CHAPTER 3

CULTURAL DISTANCE: A COMPUTATIONAL APPROACH TO EXPLORING CULTURAL INFLUENCES ON MUSIC COGNITION

STEVEN J. MORRISON, STEVEN M. DEMOREST, AND MARCUS T. PEARCE

Introduction

As with many psychological constructs, much of what has been reported in research on the cognitive processing of music is limited to data collected from individuals from a small subset of cultural contexts (Henrich, Heine, & Norenzayan, 2010). Further, the music that is typically employed for the purposes of testing and exploration tends to be drawn from a similarly small set of music practices and mostly consists of that constructed within the Western diatonic framework. This includes Western classical music as well as many North American and Western European folk and popular genres. This is striking given that music is often regarded as a particularly prominent and powerful manifestation of culture. Music is a common way for individuals to assert cultural identity (Frith, 1996) and, as such, its value arguably lies as much in its cultural and stylistic distinctiveness as in any universal qualities it may possess.

Musical systems are somewhat closed in that each describes a set of practices and conventions within which performances, pieces, or whatever might be the appropriate musical "unit" are understood and evaluated. These same practices and conventions can also serve as touchstones against which individuals push in the spirit of creativity and innovation. People come to inhabit a musical system due to various combinations of formal learning—conservatory training, for example, as a means of gaining knowledge of avant-garde art music—and informal learning—becoming steeped in Cajun music as a result of growing up in the southern region of the US state of Louisiana. In this chapter, our purpose is to emphasize music as an *inter*cultural phenomenon. As such, we will not focus on the particularities of any specific music cultural tradition, nor will we examine the concept of musical universals or the structural or acoustical candidates for such a distinction. Rather, we will dedicate our attention to interactions between music cultures, to what happens when music moves across cultural boundaries.

From a sociological perspective, it has been useful to view the construct of culture from a somewhat dichotomous perspective in which the notion of the cultural insider can be contrasted with that of the cultural outsider (Merton, 1972). Contemporary scholarship has drawn attention to the complexity of this comparison and the considerable subjectivity that lies at the heart of such an often, oversimplified bifurcation (for an examination of this issue in the field of music research, see Trulsson & Burnard, 2016). Although music is often associated with cultural identity and therefore susceptible to insider/outsider categorization, the ease with which an individual interacts with any given culture's music may be more nuanced. Culture-based differences in the way listeners and performers interact with and respond to music are often delineated by ethnic identity or geographical location which are, in turn, generally treated as categorical constructs. As such, they tend to oversimplify complex relationships, obscure considerable within-group variability, and, most critically for the present purpose, do not hold up well when considering a brain-based understanding of music processing.

The cultural dimension of music provides context for critical tests of music as a neurological phenomenon. The conclusion that particular brain regions or neurological pathways are associated with *human* music processing can be tested by examining whether such relationships are evident across musical and cultural contexts. Likewise, the strength or extent of neural activity may offer insights into the ways in which particular music parameters function within musical systems.

Cultural roots of music practice also offer a critical test of principles of formal musical learning. Teaching and learning practices often vary from culture to culture and, given that they are often directed toward within-culture music, likely interact with the idiosyncratic elements of the music being taught. The prospect of learning—even at a fundamental level—an unfamiliar music tradition as a performer or as a listener provides a context in which culture-general learning strategies or pathways might be tested. Similarly, it provides a framework in which "from the ground up" skill or schema development can be observed, particularly through more informal learning pathways in which exposure and self-directed discovery feature prominently. At the neurological level, learning within a culturally unfamiliar context might provide further evidence of experience-based neural plasticity as well as potential interactions with already-learned music conventions.

Given the incremental nature of music learning (formal or informal) and the imprecision of insider/outsider classifications, cross-cultural studies of both music perception and music learning would benefit from a more nuanced view of cross-cultural differences in musical traditions, one that is more continuous than categorical. Below we will explore the construct of *cultural distance* as one potential approach. Cultural distance has been examined at a societal level (Hofstede, 1983) through the development of a suite of measures found to effectively account for culture-based variability among workers. Since its publication, this construct has been used primarily in the fields of business and economics; however, it has also been employed in a number of cross-cultural designs including, occasionally, those related to music (Baek, 2015). The principle of cultural distance—as a way to conceptualize a culture-specific phenomenon in relation to its manifestation in other cultures—is evident in research on more specific cultural practices, as well. Kuhl (1991), for example, posited a "perceptual magnet effect" to explain early language learning processes and the manner in which infants' speech perception quickly gravitates toward commonly used phonemic prototypes. Similarly, individuals demonstrate better memory (Golby, Gabrieli, Chiao, & Eberhardt, 2001) as well as better recognition of emotional expression (Chiao et al., 2008) for same-race faces. In both instances, more differentiated face recognition correlated with increased neural activity in fusiform areas and amygdala, respectively.

In this chapter, we will provide a brief overview of cross-cultural research in music cognition. We will consider studies that have compared individuals' interactions with culturally familiar and unfamiliar music, those that have compared responses by participants from different cultural backgrounds, and those that have employed fully comparative designs in which participants of different cultural backgrounds interact with each other's music tradition. Among the previous research, we will summarize some of our own recent work that has focused on identifying musical parameters—specifically pitch and rhythm—that appear to make a particularly strong contribution to the differences arising from cross-cultural music interactions. Based on this work, we will then describe the construct of *cultural distance* as a conceptual and analytical means of interpreting and perhaps predicting cross-cultural responses to music.

Related Literature

The purpose of this review is to provide a brief overview of topics in music cognition that have been explored through a cross-cultural lens. (For more thorough treatment of this topic see reviews by Morrison & Demorest, 2009 and Patel & Demorest, 2013.) Researchers have explored the cross-cultural perception of music emotion, preference, musical structures of scale and key, rhythm and meter, and larger formal elements, as well as musical memory. Participants in these studies have spanned the gamut from infancy to adulthood offering a picture of how culture influences music cognition and how that influence changes with age and experience.

Cross-Cultural Explorations of Emotion

The single largest body of cross-cultural research in music cognition has to do with the recognition of emotion in music. With the exception of a very small number of studies (e.g., Egerman, Fernando, Chuen, & McAdams, 2015), the research has focused not on emotion induction, or how music makes you feel, but on the ability to recognize emotional states present in music stimuli. On the surface, this seems a curious choice given the somewhat flexible nature of emotion recognition even within a cultural group. However, emotion proves to be an excellent choice for exploring cultural universality versus particularity in music cognition because emotions refer not just to cognitive categories, but to physical states that can be mimicked acoustically (Juslin, 2000; Juslin & Laukka, 2003). Cross-cultural studies have explored Western listeners' perceived emotion in music of India (Balkwill, 2006; Balkwill & Thompson, 1999; Balkwill, Thompson, & Matsunaga, 2004; Deva & Vermani, 1975; Gregory & Varney, 1996; Keil & Keil, 1966), perception of Western music by non-Western listeners, including Congolese pygmies (Egermann et al., 2015) and the Mafa people of northern Cameroon (Fritz et al., 2009), Western listeners perception of Congolese pygmy music (Egermann et al., 2015), and the cross-cultural communication of emotion involving performers and listeners from Swedish, Indian, and Japanese music cultures (Laukka, Eerola, Thingujam, Yamasaki, & Beller, 2013).

The findings can be summarized briefly as follows: A limited set of emotions can be recognized in music regardless of cultural familiarity. The emotions most consistently recognized (happy, sad, angry) vary in arousal in ways that mimic physiological states. Other emotion recognition judgments show influences of cultural familiarity. There are several theories of emotion recognition that attempt to model this combination of psychophysical and cultural cues in emotion recognition judgments. One of the first theories was the Cue Redundancy Model (CRM) proposed by Balkwill and Thompson (1999). According to this model, emotions in music are decoded by attending to cues in the musical stimulus consisting of psychophysical cues (sound intensity, tempo, melodic complexity, pitch range, etc.) and culture-specific cues like the use of a certain instrument or tonality to communicate a particular emotional state. This allows in-culture listeners to use more information in their emotion recognition judgments, but it also allows out-of-culture listeners to access basic emotional information regardless of familiarity. The authors later proposed a more refined model called Fractionating Emotional Systems or FES (Thompson & Balkwill, 2010). FES attempts to explain how the culture-specific and culture-general cues proposed in CRM function in development. They propose that all emotion communication is built on a phylogenetic base of shared cues involved in being human. As we age we incorporate ontogenetic cues for both music and language prosody into our emotional vocabulary in a more culturally specific way. Fritz (2013) has proposed a "dock-in" model of emotion recognition that is consistent with previous models in stating "all music cultures contain both universal and culture-specific features" (p. 514). It differs from previous models in that it proposes that different cultures may "dock in" to only a subset of universal music codes and that cross-cultural understanding can be explained in part by the overlap in universal features employed. This notion of overlap between cultures is similar to the cultural distance hypothesis discussed below, though the basis for comparing cultural systems is based on a simulation of the cognitive processing of musical structure rather than a comparison of stimulus features.

When evaluating the findings of cross-cultural research in emotion perception it is important to keep in mind that, of all of the studies listed, only three (Egermann et al., 2015; Gregory & Varney, 1996; Laukka et al., 2013) were fully comparative, that is, featuring both listeners and musical stimuli from all cultures involved (Patel & Demorest, 2013). It may be difficult to generalize these findings to other non-Western listeners or musics. While the experience of emotions is a human universal, the notion that music contains an emotional message rather than a functional or social one, may be a somewhat culturally specific one. Given that most of the studies cited here asked listeners from Western or Western-influenced cultures to identify the emotions in non-Western music, and that much of that music came from a single non-Western culture (India), it is difficult to determine the cultural appropriateness of emotion judgments in music. As Fritz (2013) observed in relation to one specific comparison involving members of a society indigenous to a remote region of Cameroon, "the musical expression of a variety of emotions like fearfulness and sadness, while recognized in the Western stimuli by the Mafa participants, are—according to interviews with Mafa individuals—never represented in the traditional music of the Mafa people" (p. 512).

Cross-Cultural Explorations of Music Preference

Music preference research also explores affective responses to music, not in terms of how music codes affect and emotion, but rather by examining the conditions under which listeners experience pleasure when hearing music. As LeBlanc proposes in his theoretical model, "Music preference decisions are based upon the interaction of input information and the characteristics of the listener, with input information consisting of the musical stimulus and the listener's cultural environment" (1982, p. 29). Music educators have long been interested in music preference as a cross-cultural phenomenon in part due to their commitment to providing a culturally diverse music education. Researchers in music education have looked at how children's preference for music of other cultures develops and its relationship to familiarity and other musical features.

Researchers have explored the musical qualities that might influence preference judgments across cultures (Demorest & Schultz, 2004; Flowers, 1980; Fung, 1994; Morrison & Yeh, 1999; Shehan, 1981) and whether instruction in a culture's music can influence preference (Heingartner & Hall, 1974; Shehan, 1985). As with the research on emotion, the bulk of studies explore how Western listeners respond to non-Western music and are not fully comparative. Findings show that preference for culturally unfamiliar music can be increased with exposure—most of these studies were conducted in formal educational settings among school-age and college populations—but it does not extend to novel pieces from the culture. Also, students prefer music that has properties of their culture such as westernized arrangements of non-Western music (Demorest & Schultz, 2004). To summarize the findings, the more familiar sounding something is culturally, the more likely listeners are to like it. However, while exposure can increase preference for out-of-culture music, it does so only for learned pieces and does not generalize to the style as a whole (Shehan, 1985).

Cross-Cultural Explorations of Musical Structure

One of the debates surrounding music and culture is the extent to which there are deep structures in music that are relatively invariant across cultures (cf. Brown & Jordania, 2013). Given humans' shared biology and the apparent human need to engage in musical behavior, it is plausible that certain structural features would be present in most, if not all, musics. Through cross-cultural explorations of musical structure, researchers have sought to identify some of the structural features or perceptual processes that work across cultures as well as the points at which music cognition becomes more culturally bound.

Scale and Key Perception

Some of the earliest cross-cultural work done on scale perception included infants (Lynch & Eilers, 1991, 1992; Lynch, Eilers, Oller, & Urbano 1990; Lynch, Eilers, Oller, Urbano, & Wilson, 1991; Lynch, Short, & Chua, 1995). In a series of studies the authors tested whether pitch deviations could be detected when presented in the context of familiar (major/minor) versus unfamiliar (pelog) scale contexts. They found that deviations were better detected for familiar scale contexts for both adults and children with the exception of infants aged 6–12 months who performed similarly. While these studies represent an important early attempt to examine scale perception, they were hampered by methodological issues pertaining to the way in which stimuli were created and the possible interference of absolute pitch strategies.

There has been a significant amount of work examining whether tonal relationships or tonal hierarchies (Krumhansl & Shepard, 1979) can be perceived by out-of-culture listeners (Castellano, Bharucha, and Krumhansl, 1984; Kessler, Hansen, and Shepard, 1984; Krumhansl, 1995; Krumhansl, Louhivuori, Toiviainen, Jarvinen, & Eerola, 1999; Krumhansl et al., 2000). The research has included music and participants from a variety of cultures in the designs and the findings have been mixed. The general sense is that out-of-culture listeners can employ more global strategies involving tone proximity and frequency of occurrence within the stimulus materials to mimic insider tonality judgments, but only up to a point. When judgments become more complex (Krumhansl et al.,1999, 2000) or require specific cultural knowledge (Curtis & Bharucha, 2009), cultural influences on tonal cognition become more pronounced. This suggests that tonality perception, like emotion perception, provides both general and specific cues for listeners depending on their cultural background.

Two recent fully-comparative studies (Raman & Dowling, 2016, 2017) demonstrate the relative influence of global versus cultural factors in tonality judgments. In a series of four experiments across two studies the authors explored the sensitivity of Western and Carnātic trained musicians to two types of modulations in Carnātic melodies. The *rāgamālikā* modulation is more typical in Carnātic music and corresponds to the less frequent parallel minor (C major to C minor) modulation in Western music. The *grahabēdham* modulation is less common in Carnātic music, but more common in Western music as it corresponds to a modulation to the relative minor (C major to A minor). They tested modulation identification (both accuracy and speed), tonal profiles, and active probe tone response during modulation. While results varied somewhat across the different experiments, they found, in general, that cultural background influenced speed and accuracy in modulation detection with Indian listeners more accurate overall. Response time varied by the cultural familiarity of the modulation, with Indians faster for *rāgamālikās* and Westerners faster for *grahabēdhams.* They also found that Western musicians' tone profile responses, while relying on global information about frequency and distribution of tones, were sometimes influenced by a misapplication of Western major/minor judgments in Carnātic tone profiles. The authors reference the Cue Redundancy Model reviewed above as a possible explanation for the mix of global and cultural cues employed by both groups of musicians.

Other approaches to cross-cultural tonal cognition have included event-related potential (ERP) responses to tasks involving out-of-culture scale violations (Neuhaus, 2003; Renninger, Wilson, & Donchin, 2006) and melodic expectancy violations (Demorest & Osterhout, 2012). In general, listeners were less sensitive to out-of-culture scale deviations unless they could detect the deviations using a culture-specific strategy. Another area of research has addressed whether linguistic background shapes musical ability. Researchers have found that tonal language speakers are generally better at general pitch discrimination (Giuliano, Pfordresher, Stanley, Narayana, & Wicha, 2011; Pfordresher & Brown, 2009; Wong et al., 2012) and even at pitch accuracy in singing (Pfordresher & Brown, 2009) than those from non-tonal linguistic backgrounds. The authors suggest that fine-grained pitch processing is central to the acquisition of a tonal language and therefore better developed among these individuals (Pfordresher & Brown, 2009).

Rhythm and Meter Perception

Rhythm and meter perception has received much more attention in music cognition over the last ten to fifteen years, and with that attention has come a commensurate increase in cross-cultural exploration. Researchers have examined when infants' responses to meter become culturally biased (Hannon & Trehub, 2005a, 2005b; Soley & Hannon, 2010), the influence of linguistic rhythm on rhythm perception (Hannon, 2009; Iversen, Patel, & Ohgushi, 2008; Patel & Daniele, 2003; Yoshida et al., 2010), and cultural influences on rhythmic perception and performance (Cameron, Bentley, & Grahn, 2015; Drake & Ben El Heni, 2003; Polak, London, & Jacoby, 2016; Stobart & Cross, 2000).

In all of these investigations researchers have found varying degrees of cultural influence in rhythm processing in adults and infants, with infants demonstrating a preference for the meters of their home culture as early as 4–8 months (Soley & Hannon, 2010), even when those meters were more complex. Unlike adults, monocultural infants were equally responsive to metric violations within both familiar and unfamiliar meters (Hannon & Trehub 2005a) and infants as old as 12 months demonstrated enough flexibility to "reset" their perceptual responses with sufficient exposure to an unfamiliar meter (Hannon & Trehub, 2005b). While language acquisition has often been a focus of tonal cognition, several studies have found relationships between the rhythmic qualities of language and musical rhythms (Hannon, 2009; Patel & Daniele, 2003) and rhythm grouping (Iversen et al., 2008; Yoshida et al., 2010) of instrumental music from the culture.

In a recent fully comparative study, Cameron and colleagues (2015) tested Western-born and East African musicians' performance on three rhythmic tasks, discriminating between two patterns, reproducing rhythm patterns, and tapping a steady beat to rhythmic patterns. Patterns were drawn from East African and Western music and the authors predicted that musicians would show a cultural advantage for all three tasks. As with previous cross-cultural work, however, they found that while the two performance tasks (rhythm reproduction and beat tapping) showed an in-culture advantage, the groups were equally adept at rhythm discrimination. This study was particularly noteworthy for including both perception and performance measures, as many studies feature one or the other.

Phrasing and Form

Researchers have explored the influence of enculturation on phrase boundary perception (Nan, Knösche, & Friederici, 2006; Nan, Knösche, Zysset, & Friederici, 2008) and musical tension (Wong, Chan, Roy, & Margulis, 2011) through neuroscientific measures. Two fully comparative ERP studies (Nan, Knösche, & Friederici, 2009; Nan et al., 2006) tested Chinese and German musicians' and non-musicians' ability to detect phrase boundaries cross-culturally in unfamiliar excerpts. Results showed a clear in-culture advantage on the behavioral task, and early positive ERP components (100–450 ms) distinguished the two groups of participants for Chinese music (familiar only to the Chinese participants). Both groups exhibited a Closure Positive Shift neurologically suggesting they were sensitive to phrase boundaries in both cultures. A follow-up study with only German participants used an fMRI paradigm (Nan et al., 2008) to scan participants while they heard phrased and unphrased examples of Western and Chinese melodies that they were asked to classify by culture. All participants were better at recognizing in-culture examples and the researchers found that participants exhibited generally higher activation when listening to the Chinese melodies in regions associated with attention and auditory processing suggesting that out-of-culture music is more demanding for those processes.

In most of the studies reviewed thus far, there are differences with in-culture and out-of-culture responses to a variety of musical tasks from emotion and preference to basic musical structures. However, the results are almost always tempered by an awareness that some aspects of music processing can be done without relying on culturally specific strategies, using more global cues and responding to familiar sounding aspects of unfamiliar cultures. In the next section, we review a series of studies on cross-cultural music memory that have led us to propose a possible explanatory framework for musical enculturation.

Cross-Cultural Explorations of Music Memory

In a series of experiments over the last decade or so we have used recognition memory as a way of assessing how effectively in-culture and out-of-culture music is processed. The studies have explored both behavioral (Demorest, Morrison, Beken, & Jungbluth, 2008; Morrison, Demorest, Campbell, Bartolome, & Roberts, 2012; Morrison, Demorest, & Stambaugh, 2008) and neurological (Demorest et al., 2010; Morrison, Demorest, Aylward, Cramer, & Maravilla, 2003) responses to culturally familiar (Western or Turkish) and culturally unfamiliar (Turkish or Chinese) music. In addition, we explored whether memory performance was influenced by training (Demorest et al., 2008; Morrison et al., 2012) or complexity (Morrison et al., 2008). The primary finding of this research has been that there is an "enculturation effect," or cultural bias, in listening such that culturally unfamiliar music is consistently less effectively processed even when considering matters of age, training, and complexity. Further, this effect appears in both Western and non-Western born listeners. This finding was strengthened by the work of another group that tested memory and tension judgments in monomusical and bimusical participants in the United States and India (Wong, Roy, & Margulis, 2009) and found a similar recognition memory effect for monomusical, but not for bimusical, participants. It should be noted that in most cases out-of-culture recognition memory was above chance and demonstrated improvement with repeated testing (Morrison et al., 2012); however, the observed difference between in- and out-of-culture memory performance remained.

Despite the consistency of the enculturation effect, we did not have a good explanation for its cause: that is, what aspect of out-of-culture music was interfering with listeners' ability to hear and remember it? What was so unfamiliar about culturally unfamiliar music? Was it timbre, tonality, rhythm, melody, or some combination? In a recent study (Demorest, Morrison, Nguyen, & Bodnar, 2016), we sought to strip away contextual variables in an attempt to attenuate or eliminate the effects of enculturation on memory performance. We also explored the possible influence of music preference as a variable influencing attention and memory. Western-born participants ($N = 128$) were randomly assigned to conditions in which they heard the same music excerpts presented in one of three contexts: full instrumental ensemble (the original version), a single-voice melody on piano, or a single-voice isochronous pitch sequence also on piano. In each condition participants heard a block of three longer Western art music excerpts and a block of three longer Turkish art music excerpts in a counterbalanced order. After each example, they were asked to rate their preference for the excerpt. After each set of three examples they completed a twelve-item recognition memory test with six targets (taken from the excerpts heard previously) and six foils (taken from a musically different and previously unheard part of the same pieces). Regardless of the listening condition, participants demonstrated superior memory for in-culture examples suggesting that none of the contextual changes mitigated memory performance for out-of-culture music. In-culture memory performance was influenced by context, but out-of-culture memory performance was not. Preference was higher overall for in-culture music, but there was no significant correlation between preference scores and memory performance across cultures. This suggested that the process of enculturation involved a kind of informal learning of deeper structure involving commonly heard sequences of pitch relationships.

Based on these findings we concluded, "If our understandings of out-of-culture music are filtered through in-culture expectations, then a comparison of the statistical properties of a listener's home culture with that of an unfamiliar culture might yield predictive information about subsequent memory performance" (Demorest et al., 2016, p. 597). We labeled the notion of a statistical comparison between music cultures across one or more selected parameters as *cultural distance* (Demorest & Morrison, 2016) in an effort to convey the potentially continuous rather than dichotomous relationship among music cultural practices. In the next section, we will discuss the construct of cultural distance as an explanatory framework and present illustrative work in cross-cultural corpus analysis that lends support to its central premise.

Cultural Distance

Throughout the body of research that examines cross-cultural cognitive processes associated with music, the logic of the underlying design typically sets individuals and/or music examples from one cultural background in contrast with individuals and/or music from another cultural background. Such designs impose a dichotomous relationship between that which is culturally familiar or culturally similar and that which is unfamiliar or dissimilar. On one scale, this might be seen as reflecting the in-group and out-of-group dynamic. However, such bifurcation blurs the fluidity that characterizes musical interactions (Cross, 2008). That is, from the point of view of an individual encultured in a particular music tradition, the music of a culturally unfamiliar tradition may seem surprisingly accessible in one case or virtually impenetrable in another. It is this distinction—and the continuum of increasing or decreasing similarity from one's own music—that we propose can be productively explored using the concept of cultural distance (Demorest & Morrison, 2016).

The way in which an individual interacts with music is mediated by the properties common to the prevailing music of that individual's culture. The music on which one was "brought up" provides the framework by which subsequent music experiences are judged as typical or atypical. Put another way, the statistical likelihood of events that characterize the music of one's home culture governs not only the way in which one interacts with novel pieces from within that same cultural tradition, but also with music from culturally unfamiliar music traditions. One scans for common and familiar patterns both where they are likely to be found and where they may not be likely at all. This situation suggests a way in which an individual's responses to and facility with culturally unfamiliar music may be interpreted or, indeed, predicted. Specifically, we have hypothesized that

> the degree to which the musics of any two cultures differ in the statistical patterns of pitch and rhythm will predict how well a person from one of the cultures can process the music of the other. (Demorest & Morrison, 2016, p. 189)

Based on this *cultural distance hypothesis*, music cultures with considerable overlap of patterns would likely allow for more efficient and effective processing that might be observed through such responses as recognition memory, error detection, phrase parsing, or metric identification, to name a few.

In order to test this proposition, we first need a way to ascertain the statistical properties of structural parameters considered typical of a given culture's music. IDyOM (Information Dynamics of Music; Pearce, 2005) is a computational model of auditory expectation that uses statistical learning and probabilistic prediction to acquire and process internal representations of the structure of a musical style. Using the intervallic content of melody as an illustration, IDyOM generates a probability distribution over the set of possible intervals leading to each note in the melody. IDyOM generates probability distributions that are conditioned upon the preceding musical context and the prior musical experience of the model. The probability of each note can be log-transformed to yield its *information content* according to the model (MacKay, 2003), which reflects how unexpected the model finds a note in a particular context. IDyOM is a variable-order Markov model (Begleiter, El-Yaniv, & Yona, 2004; Bell, Cleary, & Witten, 1990; Bunton, 1997; Cleary & Teahan, 1997) which uses a multiple-viewpoint framework (Conklin & Witten, 1995) to represent music. This means that IDyOM has several features that go beyond the capabilities of standard Markov (or n-gram) models: first, it combines predictions from models of different order (using different length contexts for prediction); second, it adapts the maximum order used depending on the context; third, it combines predictions from a long-term model (intended to reflect effects of long-term exposure to a musical style) and a short-term model (reflecting dynamic learning of repeated structure within a given piece of music); and fourth, it is able to combine models of different representations of the musical surface (e.g., chromatic pitch, pitch contour, pitch interval and scale degree for predicting pitch; duration, duration ratio, duration contour for predicting rhythm).

IDyOM has been shown to predict accurately Western listeners' pitch expectations in behavioral, physiological, and EEG studies (e.g., Egermann, Pearce, Wiggins, & McAdams, 2013; Hansen & Pearce, 2014; Omigie, Pearce, & Stewart, 2012; Omigie, Pearce, Williamson, & Stewart, 2013; Pearce, 2005; Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010). In many circumstances, IDyOM provides a more accurate model of listeners' pitch expectations than static rule-based models (e.g., Narmour, 1990; Schellenberg, 1997). Rule-based models consist of fixed rules (e.g., a small interval is expected to be followed by another small interval in the same direction) which cannot be modified by experience and therefore do not predict any differences in perception between music cultures. Although such models may describe the perception of listeners from a given culture they do not constitute accurate models of cognition since they cannot account for the observed effects of enculturation reviewed above, and they often prove less accurate than IDyOM in accounting for within-culture perception (Hansen & Pearce, 2014; Pearce, 2005; Pearce, Ruiz, et al., 2010). Furthermore, IDyOM accounts well for other psychological processes in music perception, including similarity perception (Pearce & Müllensiefen, 2017), recognition memory performance (Agres, Abdallah, & Pearce, 2018), phrase boundary perception (Pearce, Müllensiefen, & Wiggins, 2010), and aspects of emotional experience (Egermann et al., 2013; Gingras et al., 2015; Sauvé, Sayad, Dean, & Pearce, 2017).

To illustrate the construct of cultural distance, we trained three IDyOM models to simulate listeners with enculturation in three different musical styles: first, a *Western model* trained on a corpus of European folk songs to simulate the perception of a Western listener enculturated in Western tonal music; second, a *Chinese model* trained on a corpus of Chinese folk songs to simulate the perception of a Chinese listener enculturated in Chinese traditional music; and third, a *Turkish model* trained on a corpus of Turkish Makam melodies to simulate the perception of a Turkish listener enculturated in Turkish Makam music. The corpus of Western tonal music consists of 769 German folk songs from the Essen Folk Song Collection (Schaffrath, 1992, 1994, 1995), extracted from the datasets *fink* and *erk*. The corpus of Chinese music consists of 858 Chinese folk songs from the Essen Folk Song Collection, extracted from the datasets *han* and *natmin*. The corpus of Turkish Makam music consists of 805 Makam melodies extracted from the SymbTR database (Karaosmanoğlu, 2012).[1] See Table 1 for further details of the corpora used to train the model simulations.

Empty and non-monophonic compositions were first removed from all corpora. Furthermore, we removed duplicate compositions using a conservative procedure that considers two compositions duplicates if they share the same opening four melodic pitch intervals regardless of rhythm. The pitch system used in Turkish Makam music is microtonal and does not precisely map onto the Western (approximately) twelve-fold equal division of the octave (Bozkurt, Ayangil, & Holzapfel, 2014). Since IDyOM's pitch matching is exact this would cause the Western and Chinese models to assign zero probabilities to every pitch in the Turkish corpus. A simple (though not unproblematic) way of addressing this issue is to round each pitch in the Turkish corpus to the nearest semitone, which enables comparisons to be made between the corpora. For studies with Western participants, this corresponds to the assumption that listeners perceive microtonal pitches categorically, aggregating microtonal pitches to the nearest semitone category. There is some evidence that listeners do in fact perceive pitch

[1] The Essen Folk Song Collection was retrieved from: http://kern.humdrum.org/cgi-bin/browse?l=/essen. The SymbTR database was retrieved from: https://github.com/MTG/SymbTr.

Table 1. Corpora used in modeling cultural distance and stimulus selection

Corpus	Source	Number of melodies (before duplicates removed)	Number of notes (before duplicates removed)	Mean number of notes per melody
Western	Essen Folk Song Collection (*fink*, *erk*)	769 (2,240)	37,340 (112,042)	48
Turkish	SymbTr Makam melodies	805 (1,935)	307,041 (718,380)	381
Chinese	Essen Folk Song Collection (*han*, *natmin*)	858 (1,994)	57,677 (126,321)	67

categorically in this way, at least in certain circumstances (Burns & Campbell, 1994; Perlman & Krumhansl, 1996). In this example, any responses among Western listeners that demonstrated differences between Western melodies and these "pitch-Westernized" Turkish melodies would underestimate the dissimilarity experienced between the two corpora, conservatively producing type II errors (false negatives) rather than type I errors (false positives).

Each model was used to make both within-culture and between-culture predictions. For the within-culture predictions, IDyOM estimates the information content of every event in every composition in the corpus, using ten-fold cross-validation (Kohavi, 1995) to create training and test sets from the same corpus. For between-culture predictions, IDyOM is first trained on the within-culture corpus (e.g., the Western corpus for the Western model) and then estimates the information content of every note in every composition in a different corpus representing the comparison culture (e.g., the Chinese or Turkish corpus for the Western model). IDyOM was configured to use only its long-term model (or LTM, simulating long-term exposure to a musical style) trained on the appropriate corpus; the short-term model (simulating dynamic learning of repeated patterns within a piece of music) was not used. Other than these differences regarding training corpora, all models were configured identically using the default parameters described in Pearce (2005). In all cases, information content was averaged across notes for each composition yielding a value representing the mean unpredictability of that composition for a given model.

For each comparison between cultures (Western vs. Turkish, Western vs. Chinese, Turkish vs. Chinese), we then plot the data for each composition in the two corresponding corpora: information content for one model is plotted on the abscissa while information content for the second model is plotted on the ordinate. The line of equality ($x = y$) indicates equivalence between the two models. Compositions lying on this line do not distinguish the two cultures, being equally predictable for each model; in other words, they should be equally familiar and predictable to listeners enculturated in either of the two cultures. Positions near the origin represent compositions that are simple within both cultures—that is, they are highly predictable insofar as most incidences of a selected feature are quite common—while positions far from the origin represent compositions that are complex—unpredictable, uncommon—within both cultures. Positions further away from the line of equality represent compositions that are predictable for the simulated model of one culture but unpredictable for the simulated model of the other culture. Distance from the line of equality, therefore, provides a quantitative measure of cultural distance based on information-theoretic modeling of enculturation in musical styles. Fig. 1A illustrates how cultural distance is computed for a comparison between IDyOM models trained on the Western corpus and the Chinese corpus using a pitch interval representation. By rotating the data points through 45°, Fig. 1B shows the same data with Cultural Distance on the ordinate and culture-neutral complexity on the abscissa. In this example, IDyOM correctly classifies 98 percent of the folk songs by culture (Chinese vs. Western).

As mentioned above, IDyOM is capable of modeling different attributes of the musical surface and combining the predictions made by those models. For each comparison between cultures, cultural distance is computed for models predicting pitch structure alone (using a representation of pitch interval), rhythmic structure alone (using a representation of inter-onset interval), and for models using a combined representation of pitch and rhythmic structure (for which a melodic event is represented as a pair of values, one for the preceding pitch interval and one for the preceding inter-onset interval).

For each cultural comparison and each of the three representations, ten compositions with the highest Cultural Distance were selected for each of the two cultures compared. These compositions are highlighted in Fig. 1 for the pitch interval representation. Table 2 shows the mean Cultural Distance values for each combination of cultural comparison and model representation for the corpus as a whole and for the ten selected compositions. Note that this Cultural Distance measure reflects both corpora included in the comparison. Thus, there is only partial overlap between the different comparisons (e.g., five of the ten Chinese songs selected in the German comparison are the same as those selected in the Turkish comparison; five for the two Turkish comparisons and two for the two German comparisons). Note also that this Cultural Distance measure may be asymmetrical such that one culture is on average more distant from the second than the second is from the first (e.g., in the case of the Western and Chinese comparison, see Table 2). For all three cultural comparisons, as shown in Table 2, the IDyOM simulations produce positive correlations between the cultures for rhythm predictions much more so than pitch predictions which yield no correlation (Western/Chinese), a small positive correlation (Western/Turkish), or a moderate negative correlation (Turkish/Chinese). This suggests that pitch is a more important indicator of cultural distance between these styles than rhythm. For each of the three representations used in each of the three comparisons, one-sample t-tests indicate that the mean cultural distance is significantly different from zero ($p < 0.01$) for both corpora involved in the comparison.

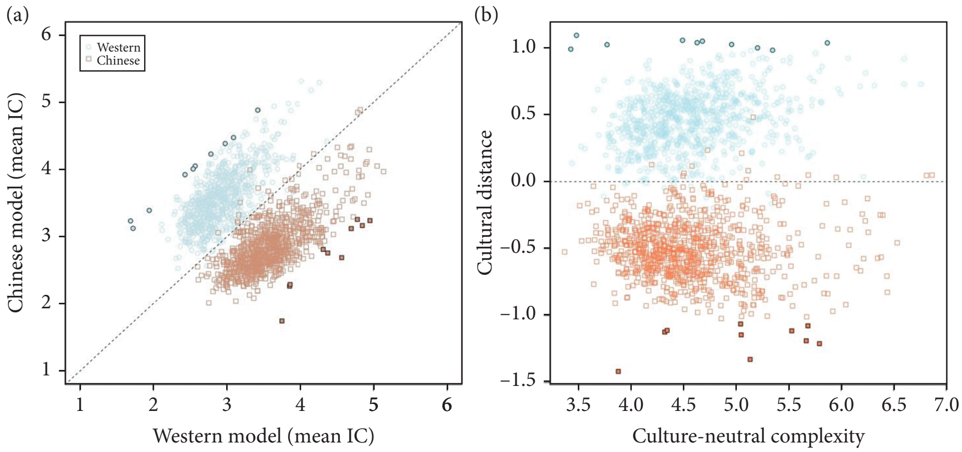

FIGURE 1. Modeling cultural distance between the Western and Chinese corpora using a pitch interval representation. **A**: The information content of the Western model plotted against that of the Chinese model with the x = y line shown. **B**: A 45° rotation of A such that the ordinate represents cultural distance and the abscissa culture-neutral complexity. For each style, the ten compositions with most extreme cultural distance are highlighted.

Limitations

The analysis of two or more types of music along any given musical parameter (for example, pitch as in the illustration above) or combination of parameters imposes the assumption that such an analysis is valid within each music type. While a music tradition such as Western art music (at least that from or deriving from the common practice period of approximately the mid-seventeenth to early twentieth centuries) has a well-established history of analysis and interpretation based, in part, on both sequential and concurrent pitch interval relationships, the same may not be said of other traditions. Tools such as IDyOM offer the flexibility to examine cultural distance according to a variety of individual or combinations of musical parameters. Nevertheless, any specific configuration runs the risk of privileging one parametric hierarchy over another. Thus, in terms of cross-cultural research, such statistical models will virtually always impose the perspective of a particular music tradition, at least to some degree.

This limitation has ramifications for fully comparative studies in that the degree to which a parameter holds primacy for one set of participants may not hold true for the other. Much as emotion recognition, so familiar to the experience of westernized listeners, did not figure meaningfully in the music tradition of the Mafa (Fritz, 2013), the statistical likelihood of patterns of pitch may contribute less to musical thinking among Rwandans and more to North Americans (as in Cameron et al., 2015) than does the complexity of patterns of rhythm. In this way, cultural distance is a tool through which one can isolate norms for one or more musical parameters as well as provide a particular perspective on musical meaning-making.

Table 2. Mean cultural distance values for the entire corpus and selected stimuli for the two styles involved in each comparison. Data are reported for models predicting pitch alone, rhythm alone and both pitch and rhythm

				Mean Cultural Distance			
				Corpus		Selected Stimuli	
Culture 1	Culture 2	Representation	IC Correlation	Culture 1	Culture 2	Culture 1	Culture 2
Western	Turkish	Pitch	0.20, $p < 0.01$	−0.39	0.41	−1.27	1.39
		Rhythm	0.52, $p < 0.01$	−0.18	0.68	−0.77	5.9
		Pitch + Rhythm	0.08, $p < 0.01$	−0.57	1.08	−1.52	6.44
Western	Chinese	Pitch	−0.04, $p = 0.11$	−0.46	0.51	−1.03	1.18
		Rhythm	0.44, $p < 0.01$	−0.4	0.11	−1.92	1.71
		Pitch + Rhythm	0.00, $p = 0.99$	−0.86	0.62	−2.35	2.23
Turkish	Chinese	Pitch	−0.52, $p < 0.01$	−0.91	1.06	−1.96	2.11
		Rhythm	0.68, $p < 0.01$	−0.4	0.07	−4.47	0.36
		Pitch + Rhythm	−0.26, $p < 0.01$	−1.32	1.13	−5.54	2.2

A related limitation is that IDyOM currently requires symbolic score-like input in which notes are represented as discrete events with discrete properties (e.g., onset time, pitch). This does not readily accommodate musical cultures which depend heavily on timbral, dynamic, or textural changes. The same is true of musical cultures that have no written tradition, where the distinction between composition and performance is blurred or nonexistent or where music is inextricably combined with other modes of communication (Cross, 2014).

Despite the emphasis here on the advantageous aspects of familiarity, without question novelty is an attractive characteristic of music. Models of musical expectancy (e.g., Huron, 2006; Meyer, 1956) describe the interest inherent in and stimulation derived from that which is unfamiliar and surprising in music. The constant curiosity for new musical ideas suggests ongoing willingness to explore less "predictable" musical scenarios. With much of the world's music readily—and in many cases instantly—accessible, such willingness leads as easily to unfamiliar music traditions as to the remoter corners of one's own. We have used cultural distance as a means of explaining processing difficulties (as operationalized by recognition memory); however, it is equally viable as a tool to examine such positive aspects of music experience as interest and surprise. Although Cook (2008) was referring specifically to musicologists, his description can arguably be construed more broadly: "Practically all of us are at least to some degree musically multilingual . . . as a result one understands even the tradition(s) in which one is most 'at home' as options amongst other options, understands them in relation to other traditions rather than as absolutes" (p. 63).

Conclusion

Research on cross-cultural music interactions has demonstrated that responses to culturally familiar and unfamiliar music, as well as responses by individuals encultured in different music traditions, can be either remarkably similar or strikingly different depending on the task and the music presented. Theoretical models such as Cue Redundancy (Balkwill & Thompson, 1999) or Fritz's (2013) dock-in model, have framed cross-cultural music interactions as consisting of culture-general and culture-specific components. The manner in which these models account for areas of overlap between music cultures and distinctions unique to each music culture fit well with recent research findings as well as with the concept of cultural distance. However, absent from their construal of shared and unique features is a middle ground of "culturally specific but similar" components that, while mutually proprietary and uniquely meaningful to each culture, may be somewhat accommodating to strategies for listening, performing, and meaning-making deployed by individuals from outside the culture.

This similar-but-not-shared aspect of the cultural distance construct can help account for memory responses, reported above, to out-of-culture music that were less successful than for in-culture music but were still above chance (e.g., Demorest et al., 2008).

Likewise, it also provides an explanation in cases where listeners have applied familiar listening strategies to culturally unfamiliar music only to encounter ultimate confusion (e.g., Curtis & Bharucha, 2009). Eventually, the trajectory of complexity within a culturally unfamiliar system takes a listener or performer past where learned patterns can accommodate. On the whole, responses to musics that demonstrate considerable overlap may show greater consistency than those to musics with very few points of commonality. Thus, one can make a distinction between the apparent "ease" with which an individual can move between music cultures and the more likely case of greater opportunities afforded by some unfamiliar music cultures to successfully deploy familiar strategies.

This is potentially useful for neurological investigations of music processing. Responses to culturally unfamiliar music have generally been reported to differ more by degree than by presence or location. That is, music appears to recruit similar neural systems regardless of its cultural familiarity, though the strength or extent of that activity may differ according to the music encountered (e.g., Nan et al., 2008; Demorest et al., 2010). The model of cultural distance is a tool that provides a continuous rather than categorical conceptualization of cross-cultural music research designs. Such a correlational approach may lend itself well to the fine-grained, incremental, and plastic manner in which neurological processes and pathways develop and are deployed.

We are not suggesting that through the learning of an unfamiliar array of patterns one can gain access to the full, rich experience of culturally situated musical contexts. Music represents a broad range of activities and relationships that may only have tenuous connections to structural parameters like melodic or rhythmic intervals. Much of music's meaning is derived from where, when, and how it occurs quite apart from how it is put together (Small, 1998). Rather, we suggest that cultural distance may be a useful lens through which specific aspects of the cognitive processing of music—particularly musical structure—may be predicted, investigated, analyzed, and interpreted.

Much of the research on cross-cultural musical interactions has involved measurement of such things as memory, affective response, detection of differences, verbal or written description, and preference. In virtually all cases these outcomes were prompted through listening tasks, a way of experiencing music that, while ecologically valid and obviating any need for previous training, is covert and arguably accommodating of varied interpretations and strategies. In contrast, investigations of cross-cultural performance contexts may yield new insights into the ways in which individuals navigate unfamiliar musical terrain. More directly observable performance-based interactions may shed additional light on the processes by which one grapples with, accommodates, or eventually gains facility with musics that are differently organized.

Earlier we posed the question of what happens when music crosses cultural boundaries. The construct of cultural distance provides a more graduated, incremental way of conceptualizing the relationship between the familiar and the unfamiliar. It allows for the fluidity characteristic of musical interactions, recognizes the porous nature of music categorization, and accounts for the variability found within any music tradition. For research purposes, cultural distance offers a way by which dichotomous models of music—insider/outsider, familiar/unfamiliar, own/other—can be refined to test a more nuanced picture of musical meaning-making. In this way, cross-cultural music interactions might be viewed less as the crossing of a boundary and more as the undertaking of a trip.

References

Agres, K., Abdallah, S., & Pearce, M. T. (2018). Information-theoretic properties of auditory sequences dynamically influence expectation and memory. *Cognitive Science 42*(1), 43–76.

Baek, Y. M. (2015). Relationship between cultural distance and cross-cultural music video consumption on YouTube. *Social Science Computer Review 33*(6), 730–748.

Balkwill, L. L. (2006). Perceptions of emotion in music across cultures. Paper presented at Emotional Geographies: The Second International & Interdisciplinary Conference, May, Queen's University, Kingston, Canada.

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

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

Begleiter, R., El-Yaniv, R., & Yona, G. (2004). On prediction using variable order Markov models. *Journal of Artificial Intelligence Research 22*, 385–421.

Bell, T. C., Cleary, J. G., & Witten, I. H. (1990). *Text compression*. Englewood Cliffs, NJ: Prentice Hall.

Bozkurt, B., Ayangil, R., & Holzapfel, A. (2014). Computational analysis of Makam music in Turkey: Review of state-of-the-art and challenges. *Journal of New Music Research 43*(1), 3–23.

Brown, S., & Jordania, J. (2013). Universals in the world's musics. *Psychology of Music 41*(2), 229–248.

Bunton, S. (1997). Semantically motivated improvements for PPM variants. *The Computer Journal 40*(2–3), 76–93.

Burns, E. M., & Campbell, S. L. (1994). Frequency and frequency-ratio resolution by possessors of absolute and relative pitch: Examples of categorical perception. *Journal of the Acoustical Society of America 96*(5), 2704–2719.

Cameron, D. J., Bentley J., & Grahn J. A. (2015). Cross-cultural influences on rhythm processing: Reproduction, discrimination, and beat tapping. *Frontiers in Psychology 6*, 366. Retrieved from https://doi.org/10.3389/fpsyg.2015.00366

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.

Chiao, J. Y., Iidaka, T., Gordon, H. L., Nogawa, J., Bar, M., Aminoff, E., . . . Ambady, N. (2008). Cultural specificity in amygdala response to fear faces. *Journal of Cognitive Neuroscience 20*(12), 2167–2174.

Cleary, J. G., & Teahan, W. J. (1997). Unbounded length contexts for PPM. *The Computer Journal 40*(2–3), 67–75.

Conklin, D., & Witten, I. H. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research 24*(1), 51–73.

Cook, N. (2008). We are all (ethno)musicologists now. In H. Stobart (Ed.), *The new (ethno) musicologies* (pp. 48–70). Lanham, MD: Scarecrow Press.

Cross, I. (2008). Musicality and the human capacity for culture. *Musicae Scientiae 12*(1 Suppl.), 147–167.
Cross I. (2014). Music and communication in music psychology. *Psychology of Music 42*(6), 809–819.
Curtis, M. E., & Bharucha, J. J. (2009). Memory and musical expectation for tones in cultural context. *Music Perception 26*(4), 365–375.
Demorest, S. M., & Morrison, S. J. (2016). Quantifying culture: The cultural distance hypothesis of melodic expectancy. In J. Y. Chiao, S.-C. Li, R. Seligman, & R. Turner (Eds.), *The Oxford handbook of cultural neuroscience* (pp. 183–194). Oxford: Oxford University Press.
Demorest, S. M., Morrison, S. J., Beken, M. N., & Jungbluth, D. (2008). Lost in translation: An enculturation effect in music memory performance. *Music Perception 25*(3), 213–223.
Demorest, S. M., Morrison, S. J., Beken, M. N., Stambaugh, L. A., Richards, T. L., & Johnson, C. (2010). Music comprehension among western and Turkish listeners: fMRI investigation of an enculturation effect. *Social Cognitive and Affective Neuroscience 5*, 282–291.
Demorest, S. M., Morrison, S. J., Nguyen, V. Q., & Bodnar, E. N. (2016). The influence of contextual cues on cultural bias in music memory. *Music Perception 33*(5), 590–600.
Demorest, S. M., & Osterhout, L. (2012). ERP responses to cross-cultural melodic expectancy violations. *Annals of the New York Academy of Sciences 1252*, 152–157.
Demorest, S. M., & Schultz, S. J. (2004). Children's preference for authentic versus arranged versions of world music recordings. *Journal of Research in Music Education 52*(4), 300–313.
Deva, B. C., & Virmani, K. G. (1975). A study in the psychological response to ragas. (Research Report II of Sangeet Natak Akademi). New Delhi, India: Indian Musicological Society.
Drake, C., & Ben El Heni, J. (2003). Synchronizing with music: Intercultural differences. *Annals of the New York Academy of Sciences 999*, 429–437.
Egermann, H., Fernando, N., Chuen, L., & McAdams, S. (2015). Music induces universal emotion-related psychophysiological responses: Comparing Canadian listeners to Congolese Pygmies. *Frontiers in Psychology 5*, 1341. Retrieved from https://doi.org/10.3389/fpsyg.2014.01341
Egermann, H., Pearce, M. T., Wiggins, G. A., & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective & Behavioral Neuroscience 13*(3), 533–553.
Flowers, P. J. (1980). Relationship between two measures of music preference. *Contributions to Music Education 8*, 47–54.
Frith, S. (1996). Music and identity. In S. Hall & P. Du Gay (Eds.), *Questions of cultural identity* (pp. 108–127). London: Sage Publications.
Fritz, T. (2013). The dock-in model of music culture and cross-cultural perception. *Music Perception: An Interdisciplinary Journal 30*(5), 511–516.
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.
Fung, C. V. (1994). Undergraduate nonmusic majors' world music preference and multicultural attitudes. *Journal of Research in Music Education 42*(1), 45–57.
Gingras, B., Pearce, M. T., Goodchild, M., Dean, R. T., Wiggins, G., & McAdams, S. (2015). Linking melodic expectation to expressive performance timing and perceived musical tension. *Journal of Experimental Psychology: Human Perception & Performance 42*(4), 594–609.
Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural

responses to pitch change. *Frontiers in Psychology 2*, 146. Retrieved from https://doi.org/10.3389/fpsyg.2011.00146

Golby, A. J., Gabrieli, J. D., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience 4*, 845–850.

Gregory, A. H., & Varney, N. (1996). Cross-cultural comparisons in the affective response to music. *Psychology of Music 24*(1), 47–52.

Hannon, E. E. (2009). Perceiving speech rhythm in music: Listeners classify instrumental songs according to language of origin. *Cognition 111*(3), 403–409.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science 16*(1), 48–55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences 102*(35), 12639–12643.

Hansen, N. C., & Pearce, M. T. (2014). Predictive uncertainty in auditory sequence processing. *Frontiers in Psychology 5*, 1–17. Retrieved from https://doi.org/10.3389/fpsyg.2014.01052

Heingartner, A., & Hall, J. V. (1974). Affective consequences in adults and children of repeated exposure to auditory stimuli. *Journal of Personality and Social Psychology 29*(6), 719–723.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences 33*(2–3), 61–83.

Hofstede, G. (1983). National cultures in four dimensions: A research-based theory of cultural differences among nations. *International Studies of Management & Organization 13*(1–2), 46–74.

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

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America 124*, 2263–2271.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance 26*(6), 1797–1812.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal emotion and music performance: Different channels, same code? *Psychological Bulletin 129*(5), 770–814.

Karaosmanoğlu, M. K. (2012). A Turkish Makam music symbolic database for music information retrieval: Symbtr. In *Proceedings of the 13th ISMIR Conference, Porto, Portugal*, 223–228.

Keil, A., & Keil, C. (1966). A preliminary report: The perception of Indian, Western, and Afro-American musical moods by American students. *Ethnomusicology 10*(2), 153–173.

Kessler, E. J., Hansen, C., and Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and the West. *Music Perception 2*(2), 131–65.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence* (Vol. 2, pp. 1137–1145). San Mateo, CA: Morgan Kaufmann.

Krumhansl, C. L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum 17*(1), 53–80.

Krumhansl, C. L., Louhivuori, J., Toiviainen, P., Jarvinen, T., & Eerola, T. (1999). Melodic expectation in Finnish spiritual folk hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception 17*(2), 151–195.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance 5*(4), 579–594.

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.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Attention, Perception, & Psychophysics 50*(2), 93–107.

Laukka, P., Eerola, T., Thingujam, N. S., Yamasaki, T., & Beller, G. (2013). Universal and culture-specific factors in the recognition and performance of musical affect expressions. *Emotion 13*(3), 434–449.

LeBlanc, A. (1982). An interactive theory of music preference. *Journal of Music Therapy 19*(1), 28–45.

Lynch, M. P., & Eilers, R. E. (1991). Children's perception of native and nonnative musical scales. *Music Perception 9*(1), 121–131.

Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Perception & Psychophysics 52*(6), 599–608.

Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science 1*(4), 272–276.

Lynch, M. P., Eilers, R. E., Oller, K. D., Urbano, R. C., & Wilson, P. (1991). Influences of acculturation and musical sophistication on perception of musical interval patterns. *Journal of Experimental Psychology: Human Perception and Performance 17*(4), 967–975.

Lynch, M. P., Short, L. B., and Chua, R. (1995). Contributions of experience to the development of musical processing in infancy. *Developmental Psychobiology 28*(7), 377–398.

MacKay, D. J. C. (2003). *Information theory, inference, and learning algorithms*. Cambridge: Cambridge University Press.

Merton, R. K. (1972). Insiders and outsiders: A chapter in the sociology of knowledge. *American Journal of Sociology 78*(1), 9–47.

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

Morrison, S. J., & Demorest, S. M. (2009). Cultural constraints on music perception and cognition. *Progress in Brain Research 178*, 67–77.

Morrison, S. J., Demorest, S. M., Aylward, E. H., Cramer, S. C., & Maravilla, K. R. (2003). fMRI investigation of cross-cultural music comprehension. *NeuroImage 20*(1), 378–384.

Morrison, S. J., Demorest, S. M., Campbell, P. S., Bartolome, S. J., & Roberts, J. C. (2012). Effect of intensive instruction on elementary students' memory for culturally unfamiliar music. *Journal of Research in Music Education 60*(4), 363–374.

Morrison, S. J., Demorest, S. M., & Stambaugh, L. A. (2008). Enculturation effects in music cognition: The role of age and music complexity. *Journal of Research in Music Education 56*(2), 118–129.

Morrison, S. J., & Yeh, C. S. (1999). Preference responses and use of written descriptors among music and nonmusic majors in the United States, Hong Kong, and the People's Republic of China. *Journal of Research in Music Education 47*(1), 5–17.

Nan, Y., Knösche, T. R., & Friederici, A. D. (2006). The perception of musical phrase structure: A cross-cultural ERP study. *Brain Research 1094*(1), 179–191.

Nan, Y., Knösche, T. R., & Friederici, A. D. (2009). Non-musicians' perception of phrase boundaries in music: A cross-cultural ERP study. *Biological Psychology 82*(1), 70–81.

Nan, Y., Knösche, T. R., Zysset, S., & Friederici, A. D. (2008). Cross-cultural music phrase processing: An fMRI study. *Human Brain Mapping 29*(3), 312–328.

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

Neuhaus, C. (2003). Perceiving musical scale structures: A cross-cultural event-related brain potentials study. *Annals of the New York Academy of Sciences 999*, 184–188.

Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in congenital amusia. *Neuropsychologia 50*(7), 1483–1493.

Omigie, D., Pearce, M. T., Williamson, V. J., & Stewart, L. (2013). Electrophysiological correlates of melodic processing in congenital amusia. *Neuropsychologia 51*(9), 1749–1762.

Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition 87*(1), B35–B45.

Patel, A. D., & Demorest, S. M. (2013). Comparative music cognition: Cross-species and cross-cultural studies. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 647–681). London: Academic Press.

Pearce, M. T. (2005). *The construction and evaluation of statistical models of melodic structure in music perception and composition* (Doctoral dissertation). Department of Computing, City University, London.

Pearce, M. T., & Müllensiefen, D. (2017). Compression-based modelling of musical similarity perception. *Journal of New Music Research 46*(2), 135–155.

Pearce, M. T., Müllensiefen, D., & Wiggins, G. A. (2010). Melodic grouping in music information retrieval: New methods and applications. In Z. W. Ras & A. Wieczorkowska (Eds.), *Advances in music information retrieval* (pp. 364–388). Berlin: Springer.

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural and neural manifestations of musical expectation. *NeuroImage 50*(1), 302–313.

Perlman, M., & Krumhansl, C. L. (1996). An experimental study of internal interval standards in Javanese and Western musicians. *Music Perception 14*(2), 95–116.

Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception, & Psychophysics 71*(6), 1385–1398.

Polak, R., London, J., & Jacoby, N. (2016). Both isochronous and non-isochronous metrical subdivision afford precise and stable ensemble entrainment: A corpus study of Malian djembe drumming. *Frontiers in Neuroscience 10*, 285. Retrieved from https://doi.org/10.3389/fnins.2016.00285

Raman, R., & Dowling, W. J. (2016). Real-time probing of modulations in South Indian classical (Carnatic) music by Indian and Western musicians. *Music Perception 33*(3), 367–393.

Raman, R., & Dowling W. J. (2017). Perception of modulations in south Indian classical (Carnatic) music by student and teacher musicians: A cross-cultural study. *Music Perception 34*(4), 424–437.

Renninger, L. B., Wilson, M. P., & Donchin, E. (2006). The processing of pitch and scale: An ERP study of musicians trained outside of the western musical system. *Empirical Musicology Review 1*(4), 185–197.

Sauvé, S., Sayad, A., Dean, R. T., & Pearce, M. T. (2017). Effects of pitch and timing expectancy on musical emotion. arXiv Preprint, 1708.03687.

Schaffrath, H. (1992). The ESAC databases and MAPPET software. *Computing in Musicology 8*, 66.

Schaffrath, H. (1994). The ESAC electronic songbooks. *Computing in Musicology 9*, 78.

Schaffrath, H. (1995). The Essen folksong collection. In D. Huron (Ed.), *Database containing 6,255 folksong transcriptions in the Kern format and a 34-page research guide* [computer database]. Menlo Park, CA: CCARH.

Schellenberg, E. G. (1997). Simplifying the implication-realization model of melodic expectancy. *Music Perception 14*(3), 295–318.

Shehan, P. K. (1981). Student preferences for ethnic music styles. *Contributions to Music Education 9*, 21–28.

Shehan, P. K. (1985). Transfer of preference from taught to untaught pieces of non-Western music genres. *Journal of Research in Music Education 33*(3), 149–158.

Small, C. (1998). *Musicking: The meanings of performing and listening*. Middletown, CT: Wesleyan University Press.

Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: A cross-cultural comparison. *Developmental Psychology 46*(1), 286–292.

Stobart, H., & Cross, I. (2000). The Andean anacrusis? Rhythmic structure and perception in Easter songs of northern Potosi, Bolivia. *British Journal of Ethnomusicology 9*(2), 63–92.

Thompson, W. F. & Balkwill, L. L. (2010). Cross-cultural similarities and differences. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 755–790). New York: Oxford University Press.

Trulsson, Y. H., & Burnard, P. (2016). Insider, outsider or cultures in-between. In P. Burnard, E. Mackinlay, & K. Powell (Eds.), *The Routledge international handbook of intercultural arts research* (pp. 115–125). New York: Routledge.

Wong, P. C. M., Chan, A. H. D., Roy, A., & Margulis, E. H. (2011). The bimusical brain is not two monomusical brains in one: Evidence from musical affective processing. *Journal of Cognitive Neuroscience 23*(12), 4082–4093.

Wong, P. C., Ciocca, V., Chan, A. H., Ha, L. Y., Tan, L. H., & Peretz, I. (2012). Effects of culture on musical pitch perception. *PloS ONE 7*(4), e33424.

Wong, P. C. M., Roy, A. K., & Margulis, E. H. (2009). Bimusicalism: The implicit dual enculturation of cognitive and affective systems. *Music Perception 27*(2), 81–88.

Yoshida, K. A., Iversen, J. R., Patel, A. D., Mazuka, R., Nito, H., Gervain, J., & Werker, J. F. (2010). The development of perceptual grouping biases in infancy: A Japanese-English cross-linguistic study. *Cognition 115*(2), 356–361.