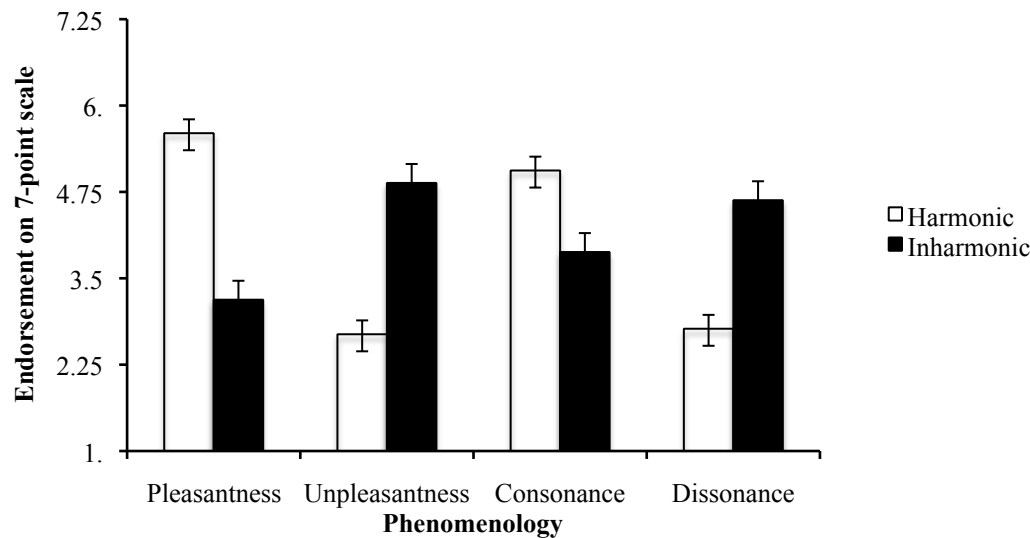

Figure 4. Mean phenomenological appraisals of the harmonic and inharmonic music in Experiment 2 (n=48). Larger numbers indicate greater experience of the rated dimension (1 = 'not at all,' 7 = 'very'). Error bars represent one standard error of the mean.

Accuracy

As in Experiment 1, accuracy analyses focused on the A' scores (shown in Figure 5) derived from participants' hit rates and false alarm rates as per Macmillan & Creelman (2005). Table 2 presents the means of the hit rates and false alarm rates in the Harmonic Music, Inharmonic Music and No Music conditions for completeness. A customary omnibus ANOVA of A' scores considering Harmonic, Inharmonic and No Music as three within-participant levels of Music confirmed a main effect of Music, $F(1,47) = 10.910$, $p < 0.0001$. In addressing the primary research hypothesis, the main interest of this ANOVA was in the difference in A' between the Harmonic Music and the Inharmonic Music conditions. Accordingly, the mean A' scores in the Harmonic Music and the Inharmonic Music conditions for each participant were submitted to a repeated measures two-tailed t-test, which revealed that participants performed

more poorly in the Inharmonic Music condition than in the Harmonic Music condition, $t(1,47) = 2.867$, $p = 0.006$ (mean difference = 0.022).

Accuracy Index	Music		
	Harmonic	Inharmonic	No Music
Hits	0.700 (0.189)	0.665 (0.188)	0.726 (0.175)
False Alarms	0.035 (0.063)	0.092 (0.145)	0.031 (0.049)

Table 2. Mean hit rates and false alarm rates (and standard deviations) for each condition in Experiment 2 (n = 48).
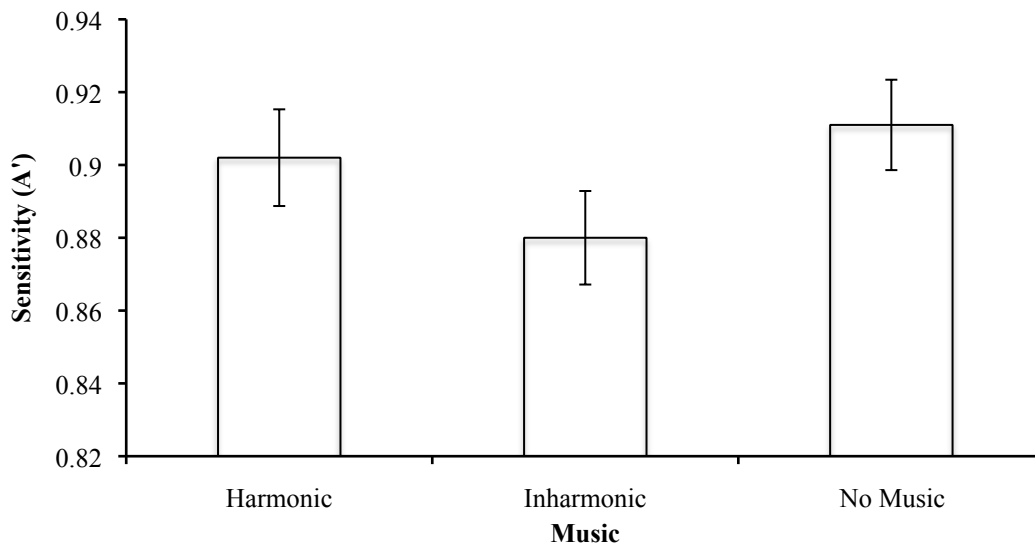

Figure 5. Mean sensitivity (A') for each condition in Experiment 2 (n=48). Error bars represent one standard error of the mean.

In addition, two repeated measures t-tests were used to compare mean A' scores in the

No Music condition with those in each of the Harmonic Music and Inharmonic Music

conditions. The analyses showed that participants performed better in the No Music condition

relative to the Inharmonic Music condition, $t(1,47) = 3.66$, $p < 0.0001$ (mean difference = 0.031),

and that there was no difference in A' scores between the No Music condition and the Harmonic

Music condition, $t(1,47) = 1.362$, $p = 0.180$ (mean difference = 0.009).

Response Time (RT)

Figure 6 shows the mean RTs for all correct responses to Targets and Distractors in the

Harmonic Music, Inharmonic Music, and No Music conditions. While the primary goal was

focus on the comparison between the Harmonic and Inharmonic Music conditions, I conducted

the customary omnibus Analysis of Variance (ANOVA) examining three within-participant

levels of Music (Harmonic, Inharmonic and No music) and two within-participant levels of Trial

Type (Distractor, Target). The ANOVA confirmed that there were main effects of Music,

$F(1,47) = 10.751$, $p < 0.0001$ and Trial Type, $F(1,47) = 35.506$, $p < 0.0001$, but no interaction

between these two factors, $F(1,47) = 0.714$, $p = 0.492$. Beginning with the planned analyses of

the RTs in the Harmonic Music and Inharmonic Music conditions, an ANOVA with the within-

participant factors of Music (Harmonic, Inharmonic) and Trial Type (Distractor, Target)

demonstrated that RTs were slower (mean difference = 40 ms) in the Inharmonic Music

condition than in the Harmonic Music condition, $F(1,47) = 7.028$, $p = 0.011$. Participants also

responded more slowly (mean difference = 86 ms) on Target trials than on Distractor trials,

$F(1,47) = 25.429$, $p < 0.0001$. The interaction between Music and Trial Type did not reach

significance, $F(1,47) = 0.920$, $p = 0.342$.
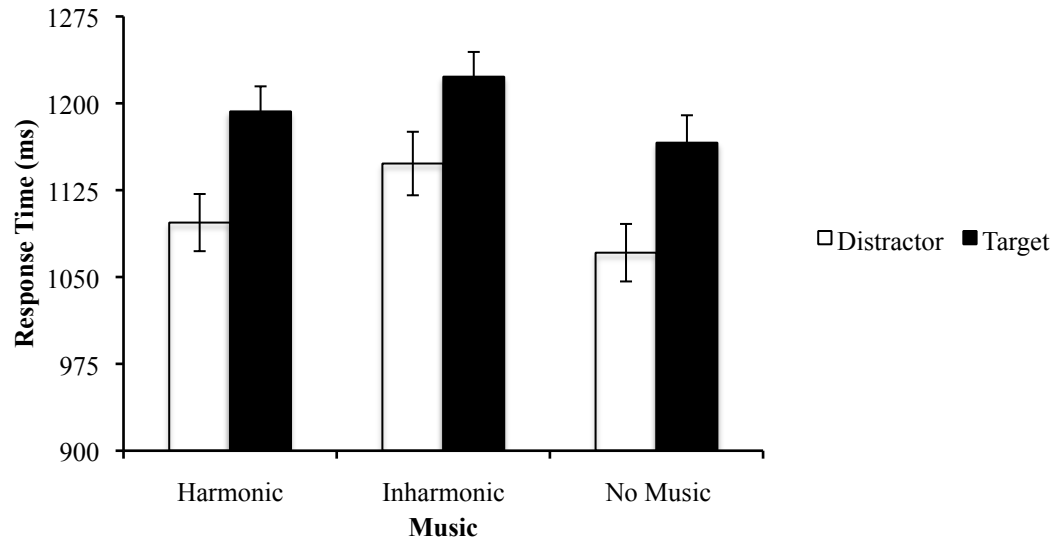
Figure 6. Mean correct response times in milliseconds for each condition and trial type in Experiment 2 (n=48). Error bars represent one standard error of the mean.

The next analyses focused on comparing the Inharmonic Music and the No Music conditions, submitting the mean RTs for each of these Music conditions (Inharmonic Music, No Music) as a within-participant factor to an ANOVA, which also included Trial Type (Distractor, Target) as a within-participant factor. RTs were slower in the Inharmonic Music condition than in the No Music condition, $F(1,47) = 19.005$, $p < 0.0001$ (mean difference = 67 ms) and slower on Target trials than on Distractor trials, $F(1,47) = 36.977$, $p < 0.0001$ (mean difference = 85 ms). There was no statistically significant interaction between Music and Trial Type, $F(1,47) = 1.208$, $p = 0.277$.

To directly compare RTs in the Harmonic Music and No Music conditions using another repeated-measures ANOVA assessing Music (Harmonic Music, No Music) and Trial Type (Distractor, Target). The main effect of Music was statistically significant, $F(1,47) = 4.256$, $p = 0.045$ (mean difference = 26 ms), with responses being slower in the Harmonic Music condition

relative to the No Music condition. The main effect of Trial Type was also significant, $F(1,47) =$ 33.147, $p < 0.0001$ (mean difference = 96 ms), with responses being slower on Target trials than on Distractor trials. The interaction between these two factors was not significant, $F(1,47) =$ 0.002, $p = 0.967$.

Summary and Discussion

Analyses of participants' phenomenological appraisals of the harmonic and inharmonic music reiterated that the inharmonic music was experienced as more dissonant than its harmonic counterpart. Both the accuracy and the response time data showed that performance on the 2-back task was poorer when dissonant (inharmonic) music was played relative to when consonant (harmonic) music was played, suggesting that dissonant music poses greater interference with cognitive processing than does consonant music. These performance effects reflect those observed in Experiment 1 despite additional instructions to bias attention towards the primary task. Together, these results suggest that the cognitive processing demands of dissonant music are to some extent automatic and evade strategic control. Poorer performance on the 2-back task was also observed when participants were presented with inharmonic music compared to when they were presented with no music. There were no detectable accuracy differences in performance on the 2-back task when participants were exposed to harmonic music compared to no music. However, responses were slightly slower in the Harmonic Music condition relative to the No Music condition despite explicit instructions to ignore the music in the Harmonic Music condition. This finding is consistent with the irrelevant sound effect literature (e.g., Tremblay & Jones, 1998) in that it might reflect a small tendency for even harmonic music to disrupt performance relative to a situation in which no music is presented, but warrants caution due its size and reliability.

Experiment 3

Introduction

The main conclusion drawn from Experiments 1 and 2 is that dissonant music not only results in negative affect as typically described, but that is also interferes with the performance of a concurrent cognitive task to a greater extent than does its consonant counterpart. In Experiments 1 and 2, however, the 2-back task and distracting musical excerpts were presented in the same sensory modality, leaving open the possibility that the measured performance decrements could be attributed to low-level sensory interference rather than cognitive processing demands. To address this possibility, Experiment 3 presented the primary 2-back task and distracting musical stimuli in different sensory modalities. Specifically, participants attended to a *visual* 2-back task while presented diotically with the harmonic or inharmonic musical distractor. This manipulation precluded any opportunity for sensory interference between the primary 2-back task and the distracting music, allowing interpretation of the measured performance interference effects, should they arise, strictly in terms of cognitive interference.

Experiment 3 also employed a modified order of the presentation of the No Music, Harmonic Music and Inharmonic Music conditions. In the previous experiments, each of these conditions was tested in a separate block of trials, fully counterbalanced between participants. A weakness of this design, however, is that variance associated with learning the 2-back task likely contaminates the responses in whichever condition is tested first, thus adding noise to the primary comparison of the Harmonic Music and Inharmonic Music conditions. To reduce this problem, participants in Experiment 3 first completed the 2-back task in the absence of music. In other words, participants were presented the No Music condition first, followed by counterbalanced blocks containing either consonant or dissonant musical distractors. This

isolated any potential decrements in performance due to learning the task to the No Music block and allowing the comparison between Harmonic Music and Inharmonic Music conditions to be uncontaminated by any such learning effects. This of course precluded conducting any meaningful statistical analyses involving the No Music condition, now confounded by order effects. As a result, no statistical analyses were used to compare performance in this condition to either the Harmonic Music or Inharmonic Music conditions. This seemed no great loss, however, as the primary comparison of interest was between the Harmonic Music and the Inharmonic Music conditions, and the spectral manipulations of the musical stimuli served as the effective experimental control in this regard. In all other ways, Experiment 3 was the same as Experiment 2.

Method

Participants

The final analysis included 48 undergraduate students (mean age = 19.01 years, SD = 1.56 years; 13 male) from the University of Waterloo. Participants were granted partial course credit after completing the thirty-minute experiment. While participants were not selected on the basis of their musical training, participants reported they received music lessons ranging from 1 to 18 years (mean = 5.00 years, SD = 3.88 years).

A sample size of forty-eight participants was predetermined for Experiment 3 before data collection began based on the results of Experiment 2. After completing data collection for an initial sample of forty-eight participants, the data from four participants were excluded from the original data set for behavioural non-compliance (three participants prematurely terminated the experiment, and one participant systematically responding "no, no, yes" for the duration of the experiment, irrespective of the targets in the to-be-attended stream). Data from two additional

participants were excluded because their response accuracy fell 2.5 standard deviations below the mean. As a result, six additional participants were recruited to complete the full counterbalance and reach the predetermined sample size of forty-eight.

Apparatus and Stimuli

The apparatus and stimuli were identical to those used in Experiment 2 except that the numbers 1–9 of the 2-back task were presented in print (80pt Helvetica font; height = 1.25cm) in the center of the computer screen in white font against a black background. Participants were seated at a normal distance from the screen but were not restricted in their head movements or viewing distance. The randomization constraints of the 2-back task were identical to those used in Experiments 1 and 2. The distracting music stimuli were identical to those in Experiments 1 and 2, with the only difference being that the music was presented diotically (i.e. with the same signal to both ears).

Procedure

Each trial of the 2-back task began with the presentation of a white fixation cross for 500-ms in the middle of a full-screen with a black background. The fixation cross was then replaced by one of the numbers of the 2-back task for 500-ms. A black background persisted for 1500-ms before the next trial began. Critically, while participants completed three blocks of trials as in Experiment 1, they always completed the No Music condition first, followed by the counterbalanced presentation of the Harmonic Music and Inharmonic Music conditions.

Results

Phenomenological Appraisals

As in the previous experiments, mean phenomenological appraisals (i.e. "Pleasant," "Unpleasant," "Consonant," and "Dissonant") for each of the harmonic and inharmonic musical

pieces were submitted as the dependent variable to separate repeated-measures two-tailed t-tests. The means of each rating are reported in Figure 7.
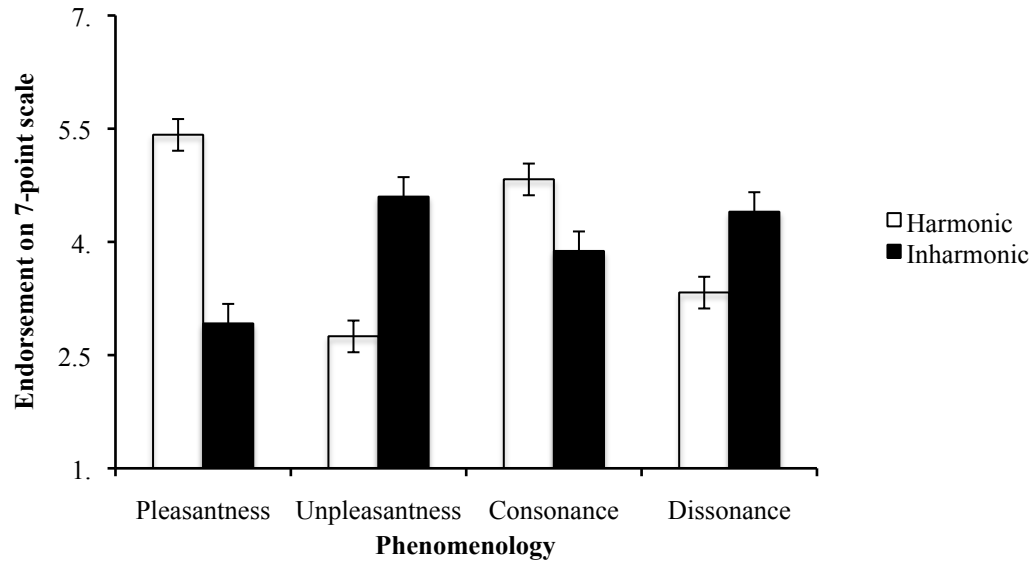
Figure 7. Mean phenomenological appraisals of the harmonic and inharmonic music in Experiment 3 (n=48). Larger numbers indicate greater experience of the rated dimension (1 = 'not at all,' 7 = 'very'). Error bars represent one standard error of the mean.

Consistent with the preceding findings, the inharmonic music was rated as less "pleasant," $t(1,47) = 10.840$, $p <0.0001$, more "unpleasant," $t(1,47) = 5.301$, $p < 0.0001$, less "consonant," $t(1,47) = 2.976$, $p = 0.005$, and more "dissonant," $t(1,20) = 2.702$, $p = 0.01$ than the harmonic music.

Accuracy

The means of the hit rates and false alarm rates from the 2-back task for the Harmonic Music, Inharmonic Music and No Music conditions are presented in Table 3. Though the descriptive statistics from the No Music condition are included for completeness, analyses focused only on comparing the A' scores between the Harmonic Music and Inharmonic Music

conditions (shown in Figure 8). Consistent with the findings in Experiments 1 and 2, analysis of

the A' scores using a repeated-measures t-test showed that performance on the 2-back task was

poorer in the Inharmonic Music condition than in the Harmonic Music condition, $t(1,47) =$

2.835, $p = 0.007$ (mean A' difference = 0.024).

Accuracy Index	Music		
	Harmonic	Inharmonic	No Music (Practice)
Hits	0.636 (0.227)	0.632 (0.213)	0.706 (0.199)
False Alarms	0.053 (0.059)	0.101 (0.075)	0.077 (0.082)

Table 3. Mean hit rates, and false alarm rates (and standard deviations) for each condition in
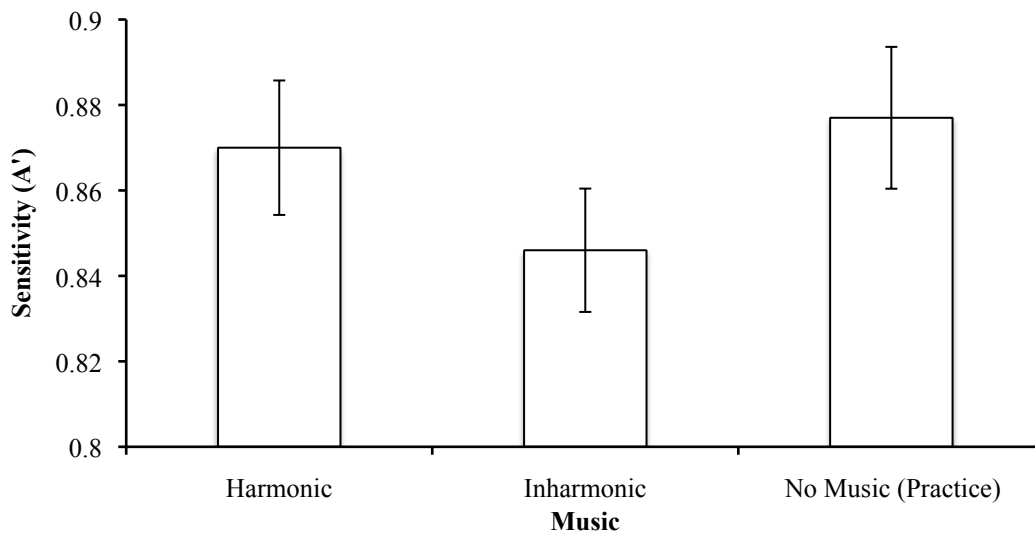
Experiment 3 (n = 48).

Figure 8. Mean sensitivity (A') for each condition in Experiment 3 (n=48). Error bars represent

one standard error of the mean.

Response Time (RT)

The mean RTs for all correct responses to the 2-back task in each condition are reported in Figure 9. Note that the RTs are much faster in this experiment than in Experiments 1 and 2. This is likely due in part because auditory stimuli must unfold over time, whereas visual stimuli are present instantaneously. Indeed, previous research has found faster RTs to visual stimuli than to auditory stimuli (e.g. Seli, Cheyne, Barton & Smilek, 2012), and this is also true specifically in the 2-back task (Owen, McMillan, Laird & Bullmore, 2005). Again, due to the fact that the No Music condition was always presented first (and not counterbalanced with the other conditions), analyses focused only on comparing the Harmonic Music and Inharmonic Music conditions, but include data from the No Music condition in the table for completeness. The mean RTs were assessed with a Music (Harmonic, Inharmonic) by Trial Type (Distractor, Target) repeated measures ANOVA. Most importantly, as in each of the previous studies, responses on the 2-back task were slower in the Inharmonic Music condition than in the Harmonic Music condition, $F(1,47) = 32.316$, $p < 0.0001$ (mean difference = 61 ms). The analysis also revealed a main effect of Trial Type, $F(1,47) = 25.429$, $p < 0.0001$ (mean difference = 80 ms), indicating that responses were slower on Target trials than on Distractor trials. Interestingly, there was also a significant interaction between Music and Trial Type, $F(1,47) = 4.647$, $p = 0.036$ indicating that the longer response times observed on Target trials relative to Distractor trials were more pronounced in the Inharmonic Music condition than in the Harmonic Music condition. As this interaction was not of primary interest, I conducted no further analyses.
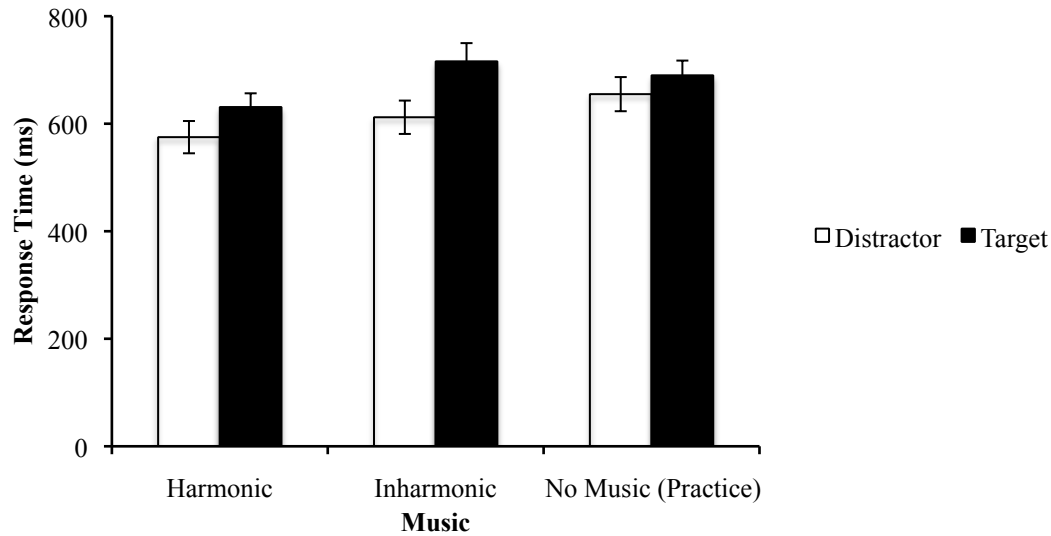
Figure 9. Mean correct response times in milliseconds for each condition and trial type in Experiment 3 (n=48). Error bars represent one standard error of the mean.

Summary and Discussion

Consistent with Experiments 1 and 2, Experiment 3 demonstrated that performance on the primary cognitively demanding 2-back task was slower and less accurate when participants were exposed to inharmonic music than when they were exposed to harmonic music. Critically, these results were observed in a cross-modal paradigm that precluded any low-level sensory interference between the music and primary cognitive task. As such they strongly suggest that the measurable task interference produced by dissonant music is a reflection of the cognitive processing load this music entails.

General Discussion

The primary goal of the present research was to evaluate the possibility that dissonant music interferes with cognitive performance to a greater extent than does consonant music. I reasoned that if these differential interference effects were potentially rooted in source dilemma, arousal, or sensory complexity, then they should occur in the absence of any response selection conflict (cf., Masataka & Perlovsky, 2013) and might best be measured under conditions of sustained cognitive processing. Consistent with these predictions, performance on the 2-back task was worse under simultaneous exposure to dissonant inharmonic music than it was under simultaneous exposure to consonant harmonic music.

Participants' phenomenological appraisals of the novel musical stimuli confirmed that the inharmonic music was experienced as "dissonant" and "unpleasant," while the harmonic music was experienced as "consonant" and "pleasant". In Experiments 1 and 2, participants completed an auditory version of the cognitively demanding 2-back task. The stimuli for this task were presented in one ear and the distracting to-be-ignored music was presented in the other ear. Experiment 3 produced the same pattern of results despite administering the primary 2-back task in the visual modality. The results of Experiment 3 provided the most convincing evidence that the cognitive interference produced by dissonant music reflects a greater cognitive processing demands of these stimuli, since the cross-modal presentation of primary and distracting tasks precluded any low-level sensory interference that might have influenced the measurements in Experiments 1 and 2.

As auxiliary points of interest, consonant music provided little to no cognitive load relative to silence (but see small effect in RT data of Experiment 2), and the interfering effects of dissonant music seem to reflect automatic cognitive processes that elude strategic control, since

they arose despite participants' repeated attempts to ignore the music while responding as quickly and accurately as possible to the primary 2-back task.

These results extend the dissonance and cognitive interference literatures in several ways. The findings are consistent with those of Masataka and Perlovsky (2013), who found that dissonant music led to slower and less accurate responses to incongruent Stroop trials than did consonant music. Unlike those of Masataka and Perlovsky (2013), however, the present results indicate the interfering effects of dissonant music reflect general cognitive processing demands and do not depend exclusively on the presence of stimulus response conflict within the primary task, as suggested by Masataka and Perlovsky (2013). Furthermore, these results contrast with those of Bodner, Gilboa and Amir (2007), which had suggested that performance on the cognitively demanding Letter Cancellation and Adjective Recall from a Story tasks was improved under exposure to dissonant musical excerpts relative to exposure to consonant excerpts or silence. The discrepancy between the present results and those of Bodner, Gilboa, and Amir (2007) might well reflect the more diligent acoustic controls implemented here during the creation of the consonant and dissonant musical excerpts.

In the Introduction I noted several reasons why one might broadly expect dissonant music to interfere with cognitive performance to a greater extent than consonant music. While I am partial to the interpretation forwarded by the SDH (Bonin, 2014), these present results are nevertheless consistent with other alternative lines of reasoning. Given the findings and methodological details of the present experiments, these possibilities can now be discussed further. Further, while at this point it is not possible to adjudicate between these accounts, it is possible to comment on their respective utilities for future research.

Preferably, a scientific account of phenomenologies, such as consonance and dissonance, would entail a mechanistic description of the psychological processes underlying the subject's experience, rather than simply a description of acoustic and psychological correlates of the stimulus that evokes the subject's experience. Accordingly, although Bodner, Gilboa and Amir's (2007) arousal account does sufficiently explain the behavioural performance in the present experiments, it lacks specificity, predicting that optimal levels of arousal will accompany optimal performance, and is devoid of any description of the cognitive mechanisms that might relate these psycho-biological correlates. Further, it does not make specific predictions about particular levels of arousal that might lead to interfering effects in a given task or situation. Thus, the arousal account is reasonably limited in both its explanatory specificity and its predictive capability. To bolster this account, future research should employ well-established physiological measures of arousal and, controlling for known confounding physiological responses, demonstrate a partial correlation between these measures, experienced phenomenology, and task performance to produce a consonance-dissonance-specific arousal index. This is to say, the arousal account is not theoretically unsound, but lacks empirically tractable methodological constraints in its current form.

Another possible explanation noted in the Introduction was a potentially higher prevalence of "salient state changes" within dissonant music compared to consonant music (c.f., Schlittmeier, Hellbrück & Klatte, 2008). While this account might explain the interfering effects of some dissonant musical stimuli, specifically those containing isolated melodic lines or salient sporadic inharmonicity, it is perhaps less useful in explaining the results reported here. This is because, as noted in the Methods, the dissonant stimuli employed here were comprised almost entirely of chorded, inharmonic intervals, resulting in a perception of dissonance that was

sustained throughout the entire composition and precluding any given dissonant moment from being experienced as particularly salient. Thus, just as the consonant stimulus implemented in these experiments exhibited sparse salient state changes by virtue of its uniform harmonicity, so too might its dissonant counterpart be considered to exhibit sparse salient state changes as a result of its uniform inharmonicity. Future perceptual research on this consideration would prove useful for both the ISE and dissonance literatures.

Finally, the present results are readily compatible with the theoretical accounts of the SDH (Bonin, 2014). Specifically, the evidence that dissonant music produces greater cognitive interference than does consonant music supports the conceptualization that dissonant music requires greater cognitive processing than consonant music. To further validate the SDH interpretation—particularly the assertion that this cognitive processing load originates from a *source dilemma*—a sensible next step for empirical inquiry would be to investigate whether the timbral and spectral manipulations of inharmonic music demonstrated by Bonin (2014) to reduce dissonance phenomenology also decrease the cognitive interference produced by inharmonic music. According to the SDH, the reason why consonant (harmonic) music would produce lesser interference with concurrent cognitive processing in the present experiments is because consonant music does not instantiate a source dilemma for the auditory system. By extension it was the presence of this source dilemma when hearing dissonant music that required greater cognitive processing, and disrupted performance on the primary tasks used in these experiments. Thus, the SDH predicts that if timbral and spatial manipulations are employed to increase the perceptual clarity of (i.e., to reduce the source dilemma resulting from) inharmonic music, the stimulus should exhibit a lesser cognitive processing load and thereby reduce the interference it presents to a concurrent cognitive task.

One advantage of this SDH account is that it provides a deeper conceptualization of how the acoustic structure of dissonant music influences performance of concurrent tasks while remaining consistent with musicological and performance psychology research that relates consonance-dissonance appraisals to the cognitive effects of enculturation and musical proficiency. Cultural (Cazden, 1980; Lundin, 1947; Vassilakis, 2005; Fritz et al., 2009) and music-theoretical (Krumhansl, 1990) norms restrict the prevalence of inharmonic intervals within most tonal music repertoires (Huron, 1994), making inharmonic moments unexpected and *bewildering* relative to their consonant counterparts (Costa, Bitti & Bonfiglioli, 2000). This bewildering quality of inharmonic music might contribute to the information complexity of dissonant intervals, producing cognitive demands on the listener that consonant, harmonic music does not. Consistent with these possibilities is a large body of literature showing that behavioural responses to complex (Patten, Kircher, Oslund & Nilsson, 2004), ambiguous (MacDonald, Just & Carpenter, 1992), conflicting (Stroop 1935; MacLeod, 1991; Eriksen & Eriksen, 1974; Sarmiento, Shore, Milliken & Sanabria, 2012), and unexpected (Jonides 1981; Posner, Snyder & Davidson, 1980; Crump, Gong & Milliken, 2006) stimuli take more time and are often less accurate than responses to simple, unambiguous, non-conflicting, and expected stimuli.

Conversely, prior experience and expertise with a particular stimulus class is known to reduce the amount of cognitive processing required by the stimuli (Wiesmann & Ishai, 2010). Accordingly, one might expect that prior experience and expertise with inharmonic music might reduce the cognitive processing demands imposed by inharmonic stimuli, reflecting a proficiency of the listener in resolving the perceptual source dilemmas produced by inharmonic music. Indeed, native listeners of non-Western tonalities readily embrace as an integral component of

41

their preferred musical aesthetic the same inharmonic intervals that Western listeners consider *dissonant* and aversive (Vassilakis, 2005). The same is observed within experts of harmonically complex Western genres, such as jazz (Dibben, 1999), employing tonalities where such inharmonic tonalities are more prevalent. An apparent next step for empirical research on the cognitive demands of musical dissonance is therefore to determine whether, regardless of objective inharmonicity, non-native musical tonalities interfere with concurrent cognitive processing to a greater extent than do native tonalities.

The SDH also leaves room for the possibility that, under specific conditions, *harmonic* musical stimuli might interfere with performance of ongoing tasks. For instance, locally harmonic chords (e.g. major chords) are experienced as dissonant when they interrupt musical passages of a separate key (Bigand, Parncutt & Lerdahl, 1996). In doing so, they violate the listener's expectations of the musical passage produced by the prevalent tonality of the auditory percept. According to the SDH, this violation is akin to that introduced by locally inharmonic chords, in that the locally harmonic chord contains frequency content that is incompatible with the global musical passage (auditory stream) and generates an unstable auditory percept (i.e., source dilemma), which in turn leads to greater processing demands. Work by McLachlan, Marco, Light & Wilson (2013) corroborates this interpretation, arguing that dissonance arises from the cognitive incongruity between perceived and expected (long-term memory) templates of common chords.

Finally, the SDH is readily compatible with the ISE. Of particular relevance is the finding by Jones, Alford, Bridges, and Macken (1999) that distracting tone sequences with timbrally segregated inharmonic components are less detrimental to concurrent cognitive performance than those with timbrally fused inharmonic components. Specifically, the authors

found that distracting tone sequences that contained intermittent oddballs deviant in either pitch *or* timbre were *more* distracting than those sequences that contained intermittent oddballs deviant in both pitch *and* timbre. Indeed, this manipulation is analogous to and provides preliminary support for the aforementioned proposal to investigate whether the timbral and spatial segregation of irrelevant inharmonic music reduces cognitive interference. The authors also found that the latter stimuli were no more distracting than a monotonous distractor devoid of any oddballs (Jones et al., 1999). Interpreting their results, the authors suggested, "Perhaps the key to understanding these contrary effects lies in an understanding of the modulating influence of auditory stream formation and its consequences for seriation (i.e., the encoding of a particular auditory stream)," concluding, in essence, that when an oddball deviates in two acoustic parameters it is most easily streamed as an independent auditory object that does not interfere with the concurrent processing of either the primary task or distracting events. Consistent with the preceding discussion, then, is the conclusion that the heightened distractibility of dissonant music relative to consonant music may be related to a perceptual dilemma instantiated by the unlikely but tenacious fusion of inharmonic frequency components within a single auditory stream.

Conclusion

While the available data do not allow one to precisely elucidate all of the links among dissonance, acoustic harmonicity, and cognitive processing demands, the present findings provide an important piece to the puzzle. The results provide novel evidence that dissonant music interferes with cognitive performance to a greater extent than consonant music on a generally demanding cognitive task that does not entail response selection conflict. This evidence corroborates predictions of the source dilemma hypothesis (Bonin, 2014) and provides empirical justification for further exploration of the model's predictions and implications. My hope is that this work provides theoretical inspiration for future empirical research on musical perception and cognitive processing.

References

Banbury, S. P., Macken, W. J., Tremblay, S., & Jones, D. M. (2001). Auditory distraction and short-term memory: Phenomena and practical implications. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 43*(1), 12–29.

Bidelman, G. M., & Heinz, M. G. (2011). Auditory-nerve responses predict pitch attributes related to musical consonance-dissonance for normal and impaired hearing. *The Journal of the Acoustical Society of America, 130*(3), 1488–1502.

Bigand, E., & Parncutt, R. (1999). Perceiving musical tension in long chord sequences. *Psychological Research, 62*(4), 237–254.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics, 58*(1), 125–141.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, *98*(20), 11818-11823.

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*(4), 382–387.

Bodner, E., Gilboa, A., & Amir, D. (2007). The unexpected side-effects of dissonance. *Psychology of Music, 35*(2), 286–305.

Bonin, T. (2014). *The source dilemma: A psychophysical account of dissonance perception in music.* Undergraduate thesis, McMaster University, Hamilton.

Bonin, T., & Smilek, D. (2015). Inharmonic music elicits more negative affect and interferes more with a concurrent cognitive task than does harmonic music. *Attention, Perception, & Psychophysics*, 1-14.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Bregman, A. S., & Pinker, S. (1978). Auditory streaming and the building of timbre. *Canadian Journal of Psychology/Revue canadienne de psychologie, 32*(1), 19.

Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *The Journal of the Acoustical Society of America, 118*(1), 471-482.

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

Cook, N. D., & Fujisawa, T. X. (2006). The psychophysics of harmony perception: Harmony is a three-tone phenomenon. *Empirical Musicology Review, 1*(2), 106–126.

Costa, M., Bitti, P. E. R., & Bonfiglioli, L. (2000). Psychological connotations of harmonic musical intervals. *Psychology of Music, 28*(1), 4–22.

Crump, M. J., Gong, Z., & Milliken, B. (2006). The context-specific proportion congruent Stroop effect: Location as a contextual cue. *Psychonomic Bulletin & Review, 13*(2), 316–321.

Dannenbring, G. L., & Bregman, A. S. (1978). Streaming vs. fusion of sinusoidal components of complex tones. *Perception & Psychophysics, 24*(4), 369-376.

DeWitt, L. A., & Crowder, R. G. (1987). Tonal fusion of consonant musical intervals: The oomph in Stumpf. *Perception & Psychophysics, 41*(1), 73-84.

Dibben, N. (1999). The perception of structural stability in atonal music: The influence of

    salience, stability, horizontal motion, pitch commonality, and dissonance. *Music*

    Perception: An Interdisciplinary Journal, 16(3), 265-294.

Ellermeier, W., & Zimmer, K. (2014). The psychoacoustics of the irrelevant sound effect.

    Acoustical Science and Technology, 35(1), 10–16.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a

    target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149.

Farndale, N. (2010, November 23). *Quincy Jones interview*. Retrieved December 17, 2014, from

    http://www.telegraph.co.uk/culture/music/rockandpopfeatures/8146333/Quincy-Jones-

    interview.html

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A.D., &

    Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current*

    Biology, 19(7), 573–576.

Hevner, K. (1935). Expression in music: A discussion of experimental studies and theories.

    Psychological Review, 42(2), 186.

Hevner, K. (1937). The affective value of pitch and tempo in music. *The American Journal of*

    Psychology, 49(4), 621-630.

Hughes, R., & Jones, D. M. (2001). The intrusiveness of sound: Laboratory findings and their

    implications for noise abatement. *Noise and Health, 4*(13), 51.

Huron, D. (1994). Interval-class content in equally tempered pitch-class sets: Common scales

    exhibit optimum tonal consonance. *Music Perception: An Interdisciplinary Journal,*

    11(3), 289-305.

Huron, D. (in press). *Voice leading: The science behind a musical art.* Cambridge, MA: MIT
Press.

Hutchinson, W., & Knopoff, L. (1978). The acoustic component of Western consonance. *Journal
of New Music Research, 7*(1), 1–29.

Jones, D., Alford, D., Bridges, A., Tremblay, S., & Macken, B. (1999). Organizational factors in
selective attention: The interplay of acoustic distinctiveness and auditory streaming in the
irrelevant sound effect. *Journal of Experimental Psychology: Learning, Memory, and
Cognition, 25*(2), 464.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement.
Attention and Performance IX, 9, 187–203.

Kameoka, A., & Kuriyagawa, M. (1969a). Consonance theory part I: Consonance of dyads. *The
Journal of the Acoustical Society of America, 45*(6), 1451–1459.

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

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch* (Vol. 17). New York, NY:
Oxford University Press.

Levenson, R. W. (1999). The intrapersonal functions of emotion. *Cognition & Emotion, 13*(5),
481-504.

Lundin, R. W. (1947). Toward a cultural theory of consonance. *The Journal of Psychology,
23*(1), 45–49.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the
processing of syntactic ambiguity. *Cognitive Psychology, 24*(1), 56–98.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide*. Mahwah, NJ: Lawrence Erlbaum Associates.

Masataka, N., & Perlovsky, L. (2013). Cognitive interference can be mitigated by consonant music and facilitated by dissonant music. *Nature Scientific Reports, 3*.

McLachlan, N., Marco, D., Light, M., & Wilson, S. (2013). Consonance and pitch. *Journal of Experimental Psychology: General, 142*(4), 1142.

Malmberg, C. F. (1918). The perception of consonance and dissonance. *Psychological Monographs, 25*(2), 93.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology, 20*(11), 1035–1041.

Moore, B. C. J. (2013). *An introduction to the psychology of hearing (6th edition)*. Bingley, UK: Emerald Group Publishing Limited.

Moore, B. C., Glasberg, B. R., & Peters, R. W. (1986). Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *The Journal of the Acoustical Society of America, 80*(2), 479–483.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping, 25*(1), 46–59.

Patten, C. J., Kircher, A., Östlund, J., & Nilsson, L. (2004). Using mobile telephones: Cognitive workload and attention resource allocation. *Accident Analysis & Prevention, 36*(3), 341–350.

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition, 68*(2), 111–141.

Plomp, R., & Levelt, W. J. (1965). Tonal consonance and critical band- width. *The Journal of the Acoustical Society of America, 38*(4), 548– 560.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*(2), 160.

Sarmiento, B. R., Shore, D. I., Milliken, B., & Sanabria, D. (2012). Audiovisual interactions depend on context of congruency. *Attention, Perception, & Psychophysics, 74*(3), 563– 574.

Schlittmeier, S. J., Hellbrück, J., & Klatte, M. (2008). Does irrelevant music cause an irrelevant sound effect for auditory items*? European Journal of Cognitive Psychology, 20*(2), 252– 271.

Seli, P., Cheyne, J. A., Barton, K. R., & Smilek, D. (2012). Consistency of sustained attention across modalities: Comparing visual and auditory versions of the SART. *Canadian Journal of Experimental Psychology, 66*(1), 44.

Siedenburg, K., Jones-Mollerup, K., & McAdams, S. (2015). Acoustic and categorical dissimilarity of musical timbre: Evidence from asymmetries between acoustic and chimeric sounds. *Frontiers in psychology, 6.*

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643.

Tenney, J. (1988). *A history of consonance and dissonance.* New York, NY: Excelsior.

Terhardt, E. (1974). On the perception of periodic sound fluctuations (roughness). *Acustica, 30,* 201–213.

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and sociobiology, 11*(4), 375-424.

Tremblay, S., & Jones, D. (1998). Roles of habituation in the irrelevant sound effect: Evidence from the effects of token set size and rate of transition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(3), 659–671.

Van Rossum, G. (2007). Python Programming Language. In *USENIX Annual Technical Conference.*

Vassilakis, P. N. (2005). Auditory roughness as means of musical expression. *Selected Reports in Ethnomusicology, 12,* 119–144.

Wiesmann, M., & Ishai, A. (2011). Expertise reduces neural cost but does not modulate repetition suppression. *Cognitive Neuroscience, 2*(1), 57–65.

White, P. (2000, October 1). Improving Your Stereo Mixing. *Sound on Sound*.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology and psychology, 18*(5), 459-482.

Yost, W. (2008). *Fundamentals of Hearing: An Introduction* (5th ed., pp. 1–338). Bingley: Emerald Publishing Group.

Zentner, M. R., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior and Development, 21*(3), 483– 492.

I. Pitch ranges for each contrapuntal voice and the pitch substitutions between harmonic

and inharmonic pieces.

Voice	Frequency range	Pitch substitutions	
		Harmonic	Inharmonic
1	F0 – A2	C1 F1	C#1 F#1
2	E1 – C3	E2 A3	Eb2 Ab3
3	F2 – G4	A3 C3 E3 C4 E4 G4	Ab3 C#3 Eb3 C#4 Eb4 Gb4
4	C3 – G5	F3 G3	F#3 G#3
5	C4 – F5	C4 E4 A5 B5 C5	C#4 Eb4 Ab5 Bb5 C#5
6	G4 – E6	No substitutions	