Cognition 154 (2016) 174–181

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

The source dilemma hypothesis: Perceptual uncertainty contributes to musical emotion

Tanor L. Bonin [a], Laurel J. Trainor [a,b,c], Michel Belyk [a], Paul W. Andrews [a,*]

[a] Department of Psychology, Neuroscience and Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada
[b] McMaster Institute for Music and the Mind, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada
[c] Rotman Research Institute, Baycrest Hospital, Toronto, Ontario, Canada

ARTICLE INFO

Article history:
Received 23 September 2015
Revised 25 May 2016
Accepted 30 May 2016
Available online 16 June 2016

Keywords:
Auditory scene analysis
Dissonance
Emotion
Evolution
Music
Perceptual uncertainty

ABSTRACT

Music can evoke powerful emotions in listeners. Here we provide the first empirical evidence that the principles of auditory scene analysis and evolutionary theories of emotion are critical to a comprehensive theory of musical emotion. We interpret these data in light of a theoretical framework termed "the source dilemma hypothesis," which predicts that uncertainty in the number, identity or location of sound objects elicits unpleasant emotions by presenting the auditory system with an incoherent percept, thereby motivating listeners to resolve the auditory ambiguity. We describe two experiments in which source location and timbre were manipulated to change uncertainty in the auditory scene. In both experiments, listeners rated tonal and atonal melodies with congruent auditory scene cues as more pleasant than melodies with incongruent auditory scene cues. These data suggest that music's emotive capacity relies in part on the perceptual uncertainty it produces regarding the auditory scene.

1. Introduction

Considerable research has focused on the induction of emotion through music (Juslin & Västfjäll, 2008). This research has engendered disagreement on the extent to which music evokes common emotions across individuals (Koelsch, 2005; Sloboda, 1996), whether musical emotions are similar to emotions evoked in other circumstances (Peretz, 2001; Trainor & Schmidt, 2003), and which specific emotions music is capable of evoking (Panksepp & Bernatzky, 2002). Despite such debates, everyday listeners cite their primary motivation for engaging with music as being the emotional responses that the music evokes (Juslin & Västfjäll, 2008). Western listeners report routinely employing music to induce desired emotions, to complement or change their current emotional state, to provide comfort, and to release stress (Behne, 1997; Gabrielsson, 2001; Juslin & Laukka, 2004; Sloboda & O'Neill, 2001; Zillmann & Gan, 1997).

Musically evoked emotions bear the physiological markers that accompany emotions evoked by other means (Trainor & Schmidt, 2003). Emotional responses to music can elicit physiological responses such as tears, shivering, lump in the throat, and chills (Sloboda, 1991; Sloboda, 2005), and changes in galvanic skin response, breathing rate, blood flow and heart rate (Krumhansl, 1997; Lundqvist, Carlsson, Hilmersson, & Juslin, 2009; Nyklíček, Thayer, & Van Doornen, 1997; Rickard, 2004). In addition, neuroimaging data indicate that music engages dopaminergic pathways and reward centers in the brain (Blood, Zatorre, Bermudez, & Evans, 1999; Koelsch, 2014; Mitterschiffthaler, Fu, Dalton, Andrew, & Williams, 2007; Salimpoor, Benovoy, Larcher, Dagher, & Zatorre, 2011). These neurophysiological responses strongly suggest that music engages the same emotion circuitry as other emotional precipitants. Behavioral studies indicate that there exists considerable agreement across listeners as to which emotion is being expressed in a particular piece of music (Gabrielsson & Lindström, 2001; Hevner, 1936; Terwogt & Van Grinsven, 1991), including to some extent listeners from different cultures (Balkwill & Thompson, 1999; Balkwill, Thompson, & Matsunaga, 2004; Fritz et al., 2009). Such evidence suggests that musical emotion relies at least partly, if not wholly, on the more general mechanisms involved in the production of emotion.

According to evolutionary theory, the biological mechanisms that produce emotion are ancient and evolved to help organisms find adaptive solutions to problems in the environment (Ekman, 1992; Levenson, 1999; Tooby & Cosmides, 1990). Cognitive appraisal processes help map environmental cues onto appropriate emotional responses (Frijda, 1993) by paying attention to those features of the environment that hold adaptive informational content, and producing emotional responses to them along with

* Corresponding author.
E-mail address: paul.andrews@psychology.mcmaster.ca (P.W. Andrews).

http://dx.doi.org/10.1016/j.cognition.2016.05.021
0010-0277/

their concomitant adaptive behaviors (Tooby & Cosmides, 1990). A key question, therefore, involves identifying the specific informational content in music that interacts with evolved appraisal mechanisms to produce emotions. Since these appraisal mechanisms are functionally linked to emotions generally, they probably did not evolve for processing music *per se*. Rather, it seems likely that music contains information that triggers appraisal mechanisms that evolved to produce emotion in other contexts.

Ideally, an explanation of how music induces emotional responses would be linked to a functional understanding of why organisms evolved to feel pleasant or unpleasant emotions. Evolutionary theory suggests that pleasant emotions arise when an organism has found an adaptive response to a situation, and the intrinsically rewarding properties of these emotions direct attention, motivation and cognition to maintaining that response. Conversely, unpleasant emotions arise when the organism lacks an adaptive response to a situation, and their function is to direct attention, motivation and cognition towards searching for and implementing adaptive responses (Carver & Scheier, 1990; Levenson, 1999; Thornhill & Thornhill, 1989; Tooby & Cosmides, 1990).

Several lines of research have identified specific aspects of music that could interact with pre-existing appraisal mechanisms to produce emotion. The *motivation-structural hypothesis* states that natural selection has shaped emotional responses to the acoustical structure of vocalizations to elicit particular behavioral responses (Morton, 1977). For example, low pitched and loud sounds are physically associated with large body size. Across species, vocal expressions in aggressive contexts (e.g., anger) include these acoustical features, presumably as a means of signaling large body size (Ohala, 1984). Similarly, it has been proposed that acoustic features of music with emotive properties may mimic those expressed by the voice in emotional contexts (Scherer, 1995). Hence, the emotions evoked by music may to some extent be predicted from acoustical features of the music (Balkwill & Thompson, 1999; Balkwill et al., 2004; Gagnon & Peretz, 2003; Hevner, 1936; Juslin, 2001).

The *expectation hypothesis* suggests that music exploits evolutionarily ancient physiological and cognitive mechanisms for detecting and responding to unexpected events (Huron, 2006; Meyer, 1956; Trainor & Zatorre, 2015). According to this idea, emotional responses to music commonly arise through manipulation of the listener's expectations. A prior musical context sets up probabilistic expectations for subsequent musical events, and the degree of confirmation or violation of these expectations, and the manner in which they are violated, gives rise to emotional responses. Huron's (2006) Imagination-Tension-Prediction-Reaction-Appraisal (ITPRA) hypothesis argues that the mechanisms involved in processing expectancy information evolved to help organisms understand, predict, and react adaptively to their environment. Trainor and Zatorre (2015) provide evidence for the physiological instantiation of expectancy processing in the brain related to musical emotion.

Finally, and of most importance here, a number of theorists have considered the role of *auditory scene analysis* (*ASA*) in musical organization and musical aesthetics (e.g., McAdams & Bregman, 1979; Wright & Bregman, 1987; Huron, 2001; Bonin & Smilek, 2015; Trainor, 2015). The auditory system evolved to accurately determine the identity and location of important objects in their environment, such as predators, conspecifics, mates, food sources, and running water, which directly affect survival and reproduction (Bregman, 1990; Fay & Popper, 2000). ASA describes the organizing principles by which the brain makes inferences about sound sources in its current auditory environment (Bregman, 1990). Each sound source emits a sound wave from some location in physical space with a characteristic spectrotemporal signature. These sound waves are summed and reach the ear as one complex wave. Thus, 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 auditory system first performs a spectrotemporal analysis of the incoming complex sound wave. It then uses a number of cues to determine which frequency components ought to be grouped together ("fused") as a single auditory object in perception because they most likely originated from the same sound source, and which should be represented as perceptually distinct ("segregated") because they most likely originated from different sound sources. ASA involves analyzing simultaneous frequency content as well as how it changes over time, as a single sound source such as a melody or spoken sentence can vary across time (Bregman, 1990).

One critical cue for the fusion or segregation of auditory objects is *temporal simultaneity* (Bregman & Pinker, 1978; Dannenbring & Bregman, 1978). Frequency components with common onset and offset times are most likely to have come from the same sound source and thus share a unified representation in auditory perception. A second cue is *parallel motion* (Bregman & Doehring, 1984; McAdams, 1982). Frequency components that move up and down in pitch together likely come from the same sound source and are consequently assigned to a unified auditory object. A third cue is *location*. Interaural time and level differences are computed in the auditory brain stem and can be used to help determine the spatial origin of different sounds (Moore, 2013). Components originating from different physical locations are most likely coming from different sound sources and are thus represented independently. A fourth cue to the number and identity of different sound sources is *timbre*. Timbre is the temporal-spectral quality of sound that allows a listener to differentiate, for example, a flute from a violin, even when those sources are playing a note of the same pitch, loudness, and duration (Caclin, McAdams, Smith, & Winsberg, 2005). Successive sounds that share a common timbre will tend to share a unified representation in auditory perception while those that differ in timbre will be represented independently in auditory perception (Culling & Darwin, 1993; Gregory, 1994).

A fifth cue is *harmonicity* (Dewitt & Crowder, 1987). Natural sounds that give rise to the sensation of pitch typically contain energy at a fundamental frequency, f_0, and at integer multiples of that frequency, called harmonics. For example, a sound with a fundamental of 100 Hz would also have energy at harmonics at 200, 300, 400, ... Hz. Thus, harmonically-related frequency components are most likely to arise from a single sound source and thus fuse in auditory perception, whereas inharmonic frequency content indicates the likely presence of more than one sound source and will tend to result in the perception of more than one sound source segregate. Interestingly, there is a long history of associating harmonic frequency relations to emotions via the concept of sensory consonance and dissonance. According to Plomp and Levelt (1965) interactions between partials close in frequency (which arise, for example, between harmonics of tones whose fundamental frequencies do not stand in simple integer ratios) cause interference patterns and beating on the basilar membrane and lead to unpleasant emotions. More recently, this idea has been modified by experiments showing that *inharmonicity* (the degree to which the frequency components present in a sound deviate from harmonics at integer multiples of a fundamental), rather than roughness and beating, is sufficient to trigger unpleasant emotion (McDermott, Lehr, & Oxenham, 2010). In any case, many studies indicate that the harmonic (or inharmonic) structure of the frequency components present in a stimulus is related to the perceived pleasantness (or unpleasantness) of a sound.

Some theorists have considered how evolved ASA functions may relate to the emotive properties of inharmonicity. Bonin and

Smilek (2015) demonstrated that sensory dissonance produces greater cognitive interference than consonant music and reasoned that unpleasant emotion might arise when conflicting perceptual inferences are derived from inharmonic music. Wright and Bregman (1987) argued that inharmonically-related frequency components would trigger less emotional unpleasantness if they could be perceptually segregated in different streams. While interesting, this account does not explain *why*, from a functional perspective, successful stream segregation should result in lower levels of negative affect. Some time ago, Huron (2001) proposed a hypothesis for the emotive properties of music that relies on the evolved system for ASA and emotion. We elaborate on this idea in this paper. Specifically, Huron (2001) proposed that Western voice leading principles are grounded in the limbic reward a listener experiences when successfully parsing multi-voice music and that the magnitude of this reward is proportional to the density of the auditory scene. Here, we elaborate on this idea that the relative success of parsing an auditory stream is related to the pleasure dimension of musical experiences—which we refer to as the *source dilemma hypothesis*—and provide the first empirical evidence consistent with the claim that clear auditory stream interpretation increases perceived musical pleasantness.

In many situations, auditory cues provide congruent information leading to a clear interpretation of the auditory scene. However, in some cases, these cues might lead to multiple, conflicting interpretations, which we refer to as a *source dilemma*. Since it is adaptive to reward a correct perceptual interpretation of the physical environment (including correctly identifying the number and location of auditory objects), the source dilemma hypothesis proposes that stimuli presenting conflicting perceptual cues to ASA will be experienced as more emotionally unpleasant than those with a more clear interpretation. The musical auditory environment can be quite artificial (Bregman, 1990; Huron, 2001) and present many non-normative combinations of psychoacoustic features. Take, for example, a dissonant chord in isolation. The temporal, timbre, and spatial characteristics of that sound may all suggest a unified perceptual representation. The (in)harmonic frequency components, however, are strongly indicative of a multiple sound sources. The system is now faced with a source dilemma. Is the veridical representation of this signal one that entails one auditory object, or many?

Successful disambiguation of a source dilemma will require an altered motivational state in which the dilemma gets prioritized access to perceptual and processing resources. However, due to the complex nature of a source dilemma, and the limited nature of processing resources, successful disambiguation of a source dilemma may be cognitively effortful (Huron, 2001), and it may force trade-offs in the processing of other stimuli. Negative emotions may facilitate these trade-offs by coordinating body systems so that processing resources are maintained on the source dilemma despite exposure to other stimuli (see also Andrews & Thomson, 2009). Importantly, pleasant emotions can also be triggered by situations that require great cognitive effort. In general, the degree of emotional aversiveness (or reward) will be proportional to the degree of effort required to search for (or maintain) an adaptive solution. Thus, complicated auditory scenes will elicit greater emotional rewards if they are successfully parsed, while they will trigger greater levels of unpleasantness to the degree they are not parsed.

Under the source dilemma hypothesis, this understanding of emotion and the auditory system is crucial to understanding at least one mechanism by which music can evoke emotional responses. Musical stimuli can artificially manipulate the listener's emotional state according to the degree of effort required to search for (or maintain) a coherent understanding of the auditory scene. The experience of unpleasant emotion will increase with stimuli that lead to perceptual uncertainty about the auditory scene, while the experience of pleasant emotion will increase with perceptual coherence. In the present paper, we use the principles of auditory scene analysis to experimentally manipulate the strength of an auditory source dilemma to determine whether or not its presence reliably influences a listener's emotional responses to music.

To test our predictions, we devised two experiments in which we measured self-reported emotional responses to simple melodies that have been manipulated to provide conflicting or congruent cues to the number of sound sources. In Experiment 1, we created a set of tonally consonant standard melodies and manipulated versions of these melodies in which: (1) some of the notes were outside of the key of the standard melody; (2) some of the notes came from a different spatial location; or (3) some of the notes were presented in a different timbre. Each of these manipulations produces uncertainty about the auditory scene, since all spectral cues in the standard melodies potentiate a congruent inference of a single sound source, thereby leading to coherent auditory perception. In contrast, the various manipulations of tonal, spatial, or timbre deviance provide a contradictory inference of multiple sound sources, thereby increasing uncertainty about the environment. We therefore predicted that manipulated melodies would be perceived as less pleasant than the standard melodies because listeners would encounter perceptual incoherence in the former. Critically, while the tonal manipulation is predicted by classical music theory to trigger an unpleasant emotional response, we expected our spatial and timbre manipulations to elicit a similar negative response, despite not being predicted by current musical theory.

In Experiment 2, we created a set of harmonically dissonant melodies, which consisted of two parallel (i.e., changing pitch up and down together), temporally synchronized melodic lines, separated by a minor ninth (13/12 octaves apart). The source dilemma hypothesis posits that the affective unpleasantness of these dissonant stimuli arises from the uncertainty in the auditory scene presented by such stimuli: while the parallel motion, temporal onset synchronization, unified timbre and unified spatial location implies one auditory object, the inharmonic frequency content produced by the dissonant minor ninth implies two auditory objects. We therefore created alternate versions of these melodies in which the two melodic lines were in either different timbres or presented from different spatial locations. While the manipulated lines still contain some conflicting cues, they both provide additional acoustic information that supports segregation into multiple sound sources. We therefore predicted that the manipulated melodies would be perceived as more pleasant than the standard melodies because they contain more congruent informational content that helps reduce uncertainty about the auditory scene.

2. Experiment 1

2.1. Method

2.1.1. Participants

Participants were 50 undergraduate students (mean age = 18.6, SD = 1.1; 15 male). Participants were not selected on the basis of musical training, but the number of years of music lessons ranged from 0 to 15 years (mean = 4.04 years; SD = 4.13 years).

2.1.2. Stimuli

Twelve 10-s standard melodies were created, one in each of the 12 Western major keys (see Fig. 1 and supplementary information). Each melody was played at a constant tempo, the tempos between melodies ranged from 72 to 106 quarter note beats per minute. The stimuli were generated in Cubase 6 with a Yamaha

Fig. 1. Schematic of the melodic stimuli used in Experiment 1. The boxes indicate the randomly selected notes that received the manipulations (constituting 20% of the notes in the melody). The three psychoacoustic manipulations are stylized above and below the staff: The upward and downward arrows represent pitch shifts of one semitone up or down, employed in the *Harmonic* condition. The Leftward and Rightward arrows represent a 90 degrees pan of that pitch in the designated direction, employed in the *Spatial* condition. Lastly, the trumpet image represents a timbre shift from piano to trumpet (or xylophone), employed in the *Timbre* condition.

KX-8 MIDI controller using the HalionSonic SE Yamaha S90ES piano sample bank. These standard melodies were rated by a pilot group of undergraduate students (n = 32) as highly pleasant (mean rating = 5.8 on a scale of 1 (highly unpleasant) to 7 (highly pleasant)). They were presented diotically (i.e., same stimulus to both ears) through headphones (dynamic closed-back Sennheiser HD280), which results in the perception of the sound as coming from the middle of the head.

Three deviant melody types were created using an "oddball" paradigm (Näätänen, Paavilainen, Rinne, & Alho, 2007; Trainor & Zatorre, 2015) by altering approximately 20% of the notes in each standard melody. In the *pitch deviant* melodies, the altered notes were raised or lowered by one semitone (1/12 octave) such that the resultant pitch did not belong in the key of the melody. In the *timbre deviant* melodies, the altered notes were presented in a different timbre (Steinberg Halion Sonic SE Trumpet). In the *spatial deviant* melodies, the altered notes were presented to either the right or left ear (i.e., 90° to the right or left of midline), randomly chosen, while the remaining notes were presented from both right and left channels so as to be perceived as coming from midline, as in the other conditions. Melodies were quasi-controlled for loudness by equating RMS amplitudes across conditions.

2.1.3. Procedure

On each trial, participants heard a standard melody and one of its corresponding deviant melodies (pitch, timbre, spatial) under computer control and were asked, in two alternative forced choice (2AFC) tasks: (1) "Which melody did you prefer" and (2) "Which melody was more unpleasant?" Each participant completed 36 trials (12 melodies × 3 deviant conditions) twice in a different random order each time, for a total of 72 trials. For each melody for each condition (key, timbre, spatial), both orders of the control and deviant melodies within a trial occurred equally often across participants.

2.1.4. Data analysis

We averaged the responses across trials for each participant. This yielded (i) a proportion of trials on which participants preferred the manipulated stimulus, and (ii) a proportion of trials on which participants selected the manipulated stimulus as more pleasant. Because the proportion data were not normally distributed, subjects were classified as to whether they, on average, preferred the manipulated or non-manipulated melody. Sign tests as implemented in R (R Development Core Team, 2014) supplemented with the BSDA package (Arnholt, 2012) were used to test whether the number of participants who exceeded an expected preference or pleasantness proportion of 0.5 was greater (or less) than expected by chance. This yields the S statistic, which is the number of participants above the expected median (i.e., with proportion scores >0.5 under the null hypothesis). Under the null hypothesis, S is equal to n/2. P-values represent the binomial probability of S statistics at least as extreme as that observed (Maxwell & Delaney, 2004).

2.2. Results

As shown in Table 1, for each of the pitch, timbre and spatial deviant conditions, participants greatly preferred the control melodies to the altered melodies. There were no significant correlations between musical training and either the preference or pleasantness measures for any of the manipulation comparisons (all Pearson's r's < 0.2, all p's > 0.20).

2.3. Discussion

Consistent with our hypothesis, participants rated the harmonic, timbre, and spatial deviant melodies as more unpleasant. The results indicate that the effect was weaker (though still statistically significant) for the spatial deviant melodies, consistent with

Table 1
Statistical results from Experiment 1. The "Preferred" column refers to the number of participants (S) out of 50 who, on average, preferred the manipulated melodies over the control melodies, and the associated p-value. The "More Pleasant" column refers to the number of participants (S) who, on average, rated the manipulated melody as more pleasant than the control melody, and the associated p-value.

Trial type	Preferred	More pleasant
Control vs. Harmonic	S = 1, $p = 2.89{\cdot}10^{-15}$	S = 1, $p = 2.89{\cdot}10^{-15}$
Control vs. Timbre	S = 1, $p = 9.08{\cdot}10^{-14}$	S = 1, $p = 9.08{\cdot}10^{-14}$
Control vs. Spatial	S = 8, $p = 3.63{\cdot}10^{-7}$	S = 6, $p = 3.24{\cdot}10^{-8}$

previous research indicating that spectral cues are more effective than localization cues for auditory segregation (Woods et al., 2001). Participants preferred melodies that did not contain occasional unexpected changes in key, timbre or spatial location. Such oddball paradigms, in which elements of a sequence are infrequently and unexpectedly changed in some dimension, are often used to examine memory traces and the automatic encoding of expectations in auditory cortex (Näätänen et al., 2007; Trainor & Zatorre, 2015). What is different about the present experiment is that, rather than measuring expectation per se, the dependent measures were ratings of preference and pleasantness. There are a number of reasons why participants might have preferred the control melodies. First, according to Huron's ITPRA theory of emotion (Huron, 2006), people have evolved so that correct predictions of the future engage the reward system and failures are (at least initially) experienced as unpleasant. Thus, the changes in key, timbre and location might have been experienced as unpleasant because they were not correctly predicted. Second, people might simply prefer simpler stimuli, without changes in key, timbre and location, whether expected or not.

A third possibility, which we are interested in exploring here, is related to the fact that the key, timbre and location manipulations introduce uncertainty about the source(s) of the input. As discussed in the Introduction, there are a number of cues that the auditory system uses to determine the number of sounds sources in an environment, and whether specific sounds emanate from the same or different sources. With the present stimuli, the melodic context cues a single source, but each of the pitch, timbre and spatial manipulations suggests that the altered notes are from a different source than the unaltered notes. Trying to integrate these conflicting cues leads to a *source dilemma.* Our hypothesis is that this source dilemma contributes to the deviant melodies being experienced as less pleasant and less preferred compared to the control melodies, for which the cues to sound source are consistent and for which there is therefore no source dilemma. We note that, in fact, this account is not at odds with those that rely on stimulus complexity or expectations, but that it does specify where this processing conflict is taking place—at the level of perceptual integration. That is, signal complexity might correlate with unpleasantness when it produces incoherent perception, and violated expectations might correlate with unpleasantness when they represent perceptual uncertainty. The specificity of our claim allows us to make falsifiable predictions about the contribution of the source dilemma to musical emotions in the presence and absence of these previously described phenomena. Stronger evidence of course would involve demonstrating the impact of the source dilemma on the listener's emotional responses by controlling for stimulus complexity and expectations. We do so in the following experiment.

Given that expectations and complexity likely play a role in emotional responses, in Experiment 2 we test whether a source dilemma alone affects ratings of pleasantness and preference. In particular, we eliminate expectation violation altogether and set up a situation where, opposite to that of Experiment 1, the introduction of different timbres or spatial locations cues *increases* stimulus complexity but *decreases* source dilemma. We do this by creating a standard stimulus that is perceived to be unpleasant, specifically, two simultaneous parallel melodies at the dissonant interval of a minor 9th (13 semitones apart or 13/12th of an octave) (see Fig. 2). This stimulus creates interference patterns and beating on the basilar membrane (Plomp & Levelt, 1965), and does not conform to a harmonic template (McDermott et al., 2010). In terms of auditory scene analysis, this stimulus has conflicting cues to the source(s) of the sounds. The common melodic movement from note to note, the common spatial location, and simultaneous sound onsets are all cues to one sound source, but the inharmonic tonal structure cues two sounds sources. Thus, according to the source dilemma hypothesis (and Western tonal theory, and sensory dissonance models), this stimulus will be rated as rather unpleasant. In the case of our altered stimulus, however, where either the timbres or the spatial locations of the two simultaneous melodies are differentiated, the brain receives additional information in favor of a two-source interpretation. These additional cues complement the interpretation derived from the stimulus inharmonicity, and we predict that this increased perceptual coherence should lessen the experience of instability and the resultant perceptual dilemma. Thus, these manipulations lead to a reduction in source ambiguity, and the source dilemma hypothesis predicts that there will be an increase in rated pleasantness and preference. Critically, the melodies of the altered stimulus are just as atonal, inharmonic, and even more complex compared to those of the standard stimulus, and are thus predicted by all current models of musical emotions to elicit the same *or a lesser* degree of positive affect compared to the standard. The source dilemma hypothesis predicts, to the contrary, that these manipulations should increase the listener's positive affect.

3. Experiment 2

3.1. Method

3.1.1. Participants

Participants were 52 undergraduate students (mean age = 18.8, SD = 0.9; 40 female. 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 = 5.5 years, SD = 5.2). Data from three participants were improperly recorded to the experiment log file and were not included in the data analyses.

3.1.2. Stimuli

Twelve 10-s stimuli, comprised of simultaneous parallel melodies at the interval of a minor 9th (13 semitones apart or 13/12th of an octave) were created in each of the 12 Western major keys. They were generated in Cubase 6 with a Yamaha KX-8 MIDI controller using the HalionSonic SE Yamaha S90ES piano sample bank. Each melody was played at a constant tempo (range = 72–106 quarter note beats per minute). These standard melodies were rated by a pilot group of undergraduate students (n = 32) as highly unpleasant (mean rating = 1.4 on a scale of 1 [highly unpleasant] to 7 [highly pleasant]). They were presented diotically (i.e., same stimulus to both ears) through headphones (dynamic closed-back Sennheiser HD280), which results in the perception of the sound as coming from midline.

Three altered melody types were created for each of the standard melodies. Altered melodies were generated by manipulating an entire voice (upper *or* lower octave) of each of the standard melodies in one of three ways: For the *timbre* manipulation, one of the voices was performed in a different timbre (six timbre manipulations used a trumpet while the other six used a xylophone) by playing the recorded MIDI data through the Halion Sonic SE virtual instrument sample banks. For the *spatial* manipulation, the upper and lower voices were presented in different stereo spatial locations, that is, one from the right and one from the left channel of the headphones. Finally, for the *timbre & spatial* manipulation, one of the voices was presented in a different timbre (Steinberg Halion Sonic SE Xylophone or Steinberg Halion Sonic SE Trumpet) from either the right or left channel while the other voice was presented in a piano timbre from the opposite channel. Melodies were quasi-controlled for loudness by equating RMS amplitudes across conditions.

3.1.3. Procedure

On each trial, participants heard a standard melody and one of its corresponding altered melodies (timbre, spatial, or timbre & spatial) under computer control and were asked, in two alternative

Fig. 2. Schematic of the parallel melodic stimuli used in Experiment 2. Here, the gray notes are used to visually distinguish one melodic voice from the other. Acoustically, this was accomplished either by placing one register in a piano timbre with the other in that of a trumpet or xylophone, or by panning one register 90 degrees to the participants' left and the other 90 degrees to the participants' right, or by employing both of these manipulations simultaneously (see box schematic to left of score).

forced choice (2AFC) tasks: (1) "Which melody did you prefer?" and (2) "Which melody was more unpleasant?" Each participant completed the 36 trials (12 melodies × 3 deviant conditions) twice in a different random order each time, for a total of 72 trials. For each melody for each condition (timbre, spatial, timbre & spatial), both orders of the control and deviant melodies within a trial occurred equally often across participants.

3.2. Results

There were no significant correlations between musical training and either the preference or pleasantness measures for any of the manipulation comparisons (all Pearson's r's < 0.2, all p's > 0.20).

As shown in Table 2, participants rated Control stimuli as significantly less preferred and less pleasant compared to those in which the two simultaneous melodies were in different timbres, different spatial locations, or both. Surprisingly, the timbre & spatial manipulation was significantly less preferred ($S = 16$, $p < 0.05$) and rated as significantly more unpleasant ($S = 11$, $p < 0.01$) than the timbre manipulation. There was a similar trend for the spatial manipulation, but it did not reach statistical significance.

3.3. Discussion

We created standard stimuli (two parallel melodies at a dissonant interval) with conflicting cues as to the number of sound sources present that were perceived to be somewhat unpleasant. Specifically, the parallel, temporally synchronized melodic lines suggested a single sound source, while their distinct (in)harmonic profiles suggested two sound sources (Bregman, 1990). According to the source dilemma hypothesis, this particular combination of auditory cues is unpleasant because it leads to uncertainty about the identity and location of auditory objects in the environment, and the timbre and spatial manipulations of one of the parallel melodic lines reduce unpleasantness because they facilitate segregation of those lines and help disambiguate the auditory scene. Consistent with our hypothesis, participants rated the timbre- and spatially-altered melodies as more pleasant. The effect was weaker (though still statistically significant) for the spatial deviant melodies, consistent with previous research indicating that spectral cues are more effective than localization cues for auditory segregation (Woods, Alain, Diaz, Rhodes & Ogawa, 2001). While the combination of timbre and spatial manipulations also reduced ratings of unpleasantness, this effect was smaller than for either the timbre or spatial manipulation alone, contrary to our expectations. It is possible that this is related to the fact that spatial cues are much less efficient than timbre or spectral cues for stream segregation leading to complex interactions between these cues. This question remains for future research.

It should be reiterated that the results cannot be explained in terms of aversion to complexity. Whereas the addition of auditory information (timbre, spatial differences, harmonic differences) to a melody was associated with *increased* unpleasantness in Experiment 1, the addition of such content was associated with *decreased* unpleasantness in this experiment. The results cannot be attributed to differences in temporal expectations (such as Huron's ITPRA theory) since the succession of sound events did not differ across conditions. Therefore, we propose that, in addition to complexity and expectation, a source dilemma can make a powerful contribution to emotional responses.

Table 2
Statistical results for Experiment 2. The "Preferred" column refers to the number of participants (S) out of 52 who, on average, preferred the manipulated melody over the control melody, and the associated p-value. The "More Pleasant" column refers to the number of participants (S) who, on average, rated the manipulated melody as more pleasant than the control melody, and the associated p-value.

Trial type	Preferred	More pleasant
Control vs. Timbre	$S = 43$, $p = 1.37 \cdot 10^{-8}$	$S = 43$, $p = 1.37 \cdot 10^{-8}$
Control vs. Spatial	$S = 40$, $p = 3.31 \cdot 10^{-6}$	$S = 40$, $p = 9.26 \cdot 10^{-6}$
Control vs. Timbre & Spatial	$S = 35$, $p = 6.60 \cdot 10^{-3}$	$S = 36$, $p = 2.60 \cdot 10^{-3}$

4. General discussion

The source dilemma hypothesis outlines a mechanism for triggering emotional responses to music that depends on the listener's degree of certainty or uncertainty about the auditory scene. Building on the ideas of Huron (2001), auditory stimuli that increase perceptual clarity of the auditory scene should trigger pleasant emotions, while stimuli that decrease perceptual clarity should trigger unpleasant emotions. We reported two experiments in which musical stimuli were manipulated to either decrease or increase listeners' perceptual clarity as regards the number of sound sources in the auditory environment, manipulations which were predicted to correspondingly decrease or increase ratings of the pleasantness of the experimental stimuli.

In Experiment 1, we introduced cues that decreased perceptual clarity into otherwise well-structured melodies. Using an oddball paradigm, a proportion of the notes in melodies were altered in pitch, timbre, or spatial location to introduce conflicting cues as regards to number of sound sources in the auditory environment. All of the manipulations reduced ratings of pleasantness and preference, which is consistent with the source dilemma hypothesis. However, there are alternative explanations for these results. For example, Huron's ITPRA theory would posit that the violations of expectation caused by the occasional altered notes would lead to emotional responses. And the increase in stimulus complexity that occurred by introducing altered notes might in itself lead to emotional responses. Indeed, previous research indicates that listeners prefer musical excerpts with a moderate degree of subjective complexity over those that are overly simplistic or overly complicated (North & Hargreaves, 1995). As we noted in our discussion of the results of Experiment 1, the source dilemma hypothesis is not incompatible with these models of expectation or stimulus complexity. Indeed, it might be the case that the "unexpected" events that listener's find most aversive tend to be those that represent an incongruent percept (as opposed to those which violate learned, cognitive appraisals) and this might be worth testing in future empirical studies. Furthermore, it could be that stimulus complexity correlates with negative affect when (and only when) it produces incoherent perception, but is not sufficient to explain the listener's emotional response. We reasoned that, regardless of these potential relations between our theory and others, the most effective demonstration of our hypothesis was to test it in a context where expectations and complexity could not explain the results.

Thus, in Experiment 2, we manipulated an auditory source dilemma in the absence of violations of expectation or increasing stimulus complexity. Specifically, we presented two temporally synchronous, parallel, but tonally inharmonic, melodies that all perspectives predict to be aversive. The source dilemma hypothesis posits that such melodies are unpleasant because the synchronous onsets, parallel motion of the melodic lines, unified timbres and unified spatial locations of the two melodies all suggest a single sound source, while the inharmonic intervals separating them suggests a two-source interpretation. The source dilemma hypothesis predicts that manipulations that facilitate the segregation of these simultaneous melodies into different auditory streams will reduce unpleasantness. Consistent with this, we found that manipulating the melodies so that they were in different timbres or came from different spatial locations, or both, reduced ratings of unpleasantness.

Because there were no violations of expectations within each stimulus based on the previous context, ITPRA cannot explain these results. Similarly, the results cannot be explained by stimulus complexity, as the increased complexity resulting from the addition of different timbres or spatial locations actually decreased stimulus unpleasantness. It also needs to be considered as to whether sensory dissonance and harmonic relations could explain the results. Because of the minor ninth pitch interval between the tones of the two parallel melodies, in the case where both melodies are presented to the same ear, there are many cases where a harmonic from a tone in one melody and a harmonic from a tone in the other melody together cause a substantial interference pattern on the basilar membrane. According to the sensory dissonance hypothesis of Plomp and Levelt (1965), these interference patterns lead to sensations of beating and roughness, resulting in feelings of unpleasantness. In the spatial separation condition, these interference patterns are eliminated because the pairs of interfering harmonics are presented to different ears. Thus, in this case, our observed reduction in unpleasantness could be the result of a reduction in sensory dissonance. In the case of our timbre manipulation, however, this explanation is insufficient, as the harmonics from both melodies are still presented to the same ear. Thus, sensory dissonance arising from interference patterns on the basilar membrane cannot explain all of the results of Experiment 2. Indeed, it is possible that sensory dissonance is simply one aspect of the source dilemma hypothesis as harmonic relations are a prominent cue to sound source separation.

One important limitation of the present research is that we did not explicitly assess participants' certainty about the number of distinct sound sources they were hearing. Although we designed our stimuli based on principles that are well supported by research on auditory scene analysis, the source dilemma hypothesis crucially predicts that the unpleasantness of acoustic stimuli will positively correlate with the level of perceptual coherence they produce. Future research should explicitly test this prediction. Future research could also examine whether people spontaneously attempt to alleviate uncertainty in the auditory scene through various means, such as visual corroboration and using head movements to better determine source location.

In sum, we have demonstrated that harmonic music can be made more unpleasant by the addition of stimulus features that are predicted by ASA to cloud the perception of the auditory scene (Experiment 1). Additionally, we have demonstrated that the unpleasantness of inharmonic music can be reduced by the addition of stimuli that are predicted by ASA to disambiguate the auditory scene (Experiment 2). These results cannot be entirely explained by ITPRA or stimulus complexity or theories of sensory consonance, but they are consistent with the source dilemma hypothesis. As discussed above, the source dilemma hypothesis also predicts that source dilemmas are cognitively effortful and can cause cognitive trade-offs. Consistent with this, Bonin and Smilek (2015) have demonstrated that inharmonic music causes greater interference on a cognitive task than harmonic music.

Altogether, these results suggest that the source dilemma account offers an important principle in understanding musically-evoked pleasure — presumably one of multiple possible hedonic factors. Such a framework, if validated, could help inform a systematic and multidimensional psychoacoustic approach to musical composition, orchestration, and performance, that extends beyond the current scope of Western tonal theory and has been sought after by composers and researchers alike for years (cf., McAdams & Bregman, 1979; Huron, 2001, Huron, 2006).

5. Conclusion

Based on principles of auditory scene analysis and the evolution of emotion, the source dilemma hypothesis predicts that uncertainty in the number, identity or location of sounding objects elicits unpleasant emotions. In two experiments we have demonstrated that (i) introducing conflicting auditory cues makes consonant melodies sound less pleasant, and (ii) introducing cues that promote a single interpretation of an ambiguous auditory scene makes the dissonant melodies sound more pleasant. These data suggest that the emotional responses of listeners may, in part, be driven by relative certainty/uncertainty in interpreting the auditory scene.

Acknowledgments

Alan Rheaume, Juan Lopez Tiboni, Danielle Mogyoros, Marley Russell, and Victoria Walaszczyk are gratefully acknowledged for assistance in various aspects of data collection and preparation. We also thank the anonymous reviewers for comments that greatly improved the manuscript. This research was funded by a grant from the McMaster Science and Engineering Research Board to PWA and an NSERC Discovery Grant to LJT (RGPIN-2014-0470).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2016.05.021.

References

Andrews, P. W., & Thomson, J. A. Jr, (2009). The bright side of being blue: Depression as an adaptation for analyzing complex problems. *Psychological Review, 116*(3), 620–654.

Arnholt, A. T. (2012). BSDA: Basic Statistics and Data Analysishttp://cran.r-project.org/package=BSDA.

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

Balkwill, L.-L., Thompson, W. F., & Matsunaga, R. I. E. (2004). Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners1. *Japanese Psychological Research, 46*(4), 337–349.

Behne, K.-E. (1997). The development of "Musikerleben" in adolescence: How and why young people listen to music. In I. D. J. Sloboda (Ed.), *Perception and cognition of music* (pp. 143–159). Hove, England: Psychology Press/Erlbaum (UK) Taylor & Francis.

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.

Bonin, T. L., & Smilek, D. (2015). Beyond the emotional impact of dissonance: Inharmonic music interferes with cognitive performance. *Attention, Perception, & Psychophysics*, 1–14.

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

Bregman, A. S., & Doehring, P. (1984). Fusion of simultaneous tonal glides: The role of parallelness and simple frequency relations. *Perception & Psychophysics, 36*(3), 251–256.

Bregman, A. S., & Pinker, S. (1978). Auditory streaming and the building of timbre. *Canadian Journal of Psychology/Revue canadienne de psychologie, 32*(1), 19–31.

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.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review, 97*, 19–35.

Culling, J., & Darwin, C. J. (1993). The role of timbre in the segregation of simultaneous voices with intersecting F0 contours. *Perception & Psychophysics, 54*(3), 303–309.

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

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

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*(3–4), 169–200.

Fay, R. R., & Popper, A. N. (2000). Evolution of hearing in vertebrates: The inner ears and processing. *Hearing Research, 149*(1–2), 1–10.

Frijda, N. H. (1993). The place of appraisal in emotion. *Cognition and Emotion, 7*(3–4), 357–387.

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., et al. (2009). Universal recognition of three basic emotions in music. *Current Biology, 19*(7), 573–576.

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In P. N. J. J. A. Sloboda (Ed.), *Music and emotion: Theory and research* (pp. 223–248). New York, NY, US: Oxford University Press.

Gabrielsson, A. (2001). Emotions in strong experiences with music. In P. N. J. J. A. Sloboda (Ed.), *Music and emotion: Theory and research* (pp. 431–449). New York, NY, US: Oxford University Press.

Gagnon, L., & Peretz, I. (2003). Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies. *Cognition and Emotion, 17*(1), 25–40.

Gregory, A. H. (1994). Timbre and Auditory Streaming. *Music Perception: An Interdisciplinary Journal, 12*(2), 161–174.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *The American Journal of Psychology, 48*(2), 246–268.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception: An Interdisciplinary Journal, 19*(1), 1–64.

Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: The MIT Press.

Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research, 33*(3), 217–238.

Juslin, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework. In P. N. J. J. A. Sloboda (Ed.), *Music and emotion: Theory and research* (pp. 309–337). New York, NY, US: Oxford University Press.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*(5), 559–575.

Koelsch, S. (2005). Investigating emotion with music. *Annals of the New York Academy of Sciences, 1060*(1), 412–418.

Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience, 15*(3), 170–180.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 51*(4), 336–353.

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

Lundqvist, L.-O., Carlsson, F., Hilmersson, P., & Juslin, P. N. (2009). Emotional responses to music: Experience, expression, and physiology. *Psychology of Music, 37*(1), 61–90.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data* (2nd ed.). New York: Taylor & Francis Group.

McAdams, S., & Bregman, A. S. (1979). Hearing musical streams. *Computer Music Journal, 3*(4), 26–60.

McAdams, S. (1982). Spectral fusion and the creation of auditory images. In M. Clynes (Ed.), *Music, mind, and brain: The neuropsychology of music* (pp. 279–298). New York: Plenum Press.

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

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.

Mitterschiffthaler, M. T., Fu, C. H. Y., Dalton, J. A., Andrew, C. M., & Williams, S. C. R. (2007). A functional MRI study of happy and sad affective states induced by classical music. *Human Brain Mapping, 28*(11), 1150–1162.

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

Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist, 111*(981), 855–869.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*(12), 2544–2590.

North, A. C., & Hargreaves, D. J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology, 14*, 77–93.

Nyklíček, I., Thayer, J. F., & Van Doornen, L. J. (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology, 11*, 304–321.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica, 41*(1), 1–16.

Panksepp, J., & Bernatzky, G. (2002). Emotional sounds and the brain: The neuro-affective foundations of musical appreciation. *Behavioural Processes, 60*(2), 133–155.

Peretz, I. (2001). Listen to the brain: A biological perspective on musical emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 105–134). New York, NY, US: Oxford University Press.

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

R Development Core Team (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rickard, N. S. (2004). Intense emotional responses to music: A test of the physiological arousal hypothesis. *Psychology of Music, 32*(4), 371–388.

Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience, 14*(2), 257–262.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9*(3), 235–248.

Sloboda, J. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music, 19*(2), 110–120.

Sloboda, J. (2005). *Exploring the musical mind: Cognition, emotion, ability, function*. Oxford: Oxford University Press.

Sloboda, J. A., & O'Neill, S. A. (2001). Emotions in everyday listening to music. In P. N. J. J. A. Sloboda (Ed.), *Music and emotion: Theory and research* (pp. 415–429). New York, NY, US: Oxford University Press.

Sloboda, J. A. (1996). Emotional responses to music: A review. In K. Riederer & T. Lahti (Eds.), *Proceedings of the nordic acoustical meeting (NAM96)* (pp. 385–392). Helsinki: The Acoustical Society of Finland.

Terwogt, M. M., & Van Grinsven, F. (1991). Musical expression of moodstates. *Psychology of Music, 19*(2), 99–109.

Thornhill, R., & Thornhill, N. W. (1989). The evolution of psychological pain. In R. Bell & N. Bell (Eds.) (pp. 73–103). Lubbock, TX: Texas Tech University.

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

Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 370*(1664).

Trainor, L. J., & Schmidt, L. A. (2003). Processing emotions induced by music. In I. Peretz & R. J. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 310–324). Oxford: Oxford University Press.

Trainor, L. J., & Zatorre, R. J. (2015). The neurobiology of musical expectations from perception to emotion. In Susan Hallam, Ian Cross, & Michael Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 285–306). Oxford Handbooks.

Woods, D. L., Alain, C., Diaz, R., Rhodes, D., & Ogawa, K. H. (2001). Location and frequency cues in auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 27*(1), 65–74.

Wright, J. K., & Bregman, A. S. (1987). Auditory stream segregation and the control of dissonance in polyphonic music. *Contemporary Music Review, 2*(1), 63–92.

Zillmann, D., & Gan, S.-L. (1997). Musical taste in adolescence. In D. J. H. A. C. North (Ed.), *The social psychology of music* (pp. 161–187). New York, NY, US: Oxford University Press.