Do you chill when I chill? A cross-cultural study of strong emotional responses to music.

Article *in* Psychology of Aesthetics Creativity and the Arts · April 2020

DOI: 10.1037/aca0000310

Eleonora Beier
University of California, Davis

Justin Conor Hulbert
Bard College

Do you chill when I chill? A cross-cultural study of strong emotional responses to music.

Eleonora J. Beier[a] (ejbeier@ucdavis.edu)*, Petr Janata[a] (pjanata@ucdavis.edu),
Justin C. Hulbert[b] (jhulbert@bard.edu), & Fernanda Ferreira[a] (fferreira@ucdavis.edu)

[a] Department of Psychology, University of California Davis, 1 Shields Ave, Davis, CA 95616, USA.
[b] Psychology Program, Bard College, 30 Campus Road, Annandale-on-Hudson, NY 12504, USA.

* Corresponding author.

Abstract

While listeners can infer the mood expressed by the music of a different culture, the question of whether strong felt emotional responses can also be induced cross-culturally remains unanswered. We address this question by measuring chill responses, sudden increases in emotional arousal, through self-report and skin conductance measures. Excerpts of Western classical, traditional Chinese, and Hindustani classical music were presented to three groups of participants, each familiar with one of these styles. Participants felt a similar number of chills to both familiar and unfamiliar musical styles, but significantly fewer chills to scrambled music, which acted as a control. Acoustic analyses revealed that sudden peaks in loudness, brightness and roughness were correlated with chills across styles, suggesting that similar acoustic elements induced emotional responses cross-culturally. The number of chills was also related to the degree to which participants paid attention to the music, rated the music as emotional and as having affected their own mood, but not to how much participants liked the music. Overall, this research counters the idea of musical emotional meaning as being entirely generated within cultural conventions and supports the claim that some aspects of the way music conveys emotion may be shared across cultures.

Keywords: chills; music; cross-cultural; emotion; skin conductance.

1. Introduction

Music has been an important aspect of human experience throughout cultures and time. Every cultural group we are aware of, no matter how connected or isolated, engages in some form of musical activity, and musical instruments possessed by our ancestors date back at least as far as 40,000 years ago (Nettl, 2000; Fitch, 2006; Conrad et al., 2009; Higham et al., 2012). Yet despite its ubiquity, music is extremely diverse in its form and sound, the context it is played in, its role in society, and its significance to those who play it, dance to it, or just listen to it. Even defining what "music" is has been a challenge, as different societies conceptualize what we refer to as music in very different ways (Nettl, 2000; Blacking, 1995).

Despite these difficulties, music has received growing interest in psychology. In particular, the cross-cultural study of music can reveal the cognitive processes that underlie music cognition at a fundamental level, helping us unravel the still somewhat mysterious origin and function of music. Linguistics and psycholinguistics routinely rely on cross-linguistic comparisons to extract the linguistic features underlying variation and the psychological mechanisms that make up human language processing; similarly, studying music cross-culturally would contribute greatly to our understanding of the cognitive underpinnings of music. However, cross-cultural comparisons in music cognition have been relatively limited, while the majority of studies and theories focus on Western music.

Cross-cultural comparisons of music are more common in the field of ethnomusicology, though with very different methods and goals. Historically, early cross-cultural comparisons tended to decontextualize non-Western music, often feeding into the idea that Western art music was superior (e.g., Bloom, 1987; for discussion see Becker, 1986; Myers, 1992). Because of this, the field of ethnomusicology turned away from comparative approaches and stressed instead the importance of understanding a musical style within its cultural context (Myers, 1992). While comparative approaches have since returned and the issue of how much of musical meaning is culturally constructed is still debated (Blacking, 1995, Clayton, 2003; Myers, 1992), the underlying idea that music can only be understood within its cultural context is generally accepted in this field (see Clayton, 2003 for an anthology of writings on this topic). In other words, musical meaning derives primarily from social and cultural conventions; without knowledge of these conventions, a listener would not be able to comprehend music from another culture.

In psychology, several studies have demonstrated the effect of enculturation – extensive experience with a particular musical style in its cultural context, especially at an early age – such that novel music from one's own culture is remembered and recognized better than music from another culture, suggesting that cultural knowledge fundamentally influences musical memory and processing (Morrison & Demorest, 2009). For instance, listeners from remote populations unfamiliar with Western music do not always show a preference for consonance over dissonance (McDermott et al., 2016), which is often thought of as a central aspect of musical perception. Thus, music of different cultures may rely on fundamentally different properties and modes of listening.

Despite these apparent cultural differences across musical styles, researchers have struggled for decades to reduce music from different areas of the world to a set of common characteristics, or musical universals (e.g., Fitch, 2006; Harwood, 1976; Lomax, 1968; Mehr et al., 2019; Savage et al., 2015). Human universals more generally have been controversial, as they typically have not been acknowledged by the predominant anthropological view that human behavior results primarily from cultural norms in absence of a universal basis (Brown, 1991; 2004). Nonetheless, more recently a range of absolute and near universals have been proposed, spanning from the use of tools to semantic categories (Brown, 1991; 2004) to some aspects of linguistic structure (Adger, 2019; but see Goldberg, 2019). With respect to music specifically, while no single property has been found in all musical styles surveyed, many characteristics appear to be shared by a large number of cultures (Savage et al., 2015). Thus, the search for musical universals may be better framed as a search for statistical patterns in the properties of music from around the world. Some well-documented cross-cultural aspects of music include structural properties such as the prevalence of isochronous rhythms divided into two or three subdivisions, the use of discrete pitches organized into scales, and the importance of the fifth-degree interval and octave equivalence—which may be traced to psychoacoustic properties of sound (Patel, 2008; Savage et al., 2015). The presence of these patterns suggests that while cultural conventions indisputably contribute to musical meaning, some aspects may rely on cognitive commonalities in how music is processed and generated. By studying the extent to which listeners can comprehend music from other cultures, and the shared musical elements they rely on, we can begin to describe the mutual psychological basis for music cognition.

1.1. Emotional meaning across cultures

Just as it is difficult to define music, it is not clear what constitutes musical meaning. Some aspects of musical meaning may be considered “intrinsic” to the music itself - what has been defined as “intra-musical” meaning, as opposed to “extra-musical” meaning derived from personal or societal associations (Patel, 2008; Blacking, 1995). These associations are what may constitute culture-specific meaning often described by ethnomusicology, such as a certain type of song signifying sadness because it is typically played at funerals, or the responses engendered by emotional contagion through collective musical experiences and rituals (e.g., Turner, 2012). Additionally, music may carry language-specific meaning when some of the emotional weight is carried through lyrics. On the other hand, intra-musical aspects of meaning may reflect broadly human features of communication (e.g., mimic prosodic features), and some may even be found across species (e.g., a shrieking sound sounds scary because it is likely to signal threats in the environment; Cross et al., 2009).

One type of musical meaning that has received much attention is emotional meaning, or the listener's ability to express and infer emotions from music. Parallel to the search for universals, then, is the question of whether people can infer the intended emotional meaning of music from a culture they are not familiar with. This would suggest that listeners from diverse backgrounds rely on a similar set of features to encode and interpret emotions in music. In support of this view, it was found that Western listeners can correctly identify the intended mood expressed by Indian ragas - an improvised melodic piece from the classical Hindustani tradition (Balkwill & Thompson, 1999). Perceived emotions were related to particular psychoacoustic dimensions: faster tempos were positively correlated with perceived joy and negatively correlated with sadness and peacefulness; rhythmic and melodic complexity were positively correlated with perceived sadness and negatively correlated with joy and peacefulness. The authors concluded that, when unable to rely on extra-musical cultural conventions, listeners make their judgments based on psychophysical aspects of the music, which can be therefore be seen as shared cues of emotion.

These findings were replicated and expanded in a later study (Laukka et al., 2013). Professional performers from four different musical traditions (Swedish folk music, Hindustani classical music, Japanese traditional music, and Western classical music) recorded short pieces from their tradition in order to convey eleven different emotions: happiness, sadness, anger, fear, affection, peacefulness, humor, longing, solemnity, spirituality, and neutral. Participants from Sweden, India, and Japan subsequently listened to those recordings and chose from the list of emotions the term that best described what they heard. All participants were familiar with the Western music, which was used as a control, while they were not familiar with music from the cultures other than their own. It was found that listeners discriminated the expressed emotion with accuracy above chance. Cross-cultural recognition rates varied for different emotions, and in some cases there was higher agreement within cultures than across cultures. Similarly to Balkwill and Thompson (1999), this study found relations between the cues used by performers to encode an emotion and by listeners to decode it across traditions. For example, higher volume was associated with anger and lower volume with affection and peacefulness. Similarly, a clear pulse and a rough timbre corresponded to anger, happiness and humor, whereas an ambiguous pulse and a smoother timbre corresponded to longing, peacefulness and sadness. At the same time, there was a better match between the cues used by performers and those used by listeners within one culture than across cultures.

It could be argued, however, that these results may be partly due to the fact that all participants involved in these studies were familiar with Western music, which could account for the shared use of emotional cues. Evidence to the contrary comes from a study conducted among the Mafa people of Cameroon, a culturally isolated population unfamiliar with Western music (Fritz et al., 2009). Mafa participants recognized happiness, sadness and fear in Western musical excerpts with accuracy above chance, though to a lesser degree than Western participants. The authors speculated that the Mafa may have relied on aspects of Western music that resembled emotional prosody. Together, these studies suggest that emotional communication in music relies both on culture-specific and on cross-cultural cues of emotion. This is in agreement with the observation that musical meaning may be partly based on communicative strategies common to all humans, in addition to culturally specific associations (Cross et al., 2009).

While this line of research points to a universality in the way certain psychoacoustic cues such as tempo are interpreted, some cross-cultural work has also been developed based on the idea that music expresses emotion through an intrinsic syntax-like structure. Tonal structure can be described through hierarchical trees of tension and relaxation (Lerdahl & Jackendoff, 1983). The incompleteness of a pattern or a deviation from expectations leads to tension, while the fulfillment of expectations leads to relaxation; their interplay gives rise to changes in arousal and emotion (Meyer, 1967; Juslin & Västfjäll, 2008; Steinbeis, Koelsch & Sloboda, 2006). While this theory was primarily based on Western music, hierarchical tonal structure is a feature of music from other cultures as well (e.g., Hindustani classical music; Rohrmeier & Widdess, 2012), though tonal systems (the set of rules governing how tones relate to each other) typically differ. A probe tone task, in which participants rated how well a tone fit the preceding melody, was used to show that listeners form similar expectations in response to different styles of music regardless of their familiarity with that style. This was found for Hindustani classical music (Castellano, Bharucha & Krumhansl, 1984), Chinese folk music (Krumhansl, 1995), and North Sami yoiks, a musical style typical of the Sami people indigenous to the Scandinavian Peninsula (Krumhansl et al., 2000). These results were interpreted in light of the idea that melodic expectancies may partly rely on Gestalt principles of grouping (Narmour, 1990), suggesting that the melodic structure of music from different cultures may be based on similar patterns. Additionally, listeners show rapid incidental learning of the pitch-distributional information and tonal hierarchies of unfamiliar tonal systems (Oram & Cuddy, 1995; Rohrmeier & Widdess, 2012). Thus, structural expectations giving rise to emotion may be driven both by statistical learning and by Gestalt principles (Morgan et al., 2019).

The literature thus provides strong support for the idea that emotional meaning in music may rely, at least in part, on intra-musical aspects independent of cultural associations. Some of these aspects include psychoacoustic properties (e.g., the association between rhythmic complexity and valence) as well as tonal and melodic expectations (through quick learning of novel tonal systems and shared Gestalt principles). This is in contrast to the view that musical meaning is entirely defined within a culture, through extra-musical associations or culture-specific tonal systems and structures. Nonetheless, while these studies found a striking ability for listeners to infer the emotional meaning and the structure of music from another culture, members of that culture often showed better performance on discrimination tasks. Thus, these results do not deny the effect of musical enculturation and the importance of cultural associations.

1.2. Felt emotion

One question that the literature arguably does not give a satisfying answer to, however, is whether people actually *feel* any emotions in response to music from other cultures. It has been argued that recognizing the emotion expressed by a piece of music is not equivalent to a listener experiencing that emotion (Juslin & Laukka, 2004); in fact, there are significant differences in the ratings of valence and arousal for a listener's own feelings in response to music, as opposed to what the listener thought the music tried to convey (Salimpoor et al., 2009), although there can be a strong relationship between the two (Evans & Schubert, 2008). The studies discussed thus far either measured people's recognition of the mood expressed by music or their structural expectations, which are thought to give rise to felt emotion, but did not address actual felt emotion per se. It could be argued that for emotional meaning to be truly conveyed across cultures, a listener must feel an emotional response and not only recognize what a piece may convey to people familiar with it.

This issue was addressed in a second experiment among the Mafa (Fritz et al., 2009). Western and Mafa participants listened to excerpts of Western and Mafa music that were either in their original form or had been spectrally manipulated to induce permanent sensory dissonance (the melody was played together with the same melody shifted one semitone upward and one tritone downward). They then rated the excerpts for pleasantness as a measure of felt emotion. Both Mafa and Westerners rated the original excerpts as more pleasant than their dissonant versions, which was interpreted as an indication of felt emotion across cultures. However, it is hard to draw this conclusion based on this manipulation, as what accounted for the results is the introduction of dissonance, which has been associated with unpleasantness starting in early age (Trainor & Heinmiller, 1998). Thus, it is not clear whether the observed pleasantness ratings were a reflection of felt emotion to the music or just a preference for consonance over dissonance across cultures, which may not specifically be tied to music (and might not be present in other remote populations; McDermott et al., 2016).

A second issue relates to the difficulty of inferring felt emotion from self-report, as self-report measures lack objectivity and can be subject to demand bias (Juslin & Laukka, 2004). Thus, simply asking participants to rate their own felt emotions (e.g., their sadness in response to the music) may reflect their ability to infer the intended emotion in music rather than their actual felt emotion. On the other hand, there are no strong associations between discrete emotion categories (e.g., “happiness” or “sadness”) and physiological measures (Mauss et al., 2004;

Mauss et al., 2005), reflecting the idea that discrete emotions may not arise through separate mechanisms but rather through the categorization of an underlying core affect (Barrett, 2006). This is likely what has prevented researchers from using physiological correlates to assess people's emotional responses to music of other cultures on top of their recognition of mood. Finally, the emotion categories typically used to describe emotions in Western music may not necessarily apply to music of other cultures (Laukka et al., 2013). While ratings of valence and arousal could be used instead, this would not successfully distinguish between emotions that fall similarly along these two continuums (e.g., anger and fear, both high in arousal and with negative valence) but that are differently communicated in music (Juslin & Laukka, 2004). Overall, this points to the difficulty of measuring felt emotion in music.

One way to address this issue is to focus on peak emotional responses such as shivers, goosebumps, and tears induced by music (Sloboda, 1991). In particular, some attention has been given to chills, also known as "thrills" or "frissons" (Harrison & Loui, 2014). Chills may be defined as pleasurable aesthetic experiences corresponding to a peak in the intensity of felt emotion to a particular point in music (Grewe & Kopiez, 2009). They correspond to sudden increases in emotional physiological arousal, often felt as a shiver down the spine or in the arms, and which can be measured physiologically through skin conductance (Craig, 2005; Grewe et al., 2007; Guhn et al., 2007; Salimpoor, 2009). Moreover, neural studies of chills have revealed their association with a network of reward and emotion processing regions (Salimpoor et al., 2011; Salimpoor et al., 2013) as well as areas associated with reward and motivation for biologically relevant stimuli such as food and sex (Blood & Zatorre, 2001). Interestingly, these areas did not respond to pleasantness induced by consonance, further suggesting that chills may be a better indication of felt emotion than pleasantness ratings (Blood & Zatorre, 2001).

The literature suggests that chills are associated with surprise, often coincident with unexpected events such as harmonic changes (Sloboda, 1991), sudden dynamic leaps (e.g., peak in loudness; Grewe et al., 2007; Guhn et al., 2007, Panksepp, 1995), the entry of a new voice and the re-introduction of a previous theme (Grewe et al., 2007). This is consistent with the view that changes in emotional arousal in music derive from patterns of tension and relaxation based on whether expectations are fulfilled (Lerdahl & Jackendoff, 1983). The specific ways in which expectations and other aspects of the music may give rise to chills are further explored in the Discussion.

What makes chills particularly appropriate for tackling the question of cross-cultural emotional communication is that, by measuring felt emotion more directly, they do not risk being confounded by emotion recognition (such as in the recognition of expressed mood) and by cross-cultural variability in the description of emotions. Moreover, being correlated with changes in physiological arousal, self-report can be corroborated by an objective indication of emotion such as a peak in skin conductance, leading to stronger conclusions. Finally, we can contribute to the search for commonalities in the way music expresses emotion around the world by describing the psychological and acoustic conditions giving rise to chills in music from different cultures and in participants from different backgrounds. It is therefore surprising that, as far as we are aware, the present study represents the first attempt to investigate chills in response to non-Western music (Harrison & Loui, 2014). To the best of our knowledge, no psychological research has documented chills in music and populations outside of Western culture, though some research suggests that strong emotional responses such as chills may be relatable to music-induced spiritual trance experiences common around the world (Penman & Becker, 2009). Other societies may conceptualize chill-like emotional responses in culturally specific ways (e.g., the *maysie* of Scottish ballads; McFadyen, 2014).

One of the first psychological investigations into chills in Western music speculated that it is unlikely people would feel chills to music of other cultures given the differences in musical structure (Sloboda, 1991). However, the literature reviewed above suggests that many aspects of this structure may be shared, and that recognition of emotions cross-culturally is possible. We therefore predict that, if familiarity and experience with a musical style from a different culture are not necessary to elicit an emotional response, people will experience chills in response to music of other cultures.

1.3. The present study

This experiment serves the double purpose of providing evidence for felt emotional responses in music from other cultures and of broadening the scope of the literature on chills by including non-Western stimuli. In a design inspired by the previous research on cross-cultural emotional communication in music, participants from different backgrounds were exposed to music from the Western classical, traditional Chinese, and Hindustani classical traditions. These styles were chosen given their prominence in the cross-cultural literature here surveyed; this experiment therefore tested whether, on top of inferring the mood and forming tonal and melodic expectations in response to music from these cultures, people unfamiliar with it do indeed experience felt emotional responses in the form of chills.

Based on the evidence for the existence of emotional communication across cultures, we expected that participants would experience a similar number of chills regardless of their experience with each musical style, but that they would experience more chills to music in its original form than to the same music scrambled in the time domain. Scrambled music has been used as a non-musical control with similar acoustic features to music but lacking a cohesive structure. Music engages different patterns of brain activity than does scrambled music (Abrams et al., 2013; Fedorenko et al., 2012; Janata et al., 2002; Jiang et al., 2013; Levitin & Menon, 2005; Menon & Levitin, 2005), specifically areas involved with processing linguistic structure (e.g., Brodmann Area 47; Levitin & Menon, 2005) as well as areas involved in reward processing (e.g., nucleus accumbens and ventral tegmental area; Menon & Levitin, 2005). Behaviorally, scrambled music is rated as less coherent and less aesthetically pleasing (Lalitte & Bigand, 2006). Thus, we expected the scrambled music to induce fewer chills than the intact music.

This study also aims to expand our understanding of what musical elements give rise to these strong emotions across cultures. Overall, we hypothesized that chills would be more likely to occur in response to unexpected acoustic events in the music, such as peaks in loudness, roughness and brightness (Grewe et al., 2007). We also hypothesized that if our participants felt chills in music of other cultures, they may have relied on at least some of these cues across styles. Finally, some of the psychological states linked to the experience of chills were attested cross-culturally. In particular, based on previous work, we hypothesized that number of chills would be related to self-reported liking of the music (Colver & El-Alayli, 2016; Panksepp, 1995; Salimpoor et al., 2009) and self-reported attention paid to the music (Nusbaum et al., 2014). We also expected it to be related to the degree to which participants rated the music as emotional and as having affected their own mood.

2. Method

2.1. Participants

Sixty-seven participants (females = 45; median age = 19.5 years) were recruited in two ways: (1) through a University of California, Davis undergraduate participant pool system in which participants are compensated with class credit, and (2) through fliers in which participants were told they would be entered in a raffle for a $50 Amazon gift card. Data from five participants were removed from the analyses due to the participants' failure to comply with the experimental instructions or due to the data not being recorded properly. Participants were screened for hearing impairments and categorized into three groups. Participants in the Western group (n = 21; females = 16; median age = 19.5 years) reported enjoying listening to Western classical music and were unfamiliar with Hindustani classical and traditional Chinese music. Participants in the Chinese group (n = 21; females = 15; median age = 20 years) reported enjoying listening to traditional Chinese music and were unfamiliar with Hindustani classical music; similarly, participants in the Indian group (n = 20; females = 14; median age = 19 years) enjoyed listening to Hindustani classical music but were unfamiliar with traditional Chinese music. Ethnicity and country of origin were not considered for group categorization. Familiarity and enjoyment of Western classical music was not a screening factor for the Chinese and Indian groups, as their familiarity with Western music in general could be assumed by virtue of living in a Western country. The validity of participants' self-reported familiarity with each musical style was confirmed through a questionnaire that assessed their knowledge of each style (see Method and Results for more details). The fact that all participants were familiar with Western music allowed us to use Western music as the baseline against which to compare participants' responses to the other two styles, which varied in familiarity. The possibility that this shared knowledge may have contributed to participants' responses is addressed in the Discussion. Lastly, since only between 56% (Grewe & Kopiez, 2009) to 75% (Craig, 2005) of people have been found to experience chills to previously unheard music, and about a fourth of the population does not experience chills at all (Nusbaum et al., 2014), all participants were screened for having previously had the experience of chills, based on self-report.

2.2. Stimuli

2.2.1. Musical Pieces. Four two-minute musical excerpts from each style (Western classical, traditional Chinese and Hindustani classical) were selected through a norming study. Previous studies indicate that unfamiliar excerpts between 2 and 3 minutes in length can induce chills in participants (Guhn et al., 2007; Salimpoor, 2009). Based on the literature and on correspondence with musical experts[1] (Mariagrazia Carlone, personal communication, November 4th, 2015; Paul Beier, personal communication, December 10th, 2015; Yun Fan, personal communication, November 17th, 2015; Heera Kulkarni, personal communication, November 23rd, 2016), six full length pieces were originally chosen from each style according to the following criteria: 1) highly emotional, 2) instrumental (no voice/lyrics), and 3) relatively unknown. Very well-known pieces were excluded to prevent familiarity with a particular piece

[1] Mariagrazia Carlone, PhD is a musicologist and archivist currently working at the Italian State Archive in Milan; her publications cover musical iconography, historical and biographical research. Paul Beier is a musician specialized in early plucked instruments of the Renaissance and Baroque periods; he has taught at the Civica Scuola di Musica Claudio Abbado in Milan, Italy since 1981. Heera Kulkarni is a retired educator, musician, a book author and Yoga instructor from New Delhi, India; she completed her Sangeet Visharad degree in Hindustani music from Gandharva Mahavidyalay, Delhi, and is the founder/director of Raga Academy of Indian Music where she has been teaching Hindustani music for the past twenty years. Yun Fan received Master's degrees in ethnomusicology from Xiamen University and Wesleyan University, as well as a Master of Library and Information Science from the University of Illinois, Urbana Champaign; she currently works as an editor for Chinese language publications in all fields of music and related subjects at RILM (International Repertory of Music Literature, New York, NY).

from confounding the effects of familiarity with the genre as a whole, as familiar pieces with personal associations have been found to elicit more chills than unknown music (Craig, 2005; Salimpoor, 2009). Two two-minute excerpts were then extracted from each piece; these sections were deemed to coincide with the most emotional events in the piece by the same musical experts, while also exhibiting some of the acoustic features typically associated with chills (e.g., a peak in loudness; Grewe et al., 2007, Guhn et al., 2007; Panksepp, 1995). A norming study presented participants with all 12 musical excerpts only for the style they had reported enjoying. Ten participants per group were selected using the same criteria as described above; none of them took part in the main experiment. For each style, the four excerpts that had elicited the most chills in the most participants, that had received low familiarity scores, and that did not come from the same piece were selected. Information about the resulting stimuli can be found in Table 1 (see Table S1 in the Supplementary Materials for more information about the musical instruments used in the traditional Chinese and Hindustani classical pieces). All stimuli were amplitude normalized using Audacity software (Audacity Team 1999–2019), so that the peak amplitude of each piece was set to -1 dB below the maximum amplitude possible without distortion of the audio signal.

Table 1
Pieces selected through the norming study for inclusion in the main experiment.[2]

[2] As the excerpts were selected based on the results of the norming study, we were unable to control for factors such as the instruments used or their era. The original six pieces selected with the help of the musical experts included a broader variety of sub-genres (e.g., Western classical music from earlier periods), but only the excerpts that induced the greatest number of chills were chosen for the experiment. Additionally, stimuli across musical styles differed in their musical features (e.g., whether they relied mostly on harmony or on single line embellishments), as well as whether they were composed, adapted from traditional songs, or improvised. This is due to the inherent differences in the way these musical styles emerge from their cultures, which we believe is an unavoidable consequence of studying music cross-culturally. While we do not believe that these differences alter the interpretation of our results, future studies may be able to circumvent this issue through the use of musical stimuli purposely created for the experiment.

Piece	Composer/Origin	Instrument
Western Classical		
String Quartet A minor op.132, 3rd movement	Beethoven	String Quartet
String Quartet A minor op.13, 4th movement	Mendelssohn	String Quartet
1st Piano Concerto, 2nd mvt	Chopin	Piano, orchestra
Kol Nidrei	Bruch	Orchestra
Traditional Chinese		
Riding to the flower festival	Traditional	Pipa
Chilly wind in spring	Hua Yanjun/Abing	Erhu, yangqin
Autumn Moon over the Han Palace	Traditional	Pipa
Royal chimes music	Traditional	Bianzhong
Hindustani Classical		
Raga Jog	Improvised	Sarangi
Raga Shivaranjani	Improvised	Shehnai
Kajri	Improvised	Shehnai
Raga Kalawati	Improvised	Santoor

2.2.2. Scrambling procedure. The literature offers no standardized scrambling procedure, with segments to be scrambled varying in length from 250 ms (Levitin & Menon, 2005) to 900 ms (Jiang et al., 2013) and up to 15 s (Lalitte & Bigand, 2006). In this study, the musical excerpts were processed using a time-domain scrambling of audio signals script (Ellis, 2010) in MATLAB (Mathworks, Natick, MA). The script was edited in order to randomly reorder segments of 250 ms over the length of the excerpt. The scrambled excerpts used in this experiment resembled their non-scrambled counterparts in their overall length.

2.2.3. Relaxation sound. Before the start of the experiment, participants listened to three minutes of a recording of waves crashing on a beach to relax them and to allow the skin conductance measurement to stabilize (Figner & Murphy, 2011).

2.2.4. Practice excerpt. The first stimulus presented to participants was a two-minute excerpt of relaxing instrumental jazz music. The purpose of the practice excerpt was to allow participants to become accustomed to the setup and to the post-piece survey. Jazz was chosen as a neutral genre different from the three styles used in the experiment. Although it shared some of the tonal properties of the Western classical pieces and the improvisational quality of the Hindustani classical pieces, we considered this style to be sufficiently different from the others, yet from a genre that should sound familiar to most participants.

2.3. Behavioral measures

Similarly to Grewe et al. (2007), chills were measured behaviorally by asking participants to press a button with their right hand (in this case, the spacebar on a keyboard) at the onset of the experienced chill, and to keep it pressed for the entire duration of the chill. The mere pressing of a button in this type of design has been shown not to lead to changes in skin conductance (Grewe et al., 2007).

Following the presentation of each excerpt, a post-piece questionnaire adapted and expanded from Panksepp (1995) assessed participants' experience of the music using 5-point Likert scales. Participants rated their liking of the music (Liking), how much they had been able to focus on the music (Attention), how emotional they perceived the music to be (Emotionality) and how much it had affected their own mood and feelings (Affect Mood). Additionally, participants rated how familiar they were with the pieces they heard and the perceived amount of happiness or sadness in the music. The purpose of the familiarity question was to confirm that none of the pieces was overly familiar to a group of participants, though all pieces had already been screened using the same measure in the norming study. The results of the happiness and sadness questions were not analyzed at this time as they did not pertain to the hypotheses of this study.

At the end of the experiment, a survey assessed the participants' level of knowledge of each musical style (see Appendix). The survey questions were determined based on correspondence with experts of each musical style (Mariagrazia Carlone, personal communication, November 4th, 2015; Yun Fan, personal communication, November 10th, 2015; Heera Kulkarni, personal communication, April 6th, 2017). Questions tested knowledge of the instruments used in each style, important composers or works, and general nomenclature. Additionally, part of Gooding et al.'s (2014) questionnaire on musical experience was adapted to test for knowledge of Western notation and music theory. While it is not clear whether this knowledge influences the prevalence of chills (Grewe et al., 2009) and musical expectancy more generally (Krumhansl, 1995), we included this measure to more clearly assess the effects of musical knowledge on felt emotional responses to each musical style.

2.4. Skin conductance

Skin conductance (SC) was monitored as an objective measure of physiological arousal, since skin conductance responses (SCRs), or peaks in skin conductance, have been linked to the chill response (Grewe et al., 2007; Grewe & Kopiez, 2009; Guhn et al., 2007). Before the start of the experiment, participants were asked to wash their hands with water, after which two passive Nihon Kohden electrodes were placed on the volar (palmer) surface of the distal phalanges of the index and middle finger of the left hand with electrode paste, as suggested by Figner and Murphy (2011). Participants were instructed to rest their left hand face down on an armrest (with the tip of their fingers off the border so that pressure was not applied to the electrode sites) and to keep as still as possible for the duration of each piece. The skin conductance data were digitally sampled at 64 Hz and low-pass filtered at 3 Hz, using a BioSemi ActiveTwo system. Event markers sent to the BioSemi system from the stimulus presentation computer were used to synchronize the physiological responses with the musical pieces.

Similarly to Grewe at al. (2007), we only included in the analyses chills corresponding to both a button press and an SCR. This provided physiological support to participants' self-reported chills and minimized the possibility of demand or desirability biases. As the maximum increase in skin conductance was found to be 2.7 seconds after the button press (Grewe et al., 2009), and SCRs are known to occur at an onset latency of one to three seconds (Figner & Murphy, 2011), we identified SCRs occurring in a window of four seconds following a button press. SCRs were identified using Ledalab, a MATLAB-based software designed to identify SCRs by means of deconvolution (Benedek & Kaernbach, 2010). This software is particularly well-suited for the analysis of overlapping SCRs, which may arise in paradigms where SCR-inducing events are not presented at regularly spaced ITIs, as in the current experiment.

2.5. Procedure

Participants first provided informed consent. A "chill" was described to participants as a strongly emotional response to a point in music, experienced as a shiver down the spine or goosebumps (a definition modeled from Grewe et al., 2007). Participants sat in a comfortable chair in a sound-attenuated chamber facing a screen, which presented the post-piece questions. The stimuli were played through two Tannoy Reveal 601p studio monitor speakers (Tannoy Ltd, Coatbridge, UK) placed approximately one meter away from participants. After the skin conductance electrodes had been applied, participants listened to the relaxation sound (waves). During this time, the volume was adjusted to the participant's preference with the instruction that it should be as high as possible without being uncomfortable.

Using Presentation software (Neurobehavioral Systems, Inc., Albany, CA), the excerpts were presented in one of three pseudo-random orders following presentation of the practice excerpt. The orders were restricted as follows: 1) no more than two scrambled or two non-scrambled excerpts in a row, 2) no more than two excerpts from the same musical style in a row, 3) equal number of scrambled and non-scrambled excerpts in the first and second halves of the experiment, 4) equal number of excerpts from each style in the first and second halves of the experiment. Each of the three orders began with a non-scrambled stimulus from a different musical style. Participants were randomly assigned to one pseudo-random order prior to the experiment. Participants took a three-minute break half-way through the experiment. The post-piece questions were presented following each stimulus. After presentation of all stimuli, participants filled out the musical knowledge survey. The entire experiment lasted approximately 100 minutes.

2.6. Acoustic analyses

All stimuli were analyzed acoustically in order to determine which types of acoustic events corresponded with the occurrence of chills across styles. Using the MATLAB-based Music Information Retrieval Toolbox (Lartillot et al., 2008), we extracted the acoustic properties of the stimuli that, based on the limited previous research (Grewe et al., 2007), we hypothesized would be associated with the occurrence of chills. Specifically, we extracted properties indexing changes over time in loudness (amplitude envelope) and timbral quality (brightness, roughness). Brightness, measured through the spectral centroid, reflects the number of high frequencies in the spectrum. Roughness measures the amount of sensory dissonance in a sound. These acoustic properties were extracted by first decomposing the audio inputs of each stimulus into halfoverlapping 50 ms frames (thus leading to a 40 Hz sample rate), and then using the *mirenvelope*, *mirroughness* and *mirbrightness* functions with default parameters.

The relationship between these acoustic measures and the occurrence of chills was tested by cross-correlating the timeseries of each acoustic measure with a chills distribution timeseries obtained by combining chills from all participants for each stimulus. Because chill events were coded as zeros and ones (a one representing the presence of a button press in time corresponding to an SCR), a chill probability density function was generated by convolving a beta probability density function with the chill events over time (see Figure 1 for a visualization of this procedure). The beta probability density function was used as a proxy for the probability a chill was experienced at a specific time in the 2-second window preceding the button press. Smaller peaks in the chill probability density function thus correspond to fewer reported chills, whereas larger peaks indicate that chills were reported by multiple participants within a narrow time window. The cross-correlation of each feature timeseries with the chill probability density function thus tested the hypothesis that transient increases in the amount of that feature increased the probability that a participant would experience a chill. A peak in the cross-correlation indicates a high degree of correspondence between the acoustic measures and the chills distribution as well as the temporal lag between the two. We hypothesized that the cross-correlations would show a peak at or near zero lag, indicating that there was a peak in the acoustic measures preceding the button press used to indicate the occurrence of a chill. The results of the cross-correlations were smoothed using a moving average, in order to better capture the average value at zero lag irrespective of noise. The code used to perform these analyses is available online (see Supplementary material).

To test the cross-correlations at zero for significance, we compared the values of the cross-correlations at zero for matching pairs of acoustic measures and chills distributions (e.g., amplitude envelope of stimulus 1 vs. chills of stimulus 1) to mismatching pairs (e.g., amplitude envelope of stimulus 1 vs. chills of stimulus 2), which were hypothesized not to show any significant peak in the cross-correlation as there was no correspondence between the chills and the acoustic changes over time. We then also compared the values of the cross-correlations at zero lag for different types of stimuli to test whether the relationship between the acoustic measures and chills varied as a function of the piece style and whether it was scrambled. We hypothesized that if the same acoustic events are related to chills across cultures to the same degree, we would not find a significant difference across piece styles. Additionally, we hypothesized that if chills in the scrambled pieces were due to the preservation of some of the acoustic features associated with chills in the non-scrambled pieces, we would observe a peak in the cross-correlations of the scrambled pieces as well, though we remained agnostic as to whether this peak would be smaller or equal to that found for the non-scrambled pieces.

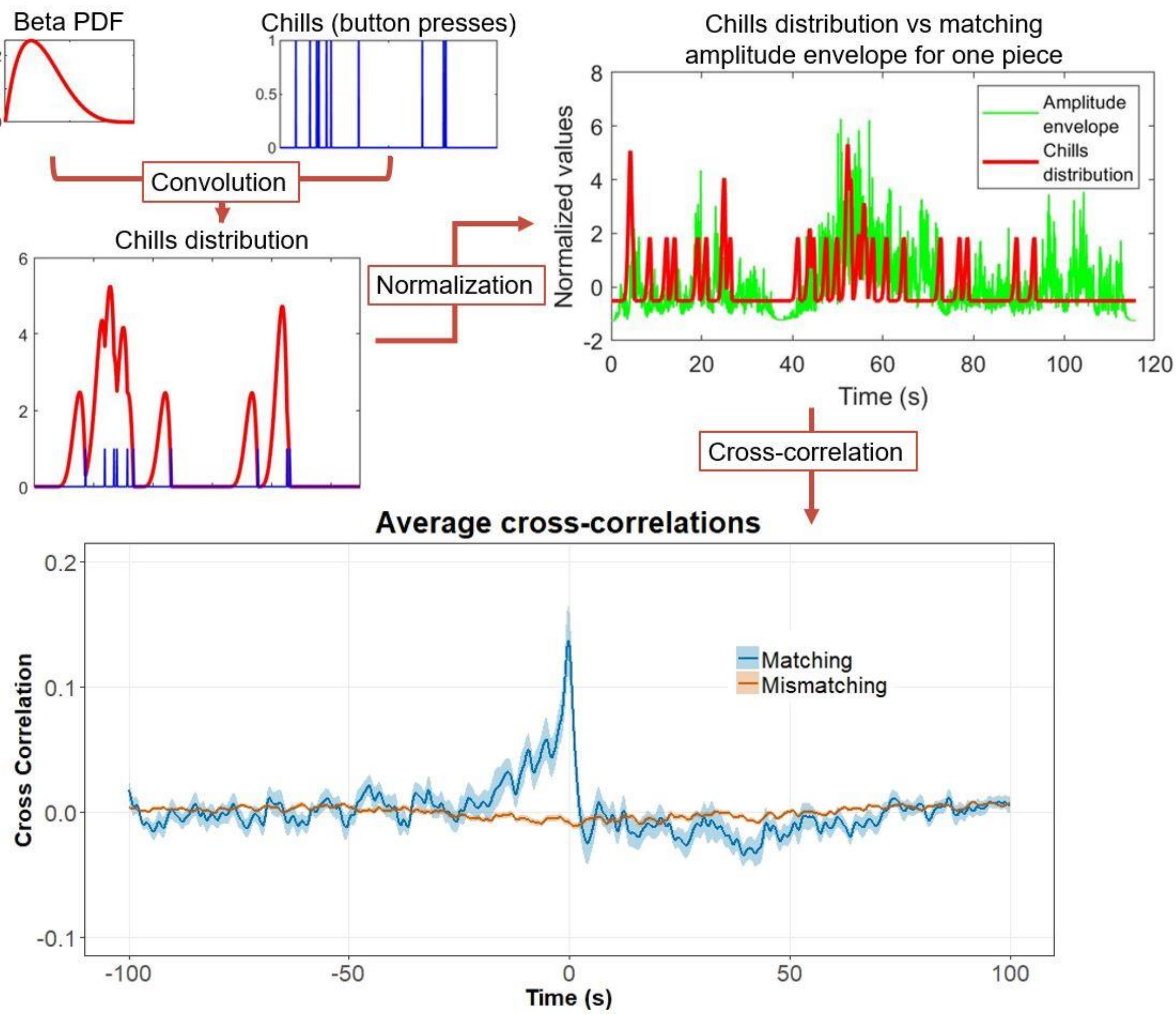

Figure 1. Visualization of the procedures involved in the acoustic analyses. For each piece, a chills distribution was obtained through the convolution of chills events and a beta probability density function. Using cross-correlation, this distribution was then compared to each acoustic measure of interest (e.g., amplitude envelope in this figure) after the two signals were normalized.

2.7. Summary of hypotheses

The primary question of this study was whether participants would feel as many chills to music from a style they are familiar with as compared to a style from an unfamiliar musical tradition. We hypothesized that participants would experience a similar number of chills to music of all styles, indicating that people not only recognize emotion in music of other cultures, but can also feel emotional responses to it. Alternatively, if people do not feel strong emotions in music of other cultures, we would expect participants to experience more chills to music from a familiar style than to music of other cultures. We also expected to find significantly fewer chills to scrambled music, which lacks musical structure.

In the analysis of the post-piece ratings, we hypothesized that if these ratings were affected by familiarity with a musical style (e.g., such that Western participants would like Western music more than other styles, and vice versa), then we would find a significant interaction of participant group and musical style on these ratings. If this interaction was significant, we planned to conduct multiple comparisons for a subset of contrasts to assess whether each group gave higher ratings to music of their own culture, and whether each musical style was rated higher by members of that group (see Results for more details on these comparisons). We hypothesized that ratings would be lower for the scrambled pieces, given that scrambled music is rated as less aesthetically pleasing (Lalitte & Bigand, 2006). We also analyzed whether there was a relationship between each of these post-piece ratings and the number of chills. We hypothesized that all measures would be positively correlated with the number of chills.

We analyzed the musical knowledge scores with the hypothesis that participants in the Chinese group would be more knowledgeable of Chinese music and participants in the Indian group would be more knowledgeable of Indian music compared to the other groups. We hypothesized that knowledge of Western music would be comparable across groups, as we assumed that all participants are familiar with Western music, while we remained agnostic as to whether knowledge of Western music theory would vary across groups. We also analyzed whether there was a relationship between musical knowledge and the number of chills to each style. Based on previous findings for Western classical music (Grewe et al., 2009), we did not hypothesize that there would be a relationship. Alternatively, if musical knowledge of a style had an effect on emotional responses to it, we expected to find that knowledge of each style would be positively correlated with the number of chills to music of that style.

Lastly, we performed acoustic analyses to assess whether the timing of sudden changes in amplitude envelope, brightness and roughness was correlated with the occurrence of chills across styles. We predicted that if changes in these acoustic measures lead to chills cross-culturally, then all three acoustic measures would be correlated to chills across styles. We also predicted that if chills to the scrambled pieces were due to the preservation of some of these acoustic features, and if these features led to chills independently of musical structure, then these measures would be correlated to chills in both the scrambled and non-scrambled music.

3. Results

Participants self-reported a total of 2,178 chills. We discarded button presses that did not correspond to a measured skin conductance response (SCR), leading to 1,004 chills that did correspond to an SCR (see the Discussion for speculations on the reason for this lower than expected correspondence between reported and measured chills). Additionally, we eliminated button presses occurring within a window of four seconds following each button press (independently for each participant), the length of the window chosen for the skin conductance response, since presses within that window could be associated with the same SCR; if multiple presses for one participant occurred within a four second window, only the first press was kept. This procedure led to a total of 910 chills included in the main analyses. Seven participants (three in the Western group, one in the Chinese group, and three in the Indian group) did not experience any chills. The percentage of participants experiencing chills in this experiment was therefore 89%, which is higher than what was found in previous studies (e.g., 55% in Grewe et al., 2007). This is likely because all participants in the present study had been screened for having experienced chills in the past.

3.1. Main analyses on number of chills

3.1.1. Effects of group, piece style, and version. A mixed-effects Poisson regression model with zero inflation was run on the number of chills, using the *glmmADMB* package (Fournier et al., 2012) in an R environment (R Core Team, 2018). Based on the hypotheses of the study, the model included the following fixed effects: piece version (Scrambled, Non-Scrambled), piece style (Western, Chinese, Indian) and participant group (Western, Chinese, Indian). Random effects were included for participant and piece. Dummy coding was used to set

Western style and group as the baseline against which the other factor levels were compared. We observed only a significant main effect of piece version (β = -1.02, SE = 0.11, $p < 0.0001$). The effects of piece style and participant group were not significant. To test for an interaction between style and group, we ran the same model as above but included this interaction. While two interaction contrasts resulting from this model were significant (Western group, Western style vs. Chinese group, Indian style: β = 0.57, SE = 0.22, $p < 0.01$; Western group, Western style vs. Indian group, Indian style: β = 0.50, SE = 0.22, $p < 0.05$), the overall significance of the interaction was determined by testing whether the model with an interaction was a significantly better fit of the data compared to the model without an interaction. A log likelihood ratio test revealed that there was no significant difference between the two models, indicating that adding an interaction did not significantly improve the model fit. Because the two significant interaction contrasts were not in a direction hypothesized a priori and because the overall interaction was not significant, we concluded that they are likely to be spurious results and they are not discussed further. For other analyses in the rest of the paper we only report the results of the model comparison as an assessment of the overall interaction and we perform post-hoc multiple comparisons when it is significant.

Overall, the results of these models support our hypotheses, suggesting that, while participants experienced significantly fewer chills to the scrambled than the non-scrambled music, the musical style and the participant group did not have an effect on their emotional responses. More importantly, as there was no significant interaction between style and participant group, we found no evidence for a difference in the number of chills participants experienced when listening to music from a familiar musical style as compared to music from other cultures with which they were not familiar. The average number of chills, along with individual data points, is visualized in Figure 2.

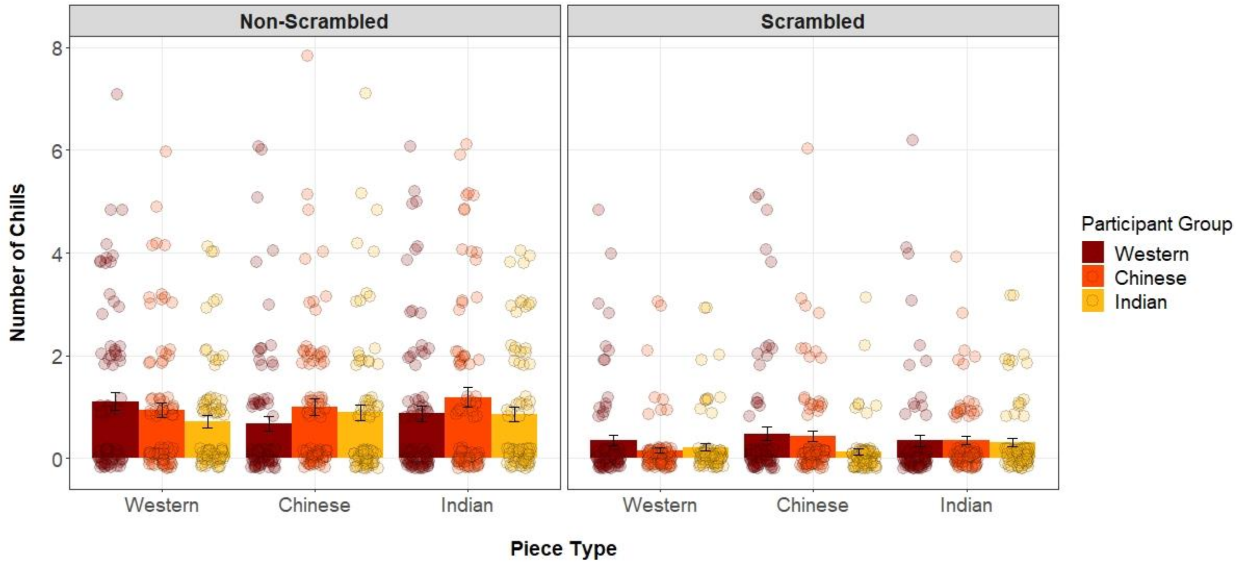

Figure 2. Effects of piece version (Scrambled, Non-Scrambled), piece style (Western, Chinese, Indian) and participant group (Western, Chinese, Indian) on chills. Error bars represent the Standard Error of the Mean. Individual data points (the number of chills for each piece for each participant) are jittered both horizontally and vertically for better visualization.

3.1.2. Bayesian analysis. As our hypotheses included null effects, we used Bayesian tests in order to further validate our observed null effects of style and group, as well as their interaction. Using the Stan computational framework (http://mc-stan.org/) accessed with the *brms* package for R (Bürkner, 2017a, 2017b), we ran a mixed-effects model with the same specifications as above. The model without an interaction found that the only effect for which the credible interval did not include 0 was that of piece version[3] (β = -1.02, 95% credible interval [-

[3] Credible intervals are the Bayesian counterpart of frequentist confidence intervals. The 95% credible interval has a 95% probability of containing the true value being estimated, such as the difference between the number of chills to the scrambled vs. non-scrambled pieces in this case. Thus, if the credible

1.29, -0.77]). We tested for the interaction using the same procedure as above, by running a new model that includes the interaction and then comparing the fit of the two models. In order to compare these two Bayesian models, we used leave-one-out cross-validation (LOO) using the R package *loo* (Vehtari et al., 2016; 2017). As the difference in the LOO information criterion (LOOIC, which is computed from the expected log predictive density) resulting from this analysis was very small (*LOOIC difference* = 1.87, *SE* = 7.71) we can conclude that the model with an interaction was not a better fit of the data than the model without the interaction (Bürkner, 2017a). Thus, the Bayesian tests yielded the same results as the conventional models discussed above, strengthening the validity of our hypothesized null findings.

3.2. Post-piece ratings analyses

3.2.1. Effects on post-piece ratings. Four separate mixed effects models were run on each of the post-piece ratings (Liking, Attention, Emotionality and Affect Mood) in an R environment using the *lme4* package (Bates et al., 2014) and *lmerTest* package (Kuznetsova et al., 2017). The models included random effects of participant and piece and fixed effects of piece version, style, and participant group. For all four ratings, scrambled pieces received significantly lower scores than non-scrambled pieces (Liking: $\beta = -1.91$, $SE = 0.08$, $p < 0.0001$; Attention: $\beta = -1.42$, $SE = 0.07$, $p < 0.0001$; Emotionality: $\beta = -1.78$, $SE = 0.09$, $p < 0.0001$; Affect Mood: $\beta = -1.16$, $SE = 0.10$, $p < 0.0001$). Other main effects are summarized in Table 2. The overall pattern is that Western pieces received higher ratings compared to the Chinese pieces in all four measures, and compared to the Indian pieces in all measures except for Affect Mood (see Table 2 for statistical values and Figure 3 for visualization of the data).

---

interval does not contain 0, we can be fairly sure that this difference is not 0 (Morey et al., 2016; Nicenboim & Vasishth, 2016).

In order to test whether the effects of participant group and musical style interacted, we re-ran these models including this interaction and then compared their fit to the models without the interaction. Log likelihood ratio tests revealed a significantly better fit of the model with an interaction for all four measures (Liking: $\chi^2(4) = 43.38$, $p < 0.0001$; Attention: $\chi^2(4) = 28.17$, $p < 0.0001$; Emotionality: $\chi^2(4) = 13.2$, $p < 0.05$; Affect Mood: $\chi^2(4) = 12.42$, $p < 0.05$). Significance for a select number of contrasts based on our hypotheses was assessed using the *lsmeans* package in R with Tukey HSD correction (Lenth, 2017). In particular, within-group contrasts tested whether each group rated higher the music they had reported being most familiar with compared to music of unfamiliar styles (e.g., whether the Western group liked Western music better than Chinese and Indian music, and vice versa for each group). Between-group contrasts tested whether each style of music was rated higher by the group familiar with it than by the groups unfamiliar with it (e.g., whether the Western music was liked better by the Western group than by the Chinese and Indian groups, and vice versa for each style). Significantly different contrasts are reported in Table 2. The overall pattern shows that the Western group gave higher Liking and Emotionality ratings to Western music than the other two groups, and the Chinese group gave lower Attention ratings to Indian music than Chinese music and as compared to the Indian group. While these particular contrasts are in accordance with the hypothesis that each group would give higher ratings to the music they are familiar with, this was not a general pattern found for all groups and all measures.

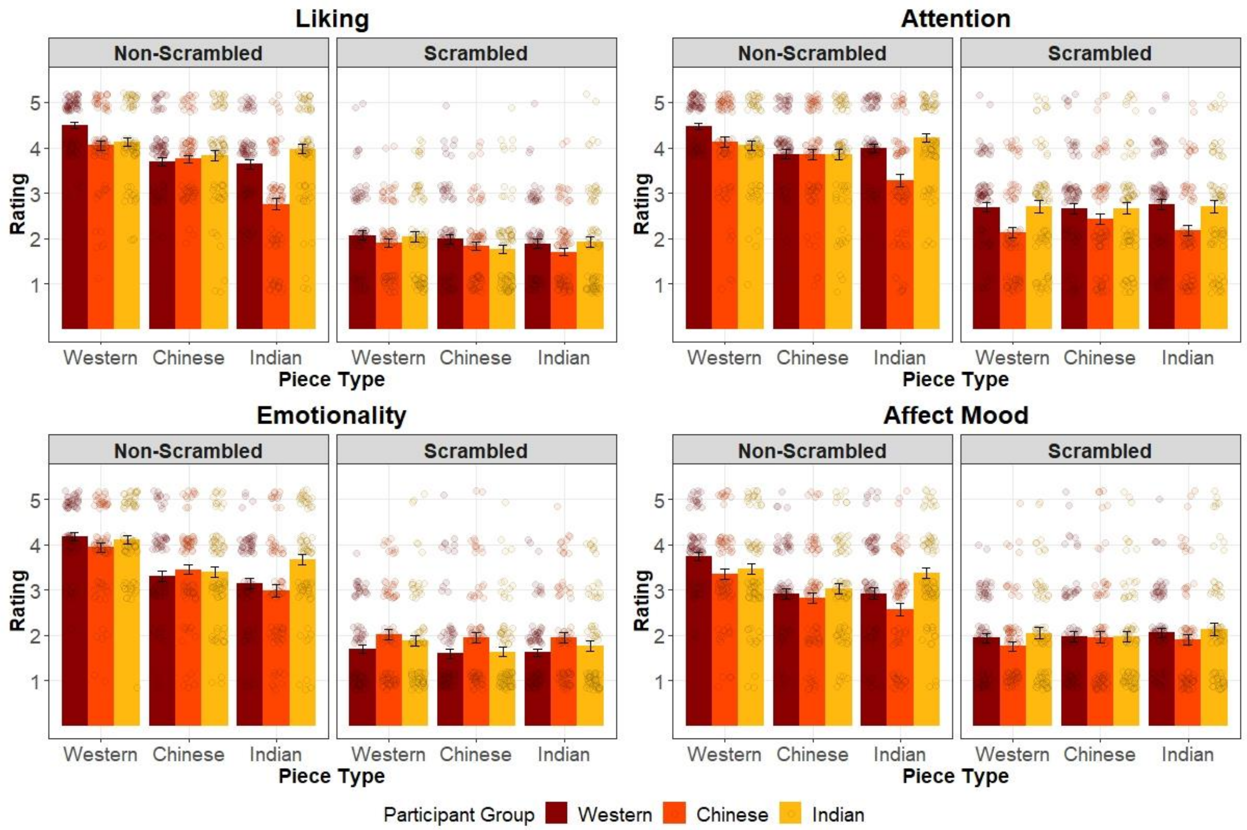

Figure 3. Effects of piece version (Scrambled, Non-Scrambled), piece style (Western, Chinese, Indian) and participant group (Western, Chinese, Indian) on post-piece ratings of Liking, Attention, Emotionality and Affect Mood. Error bars represent the Standard Error of the Mean. Individual data points are jittered both horizontally and vertically for better visualization.

Table 2

Significant contrasts for the post-piece ratings

Dependent variable	Contrast	*β*	*SE*	*p* <
Liking	Version (S - N)	-1.91	0.08	0.0001
	Style (C - W)	-0.31	0.10	0.01
	Style (I - W)	-0.47	0.10	0.0001
	Group*Style (W,W - W,C)	0.45	0.12	0.05
	Group*Style (W,W - W,I)	0.52	0.12	0.01
	Group*Style (C,C - C,I)	0.57	0.12	0.001
	Group*Style (C,I - I,I)	-0.72	0.17	0.01
Attention	Version (S - N)	-1.42	0.07	0.0001
	Group (C - W)	-0.40	0.18	0.05
	Style (I - W)	-0.18	0.08	0.05

	Group*Style (C,C - C,I)	0.41	0.11	0.05
	Group*Style (C,I - I,I)	-0.73	0.20	0.05
Emotionality	Version (S - N)	-1.78	0.09	0.0001
	Style (C - W)	-0.41	0.11	0.01
	Style (I - W)	-0.45	0.11	0.001
	Group*Style (W,W - W,C)	0.49	0.13	0.05
	Group*Style (W,W - W,I)	0.55	0.13	0.01
Affect Mood	Version (S - N)	-1.16	0.10	0.0001
	Style (C - W)	-0.27	0.12	0.05

Note: S = scrambled; N = non-scrambled; W = Western; C = Chinese; I = Indian. Each interaction contrast should be read as: [group, style] - [group, style] (e.g., W,W - W,C = Western group, Western style - Western group, Chinese style). Only significant contrasts hypothesized a priori are reported. Interaction contrast p-values are adjusted using Tukey's HSD.

3.2.2. Relationship between post-piece ratings and chills. A mixed-effects Poisson regression model with zero inflation was run on the number of chills with random effects of participant and piece and fixed effects of Liking, Attention, Emotionality and Affect Mood. It revealed a relationship between the number of chills and ratings of Attention ($\beta = 0.34$, $SE = 0.06$, $p < 0.0001$), Emotionality ($\beta = 0.11$, $SE = 0.05$, $p < 0.05$) and Affect Mood ($\beta = 0.25$, $SE = 0.05$, $p < 0.0001$), but not of Liking, indicating that the extent to which participants liked the music was not related to how many chills they experienced in response to it.

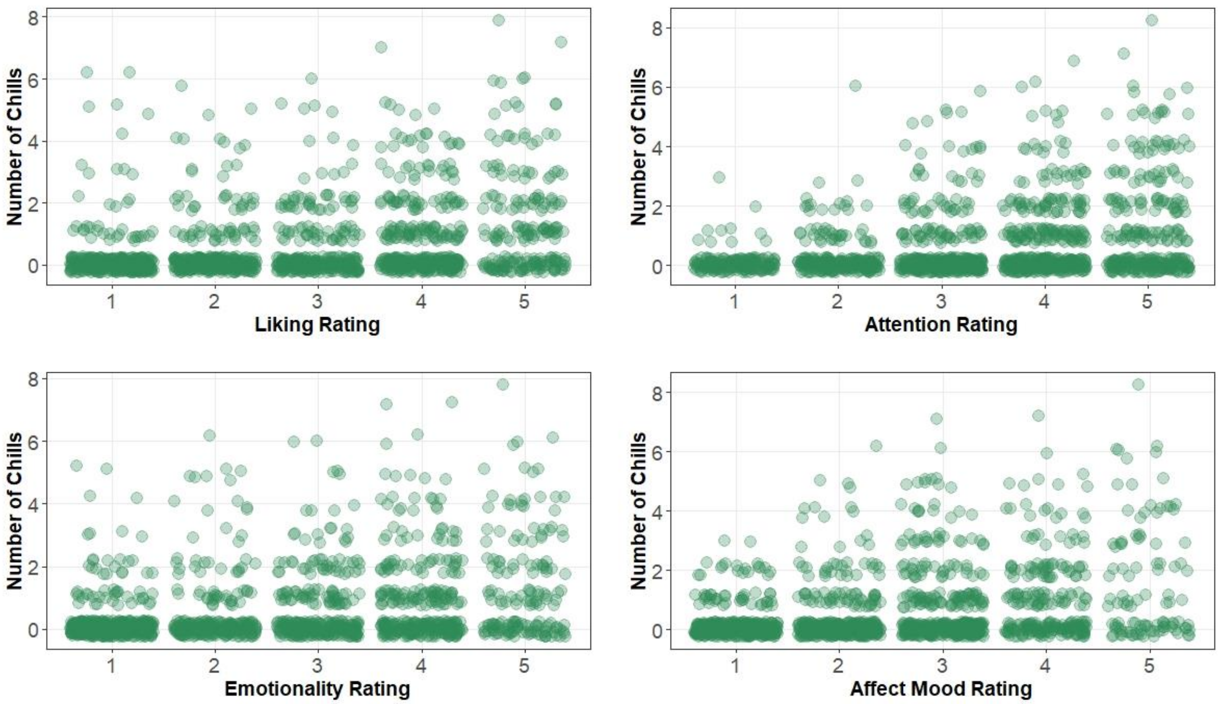

Figure 4. Relationship between post-piece ratings and number of chills. Individual data points are jittered both horizontally and vertically for better visualization.

3.3. Musical knowledge survey analyses

For each participant, scores from the knowledge survey were computed by counting the number of correct responses for each subset of the survey (Western music, Chinese music, Indian music, and Western music theory). The maximum possible score (the number of questions) was 9 for the first three, and 7 for the latter.

3.3.1. Effects of participant group on knowledge. Four ANOVAs[4] were run on Western music knowledge, Chinese music knowledge, Indian music knowledge, and music theory knowledge, respectively, using participant group as the only predictor. The ANOVA on

[4] We used ANOVAs for this analysis as opposed to mixed effects models because we had only a single data point per participant for each dependent variable.

Western music knowledge showed no significant effect of participant group, indicating that Western music knowledge did not vary significantly across groups. The ANOVA on Chinese music knowledge revealed a main effect of group ($F(2, 59) = 41.72$, $p < 0.0001$). Post-hoc comparisons using the Tukey HSD test show that the Chinese group had significantly better knowledge of Chinese music ($M = 3.76$, $SD = 2.3$) compared to the Western group ($M = 0.33$, $SD = 0.58$) and the Indian group ($M = 0.25$, $SD = 0.55$), while there was no difference between the Western and Indian groups. The ANOVA on Indian music knowledge revealed a main effect of group ($F(2, 59) = 14.81$, $p < 0.0001$). Post-hoc tests using the Tukey HSD test show that the Indian group had significantly better knowledge of Indian music ($M = 1.35$, $SD = 1.04$) compared to the Western group ($M = 0.43$, $SD = 0.68$) and the Chinese group ($M = 0.14$, $SD = 0.36$), while there was no difference between the Western and Chinese groups. Lastly, the ANOVA on music theory knowledge showed no significant effect of participant group, indicating that music theory knowledge did not vary significantly across groups. Overall, these analyses support the hypothesis that participants in the Chinese group had greater knowledge of Chinese music and that participants in the Indian group had greater knowledge of Indian music. They also support our assumption that all participants would be familiar with Western music, and indicate that this familiarity extended to knowledge of Western music theory. See Figure 5 for a visualization of the data.

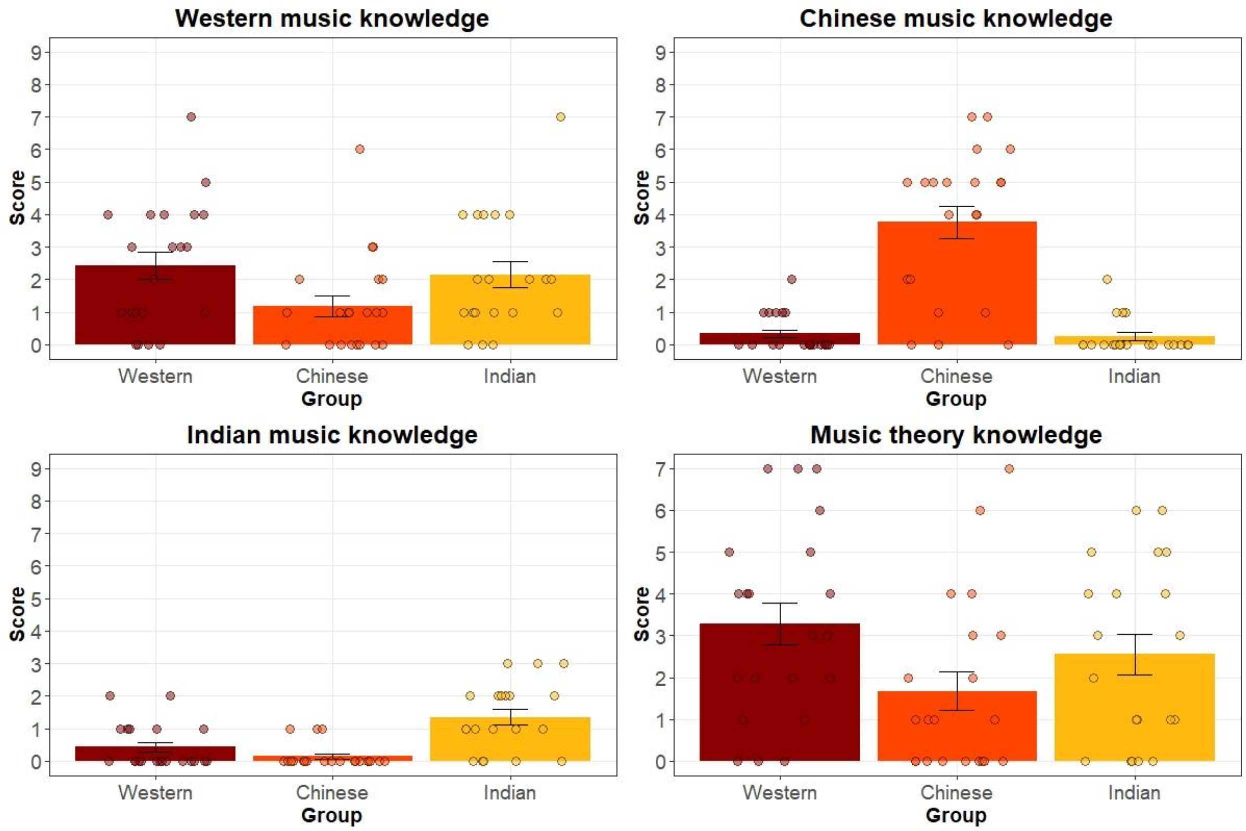

Figure 5. Musical knowledge scores for each subset of the survey, compared across groups. The highest possible score (the number of questions) was 9 for Western, Chinese and Indian music, while it was 7 for music theory. Error bars represent the Standard Error of the Mean. Individual data points are jittered horizontally.

3.3.2. Relationship between knowledge and chills. The analysis of whether musical knowledge affected the number of chills was conducted on non-scrambled pieces only. This is because scrambled pieces lack the musical and tonal structure of the style they had prior to being scrambled, and so the knowledge participants might have is not relevant or applicable. We therefore did not include chills to scrambled pieces in this analysis.

Three separate Poisson regression models with zero inflation were conducted on the total number of chills to Western, Chinese and Indian pieces, respectively. Each model included fixed effects of Western music knowledge, Chinese music knowledge, Indian music knowledge, and music theory knowledge. The model on chills to Western pieces found a significant effect of music theory knowledge (β = 0.11, *SE* = 0.04, $p < 0.01$). The model on chills to Chinese pieces found a significant effect of Chinese music knowledge (β = 0.09, *SE* = 0.03, $p < 0.01$). The model on chills to Indian music found significant effects of Chinese music knowledge (β = 0.16, *SE* = 0.03, $p < 0.001$), Indian music knowledge (β = 0.26, *SE* = 0.08, $p < 0.01$), and music theory knowledge (β = 0.09, *SE* = 0.04, $p < 0.05$). These results lend partial support to the alternative hypothesis that knowledge of each style would be associated with more chills to music from that style.

3.4. Acoustic analyses

3.4.1. Matching vs. mismatching. Cross-correlations were performed between the chills distribution and the acoustic measures for each stimulus, comparing matching and mismatching pairs (see Method; see Figure 6). For Amplitude Envelope, Brightness and Roughness, cross-correlations revealed a peak near zero lag for the matching but not the mismatching pairs. T-tests were used to compare the value of the cross-correlations at zero lag for the matching and mismatching pairs, revealing a significant difference for all three measures (Envelope: $t(23.4)$ = 5.12, $p < 0.0001$; Brightness: $t(23.5)$ = 2.1, $p < 0.05$; Roughness: $t(23.2)$ = 4.75, $p < 0.0001$).

3.4.2. Scrambled vs. non-scrambled. The cross-correlations at zero lag were compared for scrambled and non-scrambled stimuli averaging across styles (see Figure 7). For all three measures, cross-correlations revealed a peak near zero lag for the non-scrambled but not the scrambled stimuli. Paired t-tests were used to compare the value of the cross-correlations at zero lag for the non-scrambled and scrambled pairs, revealing a significant difference for all three measures (Envelope: $t(11) = 6.24$, $p < 0.0001$; Brightness: $t(11) = 2.63$, $p < 0.05$; Roughness: $t(11) = 6.5$, $p < 0.0001$).

3.4.3. Piece style. The cross-correlations at zero lag were compared for stimuli in the Western, Chinese and Indian styles (see Figure 8). This analysis averaged across scrambled and non-scrambled versions of the stimuli because our a priori hypotheses did not include a strong effect of this factor on the cross-correlations. A peak near zero lag was found for all measures across styles. For all three measures, ANOVAs revealed no significant difference in the value of the cross-correlations at zero lag across piece styles (all p's > 0.05). As we found that there was a strong effect of whether the piece was scrambled on the cross-correlations, we performed an exploratory analysis testing whether there were any effects of musical style when excluding the scrambled pieces. ANOVAs on the cross-correlations at zero lag for only non-scrambled pieces again revealed no significant differences across styles (all p's > 0.05).

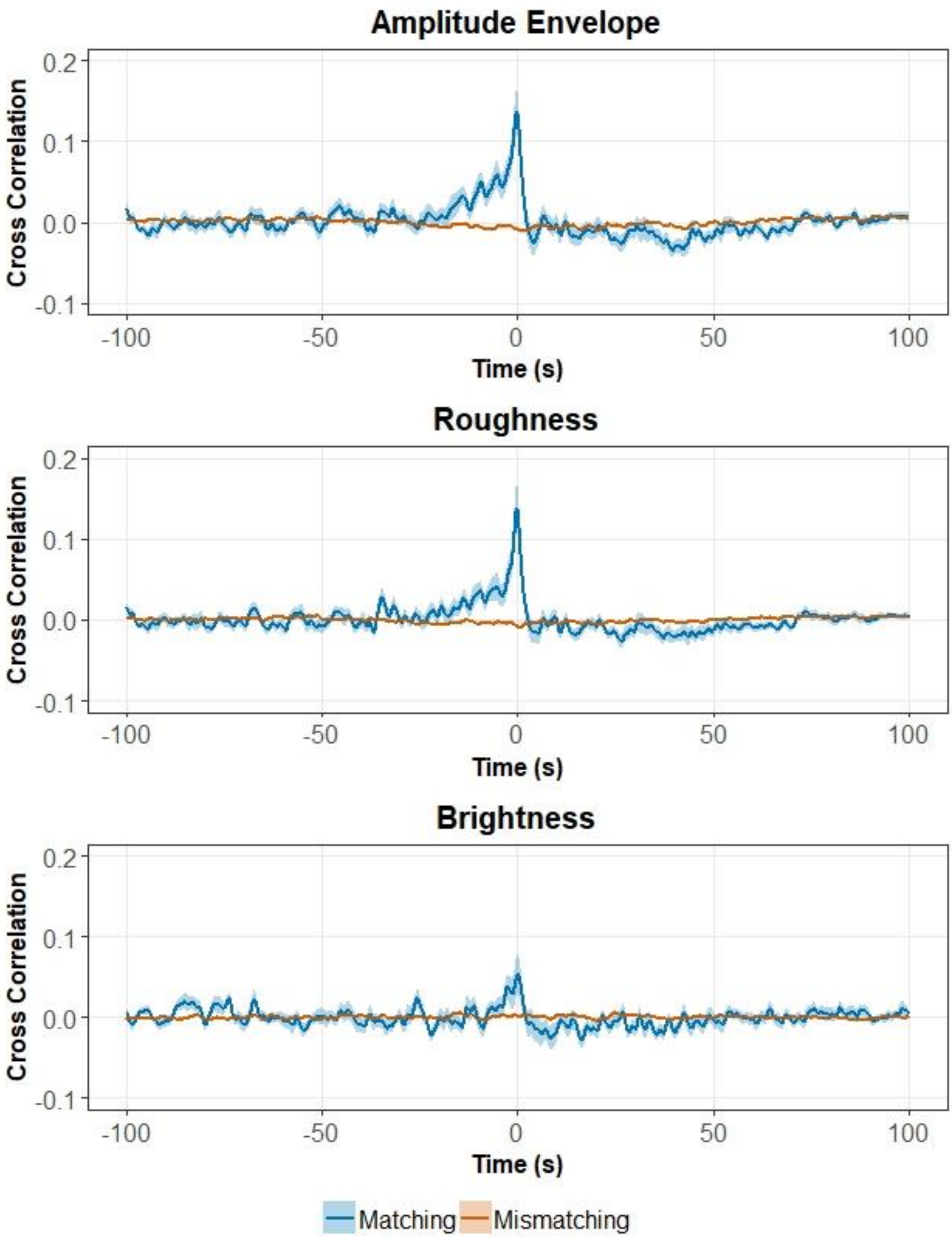

Figure 6. Cross-correlations between chills and acoustic measures for matching and mismatching

pairs, averaged across stimuli. Shaded areas represent Standard Error of the Mean. A peak at zero lag indicates that there was a sudden increase in the acoustic measures right before the button press indicating a chill (see Method).

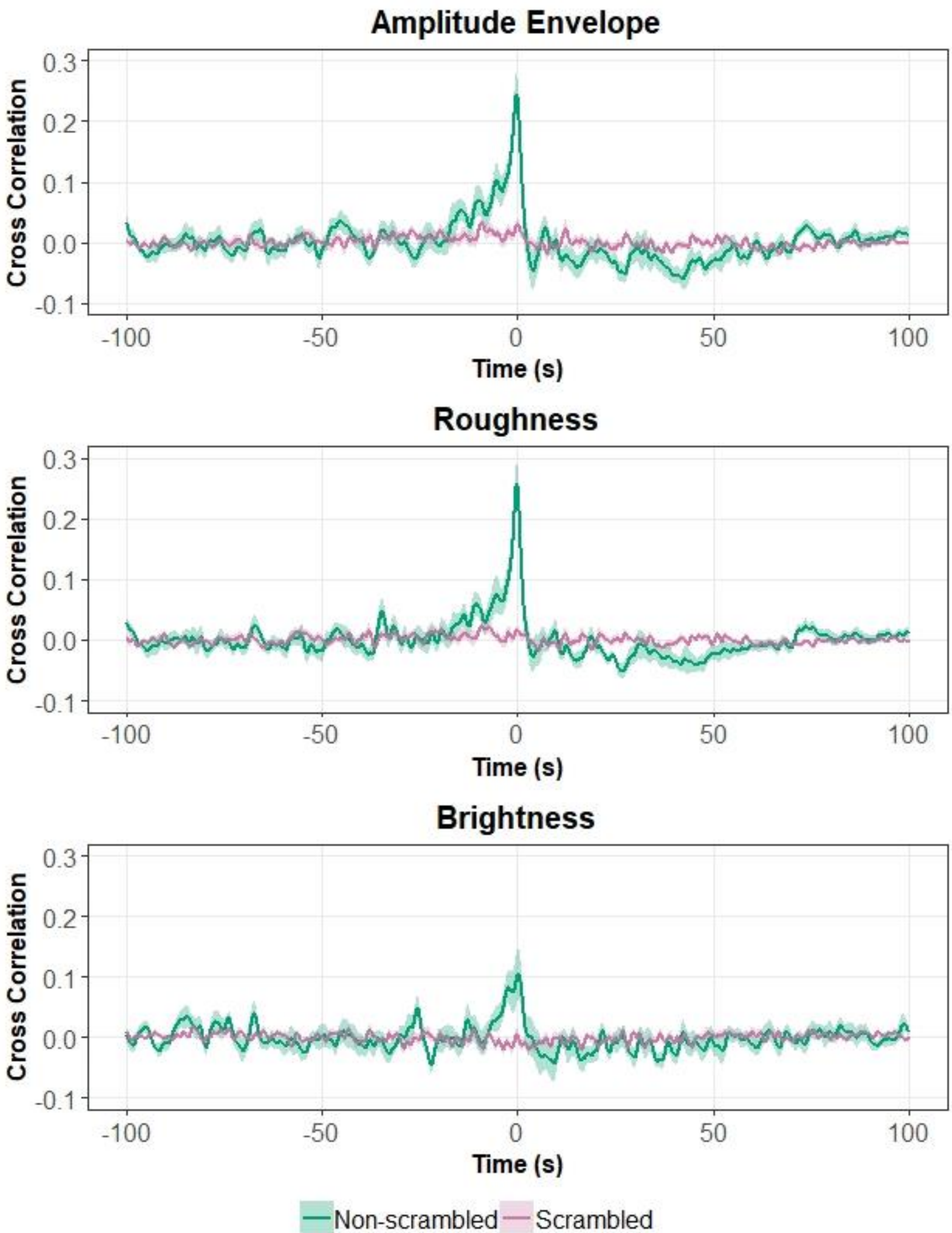

Figure 7. Cross-correlations between chills and acoustic measures averaged by version

(scrambled, non-scrambled). Shaded areas represent Standard Error of the Mean.

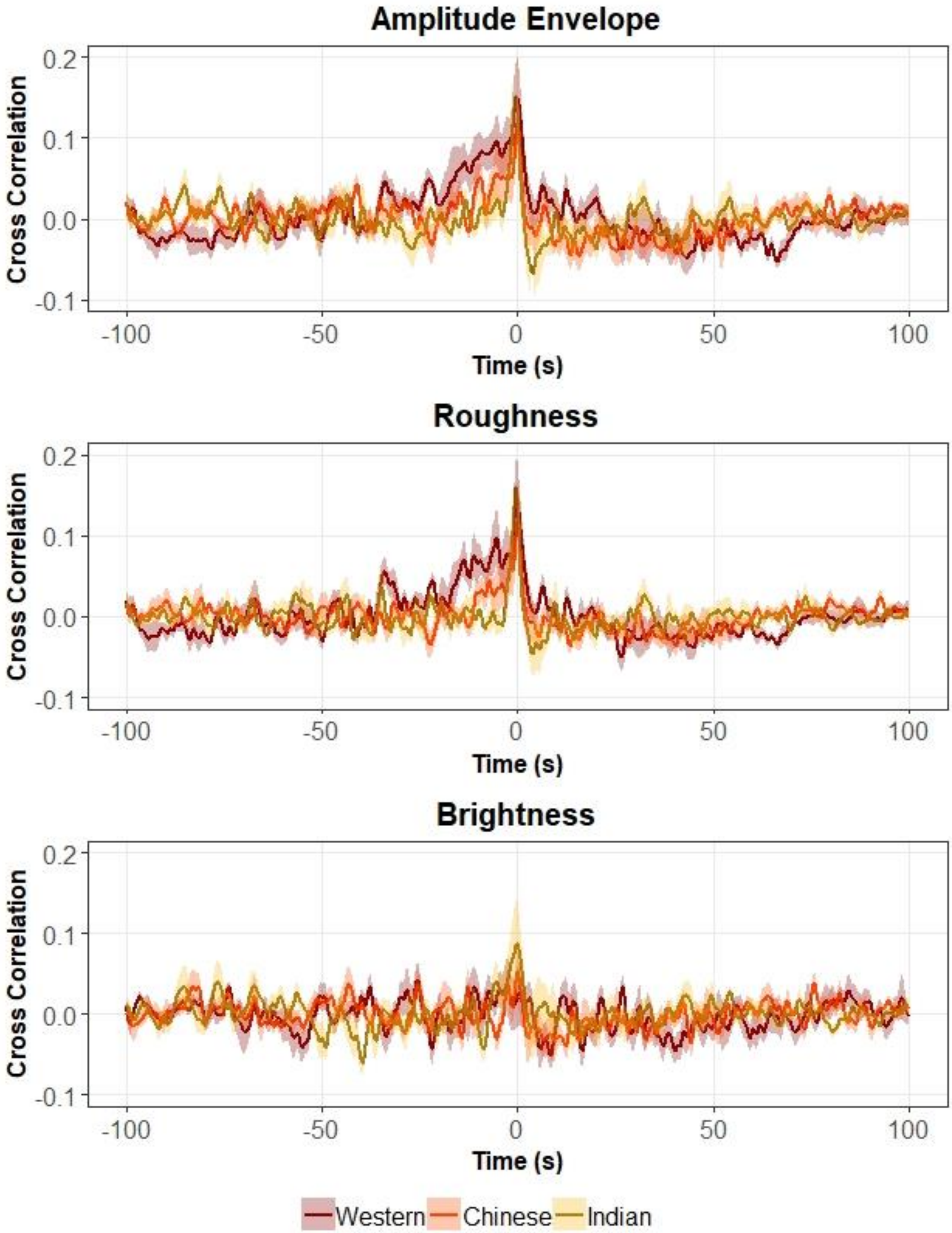

Figure 8. Cross-correlations between chills and acoustic measures averaged by musical style.

Shaded areas represent Standard Error of the Mean.

3.4.4. Exploratory analysis of scrambled pieces. We performed an exploratory analysis to determine what led to the large difference in cross-correlations at zero lag between scrambled and non-scrambled pieces. Based on the idea of "ramp archetypes" found for Western music, where attention to music is best maintained by slow increases in intensity followed by sudden drops in intensity (Huron, 1992), we expected that the acoustic events leading to chills would be relatively slow acoustic modulations lasting on the order of seconds. As the scrambling procedure randomly reordered 250 ms long segments in time, we predicted that this procedure could have disrupted such low-frequency periodicities.

In order to explore this possibility, we compared the power spectral density (psd) estimate for the Amplitude Envelope and Roughness of scrambled and non-scrambled pieces. The psd function indicates the signal's power as a function of frequency; thus, it can be used to measure the strength of energy modulations at each frequency. The psd was estimated using Welch's method with a 10-second long Hamming window and normalized by total power. We did not similarly analyze Brightness because this measure included missing values.

Visual inspection of the psd estimates for scrambled and non-scrambled pieces (see Figure 9) confirmed our expectation that the scrambling procedure disrupted low-frequency periodicities, while also suggesting that the scrambling introduced modulations at frequencies higher than 0.5 Hz. We performed paired t-tests on power averaged across two frequency bands: lower frequencies (0 - 0.5 Hz) and higher frequencies (0.5 - 1 Hz). We found that non-scrambled pieces had significantly more power at lower frequencies than scrambled pieces for both measures (Envelope: $t(11) = 11.4$, $p < 0.0001$; Roughness: $t(11) = 9.24$, $p < 0.0001$). Non-scrambled pieces also had significantly less power at higher frequencies than scrambled pieces for Amplitude Envelope (*t(11)* = -6.9, *p* < 0.0001) but not for Roughness (*p* > 0.05). These analyses suggest that the scrambling procedure did not preserve the slower modulations in Amplitude Envelope and Roughness that were found to correlate with chills.

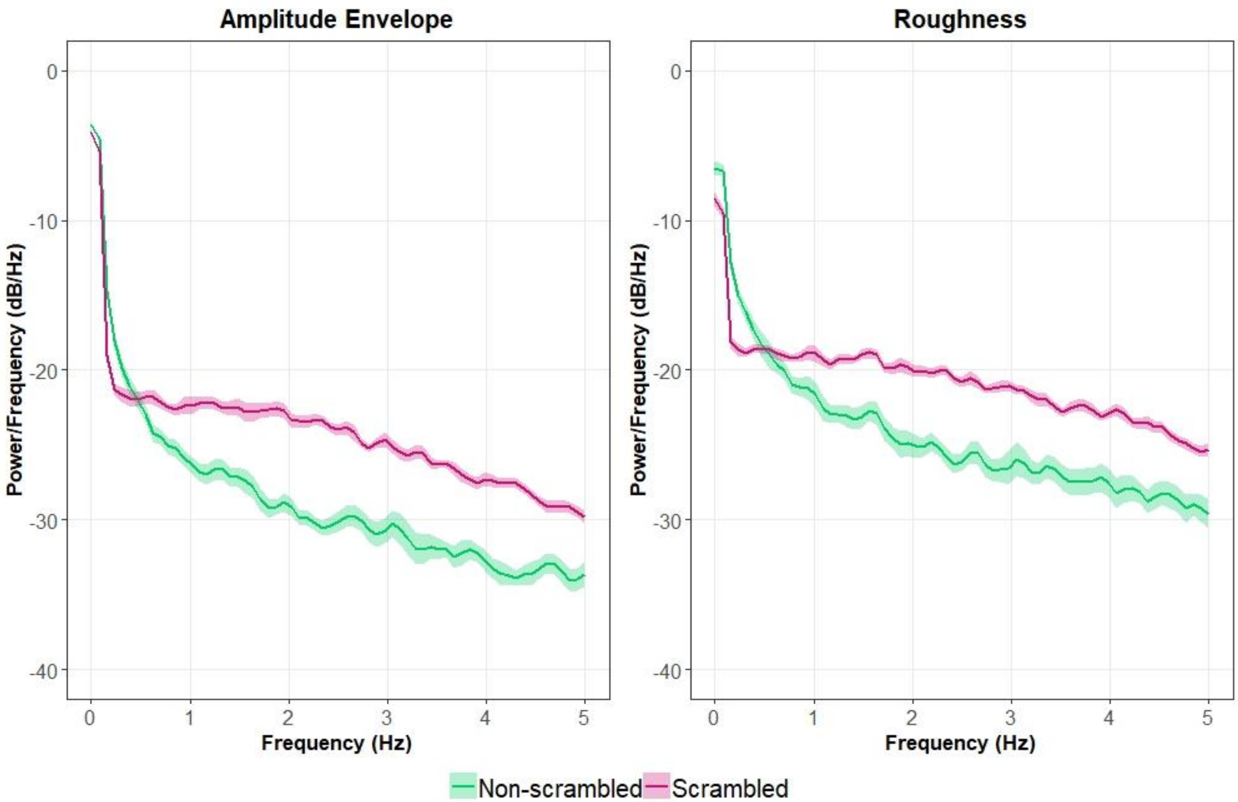

Figure 9. Power spectral density (psd) estimates for scrambled and non-scrambled pieces. Values are converted to the dB scale for plotting. Shaded areas represent Standard Error of the Mean.

3.5. Summary of results

The results of our primary analysis showed that, in support of our hypothesis, there were no significant differences in the number of chills that participants felt across styles, and participants felt significantly fewer chills to the scrambled music. The strength of the null hypothesis was validated using Bayesian analyses.

Analysis of the post-piece questions showed that participants gave higher ratings to the non-scrambled than the scrambled pieces for all measures (Liking, Attention, Emotionality and Affect Mood). While there was an interaction between participant group and musical style on these ratings, multiple comparisons only provided weak support for the hypothesis that participants would give higher ratings to music from a familiar style. Lastly, we found that the number of chills was positively correlated with ratings of Attention, Emotionality and Affect Mood, but not Liking.

Analysis of the musical knowledge scores found higher within-group knowledge of Chinese and Indian music, but not of Western music and Western music theory. We found that knowledge was related to the number of chills, partially supporting the idea that knowledge of each style leads to more chills to music of that style. In particular, we found that chills to Western music were related to music theory knowledge, chills to Chinese music to Chinese music knowledge, and chills to Indian music to Chinese, Indian and music theory knowledge.

The acoustic analyses revealed that the timing of peaks in Loudness, Brightness and Roughness was correlated with chills across styles, but not for the scrambled pieces. Exploratory acoustic analyses of the scrambled pieces suggest that this may be due to these acoustic events being disrupted through the scrambling procedure.

4. Discussion

This study addressed the question of whether listeners can experience strong emotional responses to music in absence of familiarity with the musical structures or knowledge of the cultural context and conventions through which a certain musical style has developed. We found that there were no reliable differences in the number of chills participants felt in response to styles of music of different cultures regardless of their own self-reported familiarity with each style, though all participants felt significantly more chills to non-scrambled music than to scrambled music.

The number of chills was also related to post-piece ratings measures; specifically, a higher number of chills corresponded to higher ratings of self-reported attention paid to the music, the extent to which participants found the music to be emotional, and the extent to which the music had affected their own mood. On the other hand, the number of chills was not related to how much participants reported liking the music, contrary to our hypothesis and previous literature (Panksepp, 1995; Salimpoor et al., 2009). With respect to these post-piece measures themselves, we found that all ratings were significantly higher for the non-scrambled than the scrambled music, confirming the assumption that scrambled music is less enjoyable and emotion-inducing. We found that, overall, participants gave similar ratings to music of all three styles.

We also measured participants' knowledge of each style of music through a questionnaire at the end of the experiment. Analysis of participants' knowledge scores supported the assumption that participants who had reported being familiar with either Chinese or Indian music (thus categorized into the Chinese and Indian groups) had higher knowledge of these styles compared to the Western group. On the other hand, knowledge of Western music and of Western music theory was comparable across groups, likely reflecting the fact that all participants lived in the United States. We found partial support for the idea that knowledge of each style would be associated with more chills to music of that style. This was true for chills to the Chinese music and Indian music. It is less clear why chills to the Indian music were also associated with higher knowledge of Chinese music and Western music theory. Chills to the Western music were associated with higher knowledge of Western music theory. Knowledge of Western music was not associated with chills in response to any style, suggesting that this shared knowledge likely did not impact listeners' emotional reactions in this experiment, although the fact that all participants were somewhat familiar with Western classical music (as opposed to the other styles) may pose a limitation to this conclusion. Nonetheless, this finding is in support of previous research on Western music that also found no relationship between chills and Western music knowledge (Grewe et al., 2009). This points to the idea that a better understanding of the mechanisms underlying music cognition would be achieved by sampling from as broad a range of musical styles and cultures as possible, as the results we find for Western music may not apply to other musical styles.

Lastly, we performed acoustic analyses of the pieces. We found that the timing of sudden peaks in loudness, brightness and roughness was correlated with the timing of chills across musical styles, but not in the scrambled music. Thus, we conclude that similar acoustic events can induce strong emotional responses across cultures, suggesting a commonality in the way music conveys these types of emotion.

While previous research indicates that people can infer the intended mood expressed by music of other cultures and can form tonal and melodic expectancies thought to be at the base of emotional responses, this study demonstrates that music can also lead to strong, felt emotions regardless of one’s cultural background. As felt and recognized emotion in music are linked but separable (Juslin & Laukka, 2004; Salimpoor et al., 2009; Evans & Schubert, 2008), this study adds an important piece to our understanding of the degree to which musical meaning is conveyed cross-culturally. This is supported by our data showing that chills were related to post-piece ratings of how much attention participants had paid to the music and how much it had affected their mood, but not to how much they liked it. These results suggest that the presence of felt emotion in music cannot be inferred just from asking participants to rate its pleasantness (Fritz et al., 2009). Instead, measuring strong felt emotions such as chills, which can be corroborated by physiological measures such as skin conductance, is an effective way of determining whether music induces felt emotional responses.

A second important addition to the literature is that acoustic events such as peaks in loudness, roughness, and brightness were associated with chills across our three styles of music. Cross-cultural studies of emotion recognition in music suggest that some psychoacoustic properties of music communicate emotion cross-culturally (Balkwill & Thompson, 1999; Laukka et al., 2013). Our results support this claim by showing that the acoustic events associated with chills within Western music (Grewe et al., 2007; Panksepp, 1995) also lead to chills in music from other cultures.

Interestingly, we found that there was no correlation between the acoustic measures and chills for the scrambled pieces. This was contrary to our hypothesis that chills to the scrambled music could be attributed to the preservation of some acoustic events. Exploratory analyses of the scrambled pieces suggested that this lack of correlation may be due to the fact that the scrambling procedure disrupted the types of acoustic events (slow modulations in amplitude envelope and roughness) that correlated with chills. To fully understand why acoustic events do not seem to correlate with chills in scrambled music, however, future research should carefully manipulate the degree to which acoustic events are preserved while disrupting musical structure through a range of scrambling procedures.

4.1. Possible mechanisms

One way to view the current findings is through the framework of predictive coding, whereby maximum pleasure in music (and, by extension, chills) is achieved by music that is neither too predictable nor unpredictable (Vuust & Witek, 2014; Witek et al., 2014). If music were unintelligible without knowledge of its cultural context, there would be no difference between listening to music of other cultures and scrambled music, as both would be highly unpredictable signals. However, the number of chills induced by music of other cultures was higher than that induced by the scrambled music, and similar to that induced by music from a familiar style. Thus, we may infer that music of other cultures is not an entirely unpredictable signal such as the scrambled music. However, more research is needed to assess whether chills were only induced by unexpected acoustic events such as the ones we measured across styles, or whether participants also experienced chills in response to tonal expectations formed through quick incidental learning (Oram & Cuddy, 1995; Rohrmeier & Widdess, 2012) or based on shared Gestalt principles (Narmour, 1990).

While the literature overall suggests that chills correspond to unexpected events such as peaks in loudness or an unusual harmony (Grewe et al., 2007; Guhn et al., 2007; Panksepp, 1995; Sloboda, 1991), it is not entirely clear how these events are linked to emotion. It has been posited that chills may arise through two mechanisms: a slow top-down evaluation, and a quick bottom-up reaction (Grewe et al., 2007; Harrison and Loui, 2014; Juslin & Västfjäll, 2008; Juslin, 2013). The slower process follows the encoding of other emotional responses to the music and is modulated by factors such as the listener's personality and the listening circumstances. The faster process may arise from brainstem reflexes leading to fast changes in arousal through the autonomic nervous system. While these processes can be viewed as independent, it is also possible that they interact, such that an initial reflex-like response to a surprising event is later evaluated as non-threatening, leading to a pleasurable response (Huron, 2007).

Further study of chill responses to scrambled music may be able to pull apart these mechanisms. In this experiment, there was no correlation between the timing of the acoustic events and chills in the scrambled pieces, but this was likely because the scrambling procedure did not preserve these acoustic events. Future studies could use different scrambling procedures that preserve some chill-inducing acoustic events while disrupting musical structure. Chills induced by these acoustic events, in absence of musical structure, may arise through fast bottom-up mechanisms; on the other hand, chills induced by acoustic events only when integrated in a musical structure may arise through top-down mechanisms. Thus, this is a promising avenue for research into the cognitive mechanisms leading to strong emotional responses in music, and the acoustic and structural musical properties that induce them.

4.2. Limitations and future directions

One major difference between the current study and previous studies of chills is the much lower correspondence between reported chills (button presses) and measured chills (presses corresponding to an SCR), which was 46% in the present study, as compared to 73% in Grewe et al. (2007). There may be several reasons for this discrepancy. Firstly, we utilized a much stricter criterion for determining the presence of an SCR in correspondence with a button press. Like Grewe et al. (2007), we only analyzed chills that showed both a button press and an SCR. However, previous studies do not give specific details about the exact methodology used to assess the presence of an SCR and its exact timing. We conformed to recent developments in the analysis of skin conductance by using an automated detection of SCRs, which uses thresholds for latency and amplitude of the SCRs and takes into consideration variable ITIs (Benedek & Kaernbach, 2010). We chose a window of 4 seconds following a button press based on the reports of Grewe et al. (2009) who found the maximum SC increase 2.7 seconds following a button press, and the known latency of SCRs 1 to 3 seconds following a stimulus (Figner & Murphy, 2011). Thus, this different and possibly stricter methodology led to a lower rate of correspondence between self-reported chills and SCRs than has been reported in some previous studies.

A few other factors may have contributed to this low correspondence. It is possible that the instructions given to participants for what constitutes a chill were different, and looser, in this experiment. We told participants (who had all previously reported having experienced chills in the past) to press the button when they felt "a strong emotional and pleasurable response to a particular point in the music, felt as a shiver down the spine or goosebumps" and only further clarified this to participants who asked. A large portion of our participants (27%) were non-native English speakers, so this description might not have been clear or precise enough, and participants might have pressed the button also in response to strong emotions that we might not define as chills. Additionally, participants were told to "keep the button pressed for the duration of the chill" and it was clarified to them that this meant pressing it once and releasing it at the end of the chill. Regardless of this explanation, a few participants still pressed the button repeatedly in a row, possibly having misunderstood the instruction to mean to press the button multiple times until the chill had ended. While these multiple presses were removed from the analyses as noted in the Method section, they may also have contributed to the low correspondence of self-report and SCRs. Lastly, one reason to measure SCRs in the first place was to rule out the possibility that desirability or demand biases would artificially increase the number of self-reported chills and drive the results. It is possible that the particular design of this study did, in fact, cause participants to report more chills than they felt. What is notable, and what the skin conductance helps discern, is that even if this is the case, the number of "correctly reported" chills did not vary across styles and groups. Thus, the skin conductance was successful in ruling out the possibility that the results of our analyses would reflect desirability biases rather than true emotional responses.

Another potential limitation of the current study, which it shares with much of the literature on cross-cultural music cognition, is that all participants were exposed and familiar with Western music. The inclusion of a style that all participants were familiar with provided us with a useful baseline, though it is possible that this shared knowledge may have influenced participants' emotional responses across styles. While the finding that knowledge of Western music was not related to the number of chills suggests that shared knowledge was not relevant, to truly rule out this possibility, future work should be conducted in a sample not previously exposed to Western music. Previous findings of cross-cultural emotional communication in such groups suggest that our results would likely replicate (Fritz et al., 2009). It may also be argued that, due to the prevalence of Western music around the world, its influence may be reflected in the musical features of the traditional Chinese and Hindustani classical pieces we used in this study. While we chose these particular styles due to their prevalence in previous cross-cultural investigations (Balkwill & Thompson, 1999; Castellano et al., 1984; Krumhansl, 1995; Laukka et al., 2013; Rohrmeier & Widdess, 2012) and the availability of participants familiar with these styles, future studies may also address the question of whether our results generalize to other musical styles. As a first cross-cultural study of chills, the present experiment establishes that people can feel strong emotion even to music with very different tonal structure, contrary to previous assumptions (Sloboda, 1991).

We also screened participants for having had chills in the past. This was a necessary limitation, since including participants who have never experienced chills in the past would lead to a large portion of participants not providing us with any usable data. Nonetheless, this may limit the generalizability of our findings to individuals who are prone to experiencing chills. The experience of chills has been linked with personality traits such as high openness to experience (Colver & El-Alayli, 2016; Nusbaum & Silvia, 2010; Silvia & Nusbaum, 2011). Further research is needed to assess whether aspects of personality increased the likelihood that participants would experience chills in music of other cultures.

Lastly, the questionnaire we utilized to measure the knowledge of Western music theory (adapted from Gooding et al., 2014) focused primarily on Western music notation. Future research on the potential link between emotional responses and knowledge of music theory should expand our findings through the use of a more comprehensive test of music theory knowledge, beyond music notation.

An additional step towards our understanding of how chills are generated cross-culturally would be to further analyze the pieces we used as stimuli by looking at musical structure rather than just acoustic properties. This can be achieved through models that take into account the rules governing this structure in Western, Chinese and Hindustani music. While our analyses show that unexpected acoustic events are linked to chills, a structural analysis would explain how these events are incorporated in a sensical way.

4.3. Conclusion

This experiment represents, to the best of our knowledge, the strongest evidence to date suggesting that people can experience felt emotional responses to music of other cultures even with limited explicit knowledge of their musical structure and cultural context. It is also the first cross-cultural exploration of chills, which we show to be an effective measure of felt emotion in music that can be corroborated by physiological data. We thus contribute to the limited but growing interest in approaching music cognition through a cross-cultural perspective, in order to achieve a broader and more fundamental view of the cognitive mechanisms underlying our perception of music.

Given the extreme variability found in music around the world, it has been often argued or assumed that, without knowledge of their cultural context and conventions, what listeners can infer of the meaning of music from other cultures is rather limited (Clayton, 2003; Myers, 1992; Sloboda, 1991). Evidence for the effect of enculturation (Morrison & Demorest, 2009) and the fact that some fundamental properties of Western music, such as the preference for consonance, are not always observed cross-culturally (McDermott et al., 2016) supports this view. On the other hand, there is a long tradition of attempts to reduce music of different cultures to a set of basic characteristics (e.g., Fitch, 2006; Harwood, 1976; Lomax, 1968; Savage et al., 2015). Research has shown that listeners can infer the mood (Balkwill & Thompson, 1999; Laukka et al., 2013; Fritz et al., 2009) and form tonal and melodic expectations in response to music of other cultures (Castellano et al. 1984; Krumhansl, 1995, 2000; Rohrmeier & Widdess, 2012), suggesting that part of musical meaning may rely on cross-cultural factors. However, there has not been strong evidence for the idea that listeners also experience felt emotional responses to music of other cultures.

We show that music is capable of generating strong emotional responses through intrinsic features that are present and responded to cross-culturally, independently of cultural context. Some of these intrinsic features can be captured through acoustic properties such as loudness, roughness and brightness. Sudden changes in these properties are related to chills across styles. Thus, we are in support of the view that listeners can rely on cross-cultural musical elements to extract the emotional meaning of musical styles from other cultures, resulting in strong felt emotional responses. Nonetheless, it is important to remember that music may express meaning at many levels, of which emotional meaning may just be one aspect (Cross et al., 2009). The responses measured in this study are just one of the several ways in which people react to music. Thus, this experiment does not deny the importance of cultural context and enculturation in the communication of musical meaning.

5. Supplementary material

Data and scripts for this project are available through the Open Science Framework at: https://osf.io/25h6x/. Stimuli cannot be provided due to copyright restrictions.

Table S1
Information about the musical instruments used in the traditional Chinese and Hindustani classical pieces.

Instrument	Description
Pipa	The pipa is a plucked instrument consisting of four strings and a pear-shaped wooden body. While it is still one of the most popular Chinese instruments, its origins date back to the Han Dynasty (2nd century BC) and it may have been transmitted from central Asia.
Erhu	The erhu is a two-string bowed instrument. The body consists of a long neck and a small hexagonal resonator box. Its origins date back to the Tang Dynasty (7th to 10th century AD) and it may have been transmitted from central Asia.
Yangqin	The yangqin is a hammered dulcimer consisting of a trapezoidal wooden body and 7 to 18 sets of strings over 4 to 5 bridges. It may have been modelled on the Persian santur, introduced to China in the 17th century AD from Persia.
Bianzhong	The bianzhong is a set of bronze chime bells, hung to a wooden frame and struck with a mallet. Its origins date back to the Zhou Dynasty (11th to 2nd century BC).
Sarangi	The sarangi is a North Indian short-necked bowed instrument with a rectangular body carved from a single piece of wood. Its sound is thought to best resemble the human voice. While it peaked in popularity in the 17th century AD as an accompaniment to vocal music, its popularity has since declined in favor of the harmonium and violin.
Shehnai	The shehnai is a North Indian double-reed woodwind instrument consisting of a wooden tubular body ending in a metal flared bell. While its origins are debated, it may have evolved from the Persian ney or from the South Indian nadaswaram.
Santoor	The santoor is a hammered dulcimer consisting of a trapezoidal wooden body. The number of strings and bridges is not standardized as the instrument was only recently introduced to Hindustani classical music. It

originated in Kashmir as part of the Sufiana Mausiqi tradition, and may have been adapted from the Persian santur.

6. Acknowledgments

We would like to thank Heera Kulkarni, Yun Fan, Paul Beier and Mariagrazia Carlone, the musical experts who guided us in the musical selections and in creating the music knowledge survey; David Roby, who provided insights and resources from the field of ethnomusicology; Emilio Ferrer, Shelley Blozis, Cassandra Jacobs and Gwendolyn Rehrig for advice on the statistical analyses; the members of the Janata Lab and Ferreira Lab for useful comments and discussion; the two anonymous reviewers for their detailed comments and suggestions; and Brigitte Tabaroki and Sarah Huynh, for their indispensable help with data collection.

We acknowledge support from the National Science Foundation GRFP number 1650042 awarded to EB.

7. Declaration of interest

Declarations of interest: none.

Appendix

Musical knowledge survey. Correct answers are bolded.

Western music questions

1. Which one of these operas is not by Verdi?
    a. Aida
    b. Nabucco
    c. Traviata
    d. **Madame Butterfly**
    e. I don’t know

2. Gregorian chant is typical of:
    a. ancient Rome
    b. the 1920s
    c. the 1700s
    d. **the Middle Ages**
    e. I don’t know

3. “Modulation” means:
    a. singing from memory
    b. **passing from one key to another in the same piece**
    c. tuning a wind instrument
    d. the combination of notes to produce a chord
    e. I don’t know

4. Which composer is most famous for his fugues?
    a. Mozart
    b. Vivaldi
    c. **Bach**
    d. Beethoven
    e. I don’t know

5. A string quartet is formed by:
    a. **2 violins, 1 viola and 1 cello**
    b. 4 violins
    c. 1 violin, 1 viola, 1 cello and 1 bass
    d. 1 violin, 1 viola, 2 cellos
    e. I don’t know

6. Polyphony is:
    a. music for several wind instruments
    b. vocal music
    c. music for multiple voices
    d. orchestral music
    e. I don’t know

7. The Pathetique Sonata is by:
    a. Haydn
    b. **Beethoven**
    c. Chopin
    d. Brahms
    e. I don’t know

8. The lute has:
    a. a mouth piece
    b. **strings**
    c. pedals
    d. keys
    e. I don’t know

9. The oboe has:
    a. **reeds**
    b. piston valves
    c. strings
    d. knobs
    e. I don’t know

Chinese music questions

1. Which one of the Chinese instruments below has two strings?
    a. pipa (琵琶)
    b. sanxian (三弦)
    c. **erhu (二胡)**
    d. guzheng (古筝)
    e. I don’t know

2. Cantonese music is a form of

   a. **instrumental music**
   b. folk song
   c. Chinese opera
   d. religious music
   e. I don't know

3. Pipa is a ______ instrument
   a. percussion
   b. wind
   c. **string**
   d. keyboard
   e. I don't know

4. Beijing opera (Peking opera) was formed in
   a. **19th century**
   b. 20th century
   c. 18th century
   d. 17th century
   e. I don't know

5. Mouth organ refers to which instrument?
   a. guqin (古琴)
   b. **sheng (笙)**
   c. ruan (阮)
   d. luo (锣)
   e. I don't know

6. What best describes Molihua (Jasmine flower):
   a. ensemble piece
   b. opera music
   c. instrumental music
   d. **folk song**
   e. I don't know

7. How many strings does the guqin (古琴) have?
   a. 5
   b. 3
   c. **7**
   d. 12

   e. I don’t know

8. A famous Beijing opera actor toured multiple cities the US in 1930. His name is:
   a. **Mei Lanfang**
   b. Cheng Yanqiu
   c. Shang Xiaoyun
   d. Xun Huisheng
   e. I don’t know

9. Which of the following is Abing's real name:
   a. Liu, Tianhua
   b. Xian, Xinghai
   c. **Hua, Yanjun**
   d. Yang, Yinliu
   e. I don’t know

Hindustani music questions

1. Of the following which one is a percussion instrument used in Indian music?
   a. Shehnāi
   b. Sarod
   c. **Pakhāwaj**
   d. Sārangi
   e. I don’t know

2. Sitar is a stringed instrument. How many “Tarab" or sympathetic strings does a Sitar have?
   a. 21
   b. **13**
   c. 7
   d. 32
   e. I don’t know

3. Who among the following is an accomplished vocalist?
   a. Ronu Majumdār
   b. Shiv Kumar Sharmā
   c. **Kishori Āmonkar**
   d. Nikhil Banerjee
   e. I don’t know

4. The oldest form of composition of Hindustani Vocal Music is:
    a. Ghazal
    b. **Dhrupad**
    c. Thumri
    d. Qawwāli
    e. I don't know

5. Allā Rakhā was an exponent of which of the following instruments?
    a. Shehnai
    b. Santoor
    c. Sitar
    d. **Tabla**
    e. I don't know

6. The theory of Indian classical music is discussed in which of the following Veda?
    a. Rigveda
    b. Yajurveda
    c. **Sāmaveda**
    d. Atharvaveda
    e. I don't know

7. What material is the Indian flute made of?
    a. Steel
    b. **Wood**
    c. Copper
    d. Plastic
    e. I don't know

8. How many mātrās or beats does Teen Tāl have?
    a. **16**
    b. 12
    c. 10
    d. 14
    e. I don't know

9. Who is credited with creating the world famous Rāg: Miyān Malhar?
    a. **Tānsen**
    b. Amir Khusro
    c. Meerā Bāi

d. Bhimsen Joshi
e. I don’t know

Western music theory questions (Adapted from Gooding et al., 2014)

1. There are/is________ stave(s) present in this musical example.
    a. 1
    b. **2**
    c. 4
    d. None
    e. I don’t know

2.  indicates that the piece should be played:
    a. **loudly**
    b. slowly
    c. fast
    d. softly
    e. I don’t know

3. Beat one in measure one is a:
    a. quarter note
    b. quarter rest
    c. eighth note
    d. **eighth rest**
    e. I don’t know

4. What pitch does this excerpt start on?
    a. C
    b. D
    c. A
    d. **G**
    e. I don’t know

5. The fermata occurs on what pitch or note?
    a. **E flat**
    b. D
    c. F
    d. A sharp
    e. I don’t know

6. How many flats are in this key signature?
    a. 1
    b. 2
    c. **3**
    d. 0
    e. I don’t know

7. “Allegro con brio” is a_______ marking.
    a. key
    b. **tempo**
    c. dynamic
    d. accidental
    e. I don’t know

References

Abrams, D., Ryali, S., Chen, T., Chordia, P., Khouzam, A., Levitin, D., & Menon, V. (2013). Inter-subject synchronization of brain responses during natural music listening. *European Journal of Neuroscience, 37(9)*, 1458-1469.

Adger, D. (2019). *Language Unlimited: The Science Behind Our Most Creative Power.* New York, NY: Oxford University Press.

Balkwill, L., & Thompson, W. (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.

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10*, 20-46.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Becker, J. (1986). Is Western art music superior? *The Musical Quarterly*, *72(3)*, 341-359.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. http://doi.org/10.1016/j.jneumeth.2010.04.028

Blacking, J. (1995). *Music, culture, and experience: Selected papers of John Blacking*. University of Chicago Press.

Blood, A., & Zatorre, R. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America, 98(20)*, 11818-23.

Bloom, A. (1987). *Closing of the American mind*. New York, NY: Simon and Schuster.

Brown, D. E. (2004). Human universals, human nature & human culture. *Daedalus*, *133*(4), 47-54.

Brown, D. E. (1991). *Human universals.* Philadelphia, PA: Temple University Press.

Bürkner, P. C. (2017a). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1-28.

Bürkner, P. C. (2017b). Advanced Bayesian multilevel modeling with the R package brms. *arXiv preprint arXiv:1705.11123*.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of north India. *Journal of Experimental Psychology. General*, *113*(3), 394–412. https://doi.org/10.1037/0096-3445.113.3.394

Clayton, M. (2003). *The cultural study of music: A critical introduction*. Routledge.

Colver, M. C., & El-Alayli, A. (2016). Getting aesthetic chills from music: The connection between openness to experience and frisson. *Psychology of Music*, *44*(3), 413–427. https://doi.org/10.1177/0305735615572358

Conard, N. J., Malina, M., & Münzel, S. C. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature*, *460*(7256), 737.

Craig, D. (2005). An exploratory study of physiological changes during "chills" induced by music. *Musicae Scientiae, 9(2)*, 273-287.

Cross, I., Woodruff, G. E., Road, W., & Botha, I. R. (2009). Music as a communicative medium. In R. Botha & C. Knight (Eds.), *The prehistory of language* (Vol. 1, pp. 113-144). Oxford: Oxford University Press.

Ellis, D. P. W. (2010). Time-domain scrambling of audio signals in Matlab [Software]. Available from: http://www.ee.columbia.edu/~dpwe/resources/matlab/scramble/

Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, *12*(1), 75-99.

Fedorenko, E., McDermott, J., Norman-Haignere, S., & Kanwisher, N. (2012). Sensitivity to musical structure in the human brain. *Journal of Neurophysiology*, *108*, 3289-3300.

Figner, B., & Murphy, R. O. (2011). Using skin conductance in judgment and decision making research. In M. Schulte-Mecklenbeck, A. Kühberger & J. G. Johnson, J. (Eds.). *A handbook of process tracing methods for decision research: A critical review and user's guide* (pp. 163 - 184). New York, NY: Psychology Press.

Fitch W. T. (2006) The biology and evolution of music: A comparative perspective. *Cognition* *100*(1):173–215.

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., ... & Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, *27*(2), 233-249.

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., … Koelsch, S. (2009). Universal Recognition of Three Basic Emotions in Music. *Current Biology*, *19*(7), 573–576. https://doi.org/10.1016/j.cub.2009.02.058

Goldberg, A. E. (2019). *Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions*. Princeton University Press.

Gooding, L., Abner, E., Jicha, G., Kryscio, R., Schmitt, F. (2014). Musical training and late-life cognition. *American Journal of Alzheimer's Disease and Other Dementias, 29(4),* 333-343.

Grewe, O., Nagel, F., & Kopiez, R. (2007). Listening to music as a re-creative process: Physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Perception: An Interdisciplinary Journal, 24(3),* 297-314.

Grewe, O., & Kopiez, R. (2009). The chill parameter: Goosebumps and shivers as promising measures in emotion research. *Music Perception: An Interdisciplinary Journal, 27(1),* 61-74.

Guhn, M., Hamm, A., & Zentner, M. (2007). Physiological and musico-acoustic correlates of the chill response. *Music Perception*, *24(5)*, 473–484.

Harrison, L., & Loui, P. (2014). Thrills, chills, frissons, and skin orgasms: Toward an integrative model of transcendent psychophysiological experiences in music. *Frontiers in Psychology, 5,* 1-6.

Harwood, D. L. (1976). Universals in music: A perspective from cognitive psychology. *Ethnomusicology*, *20*(3), 521–533.

Higham, T., Basell, L., Jacobi, R., Wood, R., Ramsey, C. B., & Conard, N. J. (2012). Testing models for the beginnings of the Aurignacian and the advent of figurative art and music: The radiocarbon chronology of Geißenklösterle. *Journal of human evolution*, *62*(6), 664-676.

Huron, D., (2006). *Sweet anticipation: Music and the psychology of expectation*. MIT press.

Huron, D. (1992). The ramp archetype and the maintenance of passive auditory attention. *Music Perception: An Interdisciplinary Journal*, *10*(1), 83–91.

Janata, P., Tillmann, B., & Bharucha, J. (2002). Listening to polyphonic music recruits domain-general attention and working memory circuits. *Cognitive, Affective, & Behavioral Neuroscience, 2(2),* 121-140.

Jiang, F., Stecker, G. C., & Fine, I. (2013). Functional localization of the auditory thalamus in individual human subjects. *Neuroimage, 78,* 295-304.

Juslin, P., & 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., & Västfjäll, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *The Behavioral and Brain Sciences*, *31*(5), 559-575; discussion 575-621. https://doi.org/10.1017/S0140525X08006079

Juslin, P. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266. https://doi.org/10.1016/j.plrev.2013.05.008

Krumhansl, C. L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum ,17,* 53-80

Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: cognitive methodology applied to North Sami yoiks. *Cognition*, *76*(1), 13–58. https://doi.org/10.1016/S0010-0277(00)00068-8

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13).

Lalitte, P., & Bigand, E. (2006). Music in the moment? Revisiting the effect of large scale structures. *Perceptual And Motor Skills, 103*(3), 811-828.

Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A matlab toolbox for music information retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker (Eds.), *Data analysis, machine learning and applications* (pp. 261-268). Springer, Berlin, Heidelberg.

Laukka, P., Thingujam, N., Eerola, T., Tamasaki, T., & Beller, G. (2013). Universal and culture-specific factors in the recognition and performance of musical affect expressions. *Emotion, 11*, 434-449.

Lenth, R. V. (2017). Using lsmeans. *Journal of Statistical Software*, *69*, 1-33.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Levitin, D., & Menon V. (2005). The neural locus of temporal structure and expectancies in music: Evidence from functional neuroimaging at 3 tesla. *Music Perception*, *22(3),* 563-575.

Lomax A., ed (1968) Folk Song Style and Culture (American Association for the Advancement of Science, Washington, DC).

Mauss, I., Wilhelm, F., & Gross, J. (2004). Is there less to social anxiety than meets the eye? Emotion experience, expression, and bodily responding. *Cognition & Emotion*, *18*(5), 631–642. https://doi.org/10.1080/02699930341000112

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, *5*(2), 175–190. https://doi.org/10.1037/1528-3542.5.2.175

Mcdermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535*, 547–550. https://doi.org/10.1038/nature18635

McFadyen, M. (2014). Together in sang: The embodied song experience as singularly plural. In Russell, Ian & Catherine Ingram (eds.) *Taking Part in Music: Case Studies in Ethnomusicology*. Aberdeen: Aberdeen University Press

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., ... & Howard, R. M. (2019). Universality and diversity in human song. *Science*, *366*(6468).

Menon, V., & Levitin, D. (2005). The rewards of music listening: Response and physiological connectivity of the mesolimbic system. *Neuroimage, 28*, 175-184.

Meyer, L.B. (1967). *Music, the arts, and ideas.* Chicago: University of Chicago Press.

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E. J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic bulletin & review*, *23*(1), 103-123.

Morgan, E., Fogel, A., Nair, A., & Patel, A. D. (2019). Statistical learning and Gestalt-like principles predict melodic expectations. *Cognition*, *189*, 23–34. https://doi.org/10.1016/j.cognition.2018.12.015

Morrison, S. J., & Demorest, S. M. (2009). *Cultural constraints on music perception and cognition*. *Progress in Brain Research* (Vol. 178). Elsevier. https://doi.org/10.1016/S0079-6123(09)17805-6

Myers, H. (Ed.). (1992). *Ethnomusicology, an introduction*. WW Norton.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model.* Chicago: University of Chicago Press.

Nettl, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The Origins of Music* (pp. 463-472). Cambridge, MA: MIT Press.

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational Ideas—Part II. *Language and Linguistics Compass*, *10*(11), 591-613.

Nusbaum, E. C., Silvia, P. J., Beaty, R. E., Burgin, C. J., Hodges, D. a., & Kwapil, T. R. (2014). Listening between the notes: Aesthetic chills in everyday music listening. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(1), 104–109. https://doi.org/10.1037/a0034867

Nusbaum, E. C., & Silvia, P. J. (2010). Shivers and timbres: Personality and the experience of chills from music. *Social Psychological and Personality Science*, *2*(2), 199–204. https://doi.org/10.1177/1948550610386810

Oram, N., & Cuddy, L. L. (1995). Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research*, 103–118.

Patel, A. (2008). *Music, Language and the Brain.* Oxford: Oxford UP.

Panksepp, J. (1995). The emotional sources of "chills" induced by music. *Music Perception: An Interdisciplinary Journal, 13(2),* 171-207.

Penman, J., & Becker, J. (2009). Religious ecstatics, "deep listeners," and musical emotion. *Empirical Musicology Review*, *4*(2), 49–70. https://doi.org/10.18061/1811/37474

Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: *Infant Behavior and Development*, *21*(1), 77–88. https://doi.org/10.1016/S0163-6383(98)90055-8

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rohrmeier, M., & Widdess, R. (2012). Incidental learning of modal features of North Indian music. *International Conference on Music Perception and Cognition*, 857–866. https://doi.org/10.1111/cogs.12404

Salimpoor, V., Benovoy, M., Longo, G., Cooperstock, J., & Zatorre, R. (2009) The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS ONE, 4(10),* e7487.

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), 257–62. https://doi.org/10.1038/nn.2726

Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., and Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science, 340*, 216–219. doi: 10.1126/science.1231059

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, *112*(29), 8987–8992. https://doi.org/10.1073/pnas.1414495112

Silvia, P. J., & Nusbaum, E. C. (2011). On personality and piloerection: Individual differences in aesthetic chills and other unusual aesthetic experiences. *Psychology of Aesthetics, Creativity, and the Arts*, *5*(3), 208–214. https://doi.org/10.1037/a0021914

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

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of cognitive neuroscience*, *18*(8), 1380-1393.

Turner E. (2012) Music and sport: Being in the zone. In: *Communitas: The anthropology of collective joy (pp. 43-54)* New York, NY: Palgrave Macmillan.

Vehtari A, Gelman A, Gabry J (2016). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. R package version 1.0.0, URL https://github.com/ stan-dev/loo.

Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. doi:10.1007/s11222-016-9696-4.

Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5,* 1–14. https://doi.org/10.3389/fpsyg.2014.01111.

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, 9. https://doi.org/10.1371/journal.pone.0094446