Ann. N.Y. Acad. Sci. ISSN 0077-8923

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Special Issue: Annals *Reports*

Original Article

Sweetness is in the ear of the beholder: chord preference across United Kingdom and Pakistani listeners

Imre Lahdelma, George Athanasopoulos, and Tuomas Eerola

Department of Music, Durham University, Durham, United Kingdom

Address for correspondence: Imre Lahdelma, Department of Music, Durham University, Palace Green, DH1 3RL Durham, UK. imre.d.lahdelma@durham.ac.uk

The majority of research in the field of music perception has been conducted with Western participants, and it has remained unclear which aspects of music perception are culture dependent, and which are universal. The current study compared how participants unfamiliar with Western music (people from the Khowar and Kalash tribes native to Northwest Pakistan with minimal exposure to Western music) perceive affect (positive versus negative) in musical chords compared with United Kingdom (UK) listeners, as well as the overall preference for these chords. The stimuli consisted of four distinct chord types (major, minor, augmented, and chromatic) and were played as both vertical blocks (pitches presented concurrently) and arpeggios (pitches presented successively). The results suggest that the Western listener major-positive minor-negative affective distinction is opposite for Northwest Pakistani listeners, arguably because of the reversed prevalence of these chords in the two music cultures. The aversion to the harsh dissonance of the chromatic cluster is present cross-culturally, but the preference for the consonance of the major triad varies between UK and Northwest Pakistani listeners, depending on cultural familiarity. Our findings imply not only notable cultural variation but also commonalities in chord perception across Western and non-Western listeners.

Keywords: music; cross-cultural; consonance; dissonance; harmony; chord

Introduction

Music is often described as a universal language. While music exists in every society,[1] it has remained unclear which aspects of music perception are culture dependent and which are universal. To date, the vast majority of music research has been conducted exclusively with Western participants and Western music, and it has been proposed that this tendency biases the understanding of human mechanisms for music perception.[2] The Western notion of the major-positive minor-negative affective distinction has been historically put forward as "natural" and arising from the overtone series as it exists in nature (e.g., see Ref. 3), and this affective distinction has also been proposed to be based on universal vocal characteristics of different affective states.[4] There is evidence that the minor and major thirds communicate sad and happy affect, respectively, in speech for Western people, mirroring its use in music.[5] Also, smaller musical *intervals* (i.e., two concurrent/successive pitches) express sad/subdued emotion across South Indian and Western musical cultures,[4] which is in line with the minor mode's smaller third-, sixth-, and seventh-scale degrees compared with the major mode. This effect has also been shown through exposing individuals to an unconventionally tuned scale where participants subsequently perceived melodies with scale degrees that were lowered in frequency relative to those in the exposure scale as expressively sadder.[6] However, later research has not fully succeeded in replicating this finding.[7]

Very few cross-cultural studies have directly investigated the perception of this affective mode

doi: 10.1111/nyas.14655

distinction, and none of them have investigated it specifically with single isolated *chords* (i.e., three or more concurrent/successive pitches)—Fritz *et al.* found that Mafa tribe members in Cameroon without exposure to Western music classified the majority of major pieces as happy, and most of the minor pieces as scared/fearful.[8] However, this study did not investigate the role of mode isolated from other musical cues (for example, tempo, timbre, and note density) that covaried with mode, although the important effect of covarying cues has been empirically demonstrated (e.g., see Ref. 9) and is dependent on musical style and historical period.[10] A later study by Fang *et al.* found that Chinese listeners make the Western-style major-happy minor-sad affective mode distinction, although this tendency was influenced by previous exposure to Western music.[11]

Another important aesthetic distinction in Western music in addition to the major/minor mode, namely, the preference for consonance over dissonance (i.e., the relative agreeableness/stability versus disagreeableness/need of resolution of pitch combinations), has been surmised to be either a biologically defined universal[12] or, alternatively, exclusively a product of enculturation[13] that is unique to Western music.[14] It has been indicated that infants may have a biological predisposition to prefer consonance over dissonance,[15,16] but the notion of the innateness of consonance preferences in infants has also been challenged.[17] The handful of cross-cultural consonance/dissonance studies conducted so far have been inconclusive in their findings.[14,18–20] There is lamentably little methodological consistency between these cross-cultural experiments: two of them used only intervals as the experiment stimuli,[18,19] one used both intervals and chords,[14] and one only chords.[20] The used chords are also somewhat misrepresentative in the two latter studies: in the study by McDermott *et al.*, highly dissonant chords were omitted from the stimuli,[14] and in the study by Prete *et al.*, the augmented triad, which in Western music is conventionally a dissonance, was misleadingly labeled as a "consonance."[20] The study by McDermott *et al.* investigated how consonance and dissonance are perceived among the Tsimané, an indigenous population living in the Amazon rainforest (Bolivia) with limited exposure to Western culture.[14] The study concluded that the Tsimané are indifferent to consonance/dissonance, although looking closely at the data presented, it seems that the Tsimané also had an aversion to the minor second (the most dissonant interval in Western music) when it was presented diotically (simultaneous presentation to each ear),[14] and the minor second itself is not nearly as objectively dissonant as some of the most dissonant chords in Western music.[21] This, however, was not discussed in the original paper, and later some other methodological confounds[22] have also been raised concerning this seminal investigation into the universality of consonance/dissonance perception.

Recently, a consensus is emerging that the Western notion of consonance/dissonance is a combination of the acoustic phenomena of *roughness* and *harmonicity* and the cultural effect of *familiarity* (e.g., see Ref. 23). The concept of roughness is used to denote the sound quality that arises from the beating of frequency components,[24] and it has a biological substrate: beating occurs at the level of the basilar membrane in the inner ear when the frequency components are too close together to separate (e.g., see Ref. 25). The sensitivity to roughness seems to be present cross-culturally,[14] but its appraisal differs across musical styles and cultures. While a typical Western listener hears roughness as disagreeable,[21] it is harnessed for aesthetic ends (in moderate amounts) in the vocal practice of *beat diaphony* in, for example, the Baltic and Balkan regions of Europe (e.g., see Ref. 26) and in Papua New Guinea.[27] Harmonicity, in turn, indicates how closely a sonority's spectrum corresponds to a harmonic series.[28] Harmonicity has been found to correlate with the preference for consonance,[29] and it has been suggested that humans have an attraction toward the harmonic series that characterizes conspecific vocalizations on the basis of the biological importance of social sound signals.[12] In addition to the acoustic phenomena of roughness and harmonicity, familiarity on both cultural and individual levels has been put forward as an essential contributor to consonance (e.g., see Ref. 30), and the important role of these has indeed been empirically demonstrated.[31,32] These findings imply that consonance perception may have a biological foundation, but it is evidently also notably shaped by culture.

As of recent, some studies[33,34] have applied the Bohlen–Pierce chromatic just intonation tuning scale in an attempt to bypass the effect

Table 1. The chord stimuli and their respective roughness, harmonicity, familiarity, and tonal dissonance values

Chord	MIDI numbers	Roughness	Harmonicity	Tonal dissonance	Familiarity (Western)	Familiarity (Khow)	Familiarity (Kalash)
Major	{54, 58, 61, 66}	0.173 (<25%)	0.929 (>75%)	1	10.00	5.91	1.00
Minor	{54, 57, 61, 66}	0.172 (<25%)	0.929 (>75%)	1	8.72	10.00	10.00
Augmented	{53, 57, 61, 65}	0.201 (<25%)	0.864 (>75%)	2	4.17	5.91	2.20
Chromatic cluster	{56, 57, 58, 67}	0.844 (>75%)	0.605 (<25%)	3	1.00	1.00	7.00

Note: Roughness values were calculated with Hutchinson and Knopoff's model[24] and harmonicity with Harrison and Pearce's model.[45] Values in brackets refer to the quantiles of all possible trichords ($N = 66$). Tetrachords ($N = 220$) were calculated with the respective acoustic models. Tonal dissonance values were taken from Johnson-Laird *et al.*[25] Western familiarity was calculated with Harrison and Pearce's familiarity model[23] and Northwest Pakistani familiarity with a postexperiment survey conducted on Northwest Pakistani participants (Khow and Kalash). Both familiarity indices have been rescaled to values between 1 and 10 for direct comparison. The chromatic cluster was presented in the arpeggiated delivery without the melodic leap by altering the last pitch to be one semitone from the preceding tone {58, 59, 60, 61}.

of enculturation. While this is a commendable effort, it is unclear how "unfamiliar" Bohlen–Pierce chords eventually are. Familiarity with Bohlen–Pierce stimuli is hardly binary but rather a continuum as listeners have been shown to tolerate rather large (between 20 and 45 cents) deviations from equal temperament.[35,36] This observation is also in line with the fact that people can easily adjust to deviations from equal temperament, for example, when listening to vibrato singing, blues intervals, or historical tunings. Hence, we aim to bypass any such possible confounds and will target listeners who are not familiar with Western music practices.

Rationale and hypotheses

The current study aims to compare how participants unfamiliar with Western music—people from the Kalash and Khowar tribes native to Northwest Pakistan with minimal exposure to Western music—perceive affect (positive versus negative) in musical chords compared with UK listeners, and also the overall preference of these chords as a proxy for consonance and dissonance. The current study is a part of a bigger project investigating the perception of music among these two remote Northwest Pakistani tribes—the aim of this paper is to investigate how *single isolated chords* (vertical harmony) are perceived as opposed to looking at background harmonizations in longer musical excerpts (horizontal harmony) across these two non-Western tribes and UK listeners.[37] The Kalash are a remote Indo-European/Aryan polytheistic community, and the Kho people (the tribe is referred to as *Khow*, and the people are referred to as *Kho*) are the native Muslim population of the region of Chitral (Khyber Pakhtunkhwa Province) in Northwest Pakistan. To investigate the research question, we have chosen four different chord types (see Table 1) that vary considerably in relation to their conventional affective connotations and the amount of consonance/dissonance contained in them: major, minor, and augmented triads, and the chromatic cluster, which is a cluster of adjacent pitch-classes (one semitone apart). Previous research has demonstrated that for Western listeners, the major triad typically conveys positive valence, and the minor triad conveys negative valence when the chords are presented individually without a musical context (e.g., see Refs. 38 and 39). However, the universality of this mode distinction has so far remained somewhat unclear in empirical settings.[8,11] The current experiment's Northwest Pakistani tribes are especially interesting in this regard, as in their musical culture, the minor third is significantly more common than the major third, possibly influencing its affective connotations through exposure. With reference to consonance/dissonance, the order of preference from most to least consonant when the chord types differ notably in the amount of consonance/dissonance has been empirically established for Western listeners (e.g., see Ref. 40), but its universality has also remained contentious.[14,18]

We hypothesize that the perception of both affect (positive versus negative) and the overall preference for the chords will differ between UK and Northwest Pakistani listeners in the following ways:

1. UK listeners will perceive the major triad as positive and the minor as negative, but this is unlikely to be the case for Northwest Pakistanis; previous research implies that the affective distinction of major/minor is a learned convention in the West[41] and that exposure shapes how it is perceived outside the West.[11] Hence, we expect the responses to this affective distinction will reflect the musical conventions of Northwest Pakistan instead of representing a universal pattern of perception.
2. UK listeners will prefer the consonant chords (major and minor triads) over the dissonant chords (augmented triad and chromatic cluster). The Northwest Pakistani listeners will have a clear aversion only to the maximally dissonant chromatic cluster but not for the augmented triad, as the latter's dissonance is culture dependent[25] and not related to acoustic roughness (see Table 1). Both groups will prefer the less rough sonorities (major, minor, and augmented triads) over the roughest sonority (chromatic cluster); the aversion to maximal roughness is expected to be present in both UK and Northwest Pakistani listeners as the sensitivity to roughness has been demonstrated to be universal.[14]

Methods

Ethics approval for the study was obtained from the host institution (MUS-2019-01-28T14:54:07-kxhw42). The research project and its methods were further assessed by the High Commission of Pakistan in London and the local authorities in Pakistan (Government of Khyber Pakhtunkhwa, Home & Tribal Affairs Department, Peshawar). Informed consent was a prerequisite to start the experiment. In Northwest Pakistan, it was provided verbally and recorded through a digital voice recorder. Both Kalash and Khow participants were recruited in the field by one of the authors. They predominantly originated from Bumburet and Ayun Valley in the Khyber Pakhtunkhwa Province of Pakistan in the Northwest. The authors recruited the help of a local translator who was present in all experimental and interview sessions with the participants to translate in their own language and in terms meaningful to them throughout the procedure. The participants were selected at random, with no prejudice as to their age, linguistic abilities, or educational background. Owing to local cultural norms, it was not possible to recruit any female Khow participants, as the latter do not readily interact with members outside their own tribe. In the UK, the study was administered in English, and the data were collected online via the *Qualtrics Survey Software*, a web-based survey tool. Informed consent was required at the beginning of the online data collection. In both settings (the UK and Northwest Pakistan), participants who provided consent were given a detailed description as to how the experimental procedure would unfold and were briefed at the end and compensated for their time.

Participants

UK participants. Forty participants from the UK contributed to the research, 24 of whom were females. All participants were native, monolingual English speakers, with a mean age of 29.2 years (SD = 11.3). Only three had played a musical instrument for more than 4 years, and none expressed musical proficiency above an amateur level.

Kho participants. Nineteen Kho participants contributed to the research, all of whom were males. All were native Kho speakers (additional languages spoken at the moderate or fluent level: Urdu, 100%; Pashto, 84%; and Kalash, 33%, among others) with a mean age of 25.3 years (SD = 5.4). Kho participants professed a music performance–based experience of M = 7.7 years. The most common performance instruments included drums (gilliken, dumduma, and doll), group and solo singing, Chitrali sitar, and the Chitrali flute (belu). The participants expressed having limited exposure to any form of Western music and no exposure to music notation; this is mainly due to the lack of a stable electricity grid and communication infrastructure, meaning that people residing in the area have very limited access to the internet. Any rare exposure that they have in relation to Western music is through the occasional contact with a small number of Westerners (health aid workers, and in more recent years, small numbers of tourists) present in the region. Owing to local cultural customs, it was impossible to recruit any female participants among Kho speakers.

Kalash participants. Twenty Kalash participants contributed to the research, out of whom two were

females. All were native Kalash speakers (additional languages spoken at the moderate or fluent level: Urdu; Khow, 100%; and Pashto, 65%, among others), with a mean age of 24.2 years (SD = 4.8). The Kalash participants professed a music performance–based experience in music of M = 9.4 years. Most common performance instruments included drums (gilliken, dau, and wac), singing in groups and solo performances, and the Kalash flute (ispnoe). As with the Kho participants, the Kalash participants were not familiar with any form of Western music and any form of music notation.

Materials

The stimuli consisted of four chord types varying with reference to their affective connotations (the Western distinction of major-positive versus minor-negative) and their acoustic roughness. To diversify the stimuli, two separate timbres were chosen familiar to both UK (the piano) and Northwest Pakistani (the sitar) participants, and the chords were presented with two types of delivery: they were played as both vertical chords (simultaneous presentation of pitches) and as arpeggios, as the latter is more ecologically valid with reference to Northwest Pakistani musical practice; vertical harmonic orientation is often seen as one of the major differences between Western and much of non-Western music.[42,43] Specific to the Kalash tribe (with the exception of one genre of vocal ritual music in which drone clusters, homophonic polyphony, and parallel melodic harmonies are observed), the entire corpus is based on melodic organization that makes use of minor and major pentatonic and heptatonic modes. In relation to Khowari mahfil music, similar to the stylistic organization of the music found in the region of Northwest Pakistan and eastern Afghanistan, it is also melody based, utilizing a more stylistically complex corpus of scales/modes. As such, an arpeggiated version of the chords was expected to be perhaps more relatable to both Kalash and Khow participants. The four chosen chord types were major, minor, and augmented triads, and a chromatic cluster chord. To characterize the qualities of the chosen chords, we extracted a few key variables, such as roughness,[44] harmonicity,[45] and familiarity (on the basis of a popular music corpus[23] and a custom analysis of the two tribes in Northwest Pakistan) of each chord (see Table 1). The triads were played in their root positions and contained the root doubled in the octave. All chords had identical pitch numerosity (four pitches in each chord).

Major and minor triads were chosen since these chords are highly conventionalized indices of positive and negative valence for Western listeners.[38,39] Conversely, in the most common "pop" genres of the two Northwest Pakistani tribes (Khow mahfil music; and Kalasha nonsecular flute and drum instrumental music), the minor third is significantly more common than the major third (see Supporting Information, available at http://bit.ly/3gURpvH). The Khowar music corpus was put together from commercially available cassette tapes ($N = 20$) obtained from the neighboring town of Chitral, while the Kalash music corpus was constructed from field recordings by one of the authors. The familiarity with the minor third is expected to yield positive valence for the minor triad among Northwest Pakistani listeners (contrary to the Western affective convention) as per the mere exposure effect.[46]

The augmented triad was chosen as its perceived dissonance is expected to be culture dependent: the chord's acoustic roughness value is quite similar to major and minor triads (see Table 1); however, it is consistently rated as the most dissonant of the four triads (major, minor, diminished, and augmented) among Western listeners (e.g., see Ref. 47). The perceived dissonance of the augmented triad has been linked to its intervallic equidistance[48] as the chord is made up of two consonant major thirds, and its dissonance is hence hard to pinpoint acoustically. However, a more plausible explanation is provided by the concept of *tonal dissonance* by Johnson-Laird *et al.*[25] On the basis of empirical findings, they propose that tonal dissonance depends on the scales in which chords can occur: chords occurring in a major scale (Tonal Dissonance Level 1) are less dissonant than chords occurring only in a minor scale (Tonal Dissonance Level 2), which, in turn, are less dissonant than chords occurring in neither sort of scale (Tonal Dissonance Level 3). Johnson-Laird *et al.* suggest that Western listeners—even without musical training—acquire a tacit knowledge of tonality, and violations of these tonal principles lead to the perception of additional dissonance in chords.[25] As the augmented triad is present in the (harmonic) minor scale (Tonal Dissonance Level 2) but not in the major scale, this theory offers a

credible explanation for its perceived dissonance despite the fact that it is acoustically not very rough.

The chromatic cluster's vertical voicing was chosen on the basis of an empirical experiment by Bowling *et al.* as this particular chord was perceived as the most dissonant one among all the possible tetrachords that can be formed within an octave.[40] Moreover, this chord's notably high amount of acoustic roughness has been demonstrated to be perceived as automatically negative among Western listeners in an affective priming setting,[21] making it a strong candidate for a universally unattractive sonority; a sensitivity to roughness has been demonstrated to exist across cultures.[14]

The three chord types of major, minor, and augmented were presented identically in their respective vertical and horizontal (arpeggiated) types of representation, and the fundamental frequencies (F_0s) of the pitches in each chord were adjusted so that the mean F_0 of all pitches was C_4 (261.63 Hz). Chromatic clusters, on the other hand, were delivered in slightly different versions in their respective vertical and horizontal delivery: the chord's large interval leap (a major sixth between the third and fourth pitches) in the vertical presentation was deemed as a possible confound as it would have created a disjunct melodic movement (one semitone progression of three tones with a leap of nine semitones); hence, the arguably more representative example of strict chromaticism was chosen that omits the large interval leap (see Table 1 for all chords represented in MIDI numbers). The piano stimuli were generated with *Ableton Live 9* (a music sequencer software), using the *Synthogy Ivory Grand Pianos II* plug-in. The applied sound font was *Steinway D Concert Grand*. No reverb was used, and the chords had a fixed velocity (65) for a neutral and even sound. The sitar stimuli were generated with *Logic Pro 10.5*, using the *Logic Pro Legacy Strings* plug-in. The applied sound font was *sitar*. As with the piano stimuli, no additional reverb was used. The length of both the vertical and arpeggiated (presented in an ascending and then descending order, seven distinct notes, with 250-ms inter-onset intervals) chords was 2 seconds. For both types of chord stimuli, the sound files were presented binaurally (the same signal in both channels). All stimuli were presented in equal temperament; this was the procedure applied in a similar cross-cultural study conducted by McDermott *et al.*[14] The stimuli can be found online at http://bit.ly/3gURpvH.

Procedure

Valence ratings were collected with the Self-Assessment Manikin model (SAM) by Bradley and Lang.[49] This method was developed specifically to be versatile in use by non-English speakers of different backgrounds and has successfully been used in various cross-cultural studies (e.g., see Refs. 11, 50, and 51) fulfilling its purpose (i.e., dimensional rating of emotion without the use of language). Most importantly, as in both Khow and Kalash languages, the term *valence* is untranslatable, so the word *pleasantness* was used as a synonym. In Northwest Pakistan, this was implemented by cutouts of the images, which were placed face-down and in random order in front of the participants. For UK participants, the SAM model was presented as images online through the Qualtrics platform. For validation, the participants were asked to place the valence cards in order of magnitude after the term was explained to them with the assistance of an onsite translator and in terms meaningful to the participants' own cultural norms. In both cases, valence was assessed via the pleasantness of the stimulus (how pleasant or unpleasant it is; for successful implementations of this method for collecting valence ratings, see, e.g., Refs. 52–54). After participants passed this initial assessment test, they heard all stimuli in a randomized order and proceeded to rate them with reference to valence by selecting one image from the SAM model that they thought matched the stimulus. In Northwest Pakistan, this was done by placing a small magnetic chip directly on one of the images. In the online version, UK participants indicated their preference by clicking on one image of choice. After this stage, all participants heard the paired stimuli according to their structural design (i.e., vertical chords and arpeggios) and timbre (piano and sitar). During their presentation to participants, the stimuli were grouped together in four blocks, on the basis of their structural design (arpeggio or chord) and their timbre (piano or sitar). The order of the blocks, as well as the order of the pairs within each block and the stimuli within each pair, was randomized. Within each block, each stimulus pair was presented once. Pakistani participants indicated their preference by

providing a verbal response in their native language (Kalash: Hawel, Duo; and Khowar: Hawelo, Duo), while participants in the United Kingdom indicated their preferences by clicking on their stimulus of choice. The total duration of the experiment was around 15 minutes. Following the main part of the experiment, we conducted postexperiment interviews with focus groups by members of both tribes in Northwest Pakistan. Seven Kalash and five Khow participants listened to the four chord stimuli again (minor, major, augmented, and chromatic chords) played on the piano timbre (vertical type of delivery) and were asked how well the chords matched the music of their own or one of the neighboring tribes on a scale of 1–5, where 5 would be a good match and 1 no match at all. Each chord was played twice in the session. These responses were then aggregated to create a mean familiarity value for each chord for both tribes. Additional qualitative information collected during the postexperiment interviews may be found in the Supporting Information (online only).

Results

We initially compared the internal consistency of the valence ratings within each cultural group. For UK participants, the consistency was high ($\alpha = 0.959$) and somewhat lower for the Khow ($\alpha = 0.812$) and the Kalash ($\alpha = 0.782$), albeit sufficient for further analyses. We first break down the results of the valence ratings and then move on to the preference task of paired comparison.

We subjected the ratings to a linear mixed model where we had four fixed effects (Type: arpeggio versus vertical chord; Timbre: piano versus sitar; Chord: major, minor, augmented, and chromatic; and Culture: UK, Kalash, and Khow) and one random effect (participant) allowing random intercepts for participants and specified main effects and all first-order interactions with Culture. *P* values are obtained by restricted maximum likelihood option and utilizing Satterthwaite approximations for degrees of freedom offered by lmerTest package[55] that provides robust estimations of *P* values.[56] This yielded a significant main effect of Timbre ($\beta = -0.46$, $t(1179) = -2.85$, $P = 0.0045$), where the piano (overall average valence M = 4.00) was typically perceived as more pleasant than the sitar (overall average valence M = 3.04) across all cultural groups and chord types, but the other differences in ratings were not significant, Chord ($\beta = 0.01$, $t(1179) = 0.17$, $P = 0.867$), Type

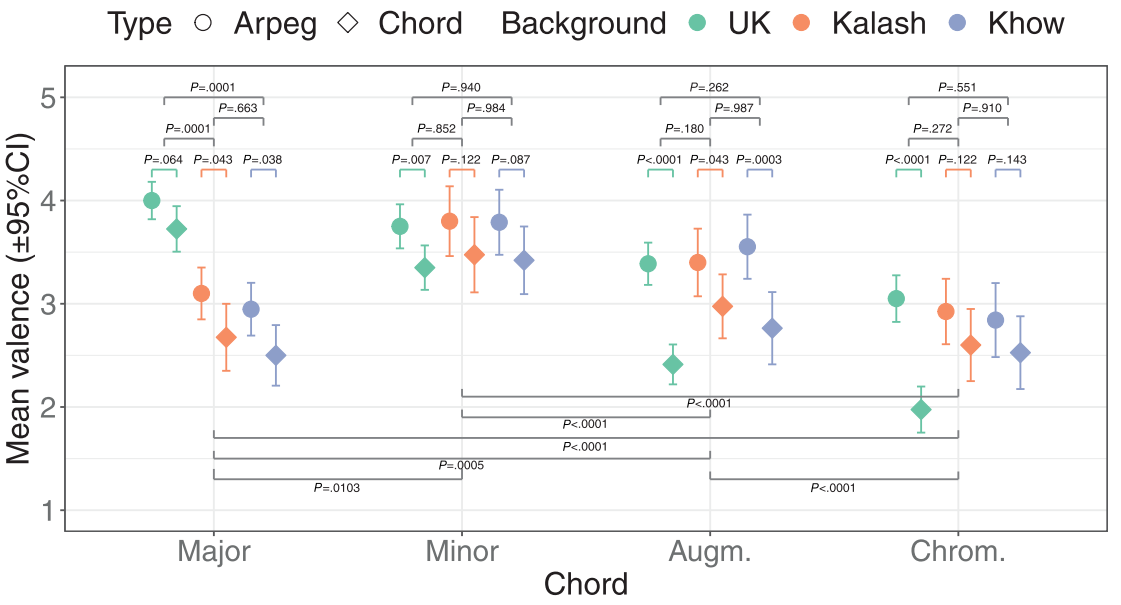

Figure 1. **Mean valence ratings across Culture and Type (arpeggio versus vertical chord) for each chord. *P* values reflect contrast analyses with GLMM with Tukey's method for adjusting multiple corrections that have been carried out between chords, Type, or Culture.**

($\beta = -0.20$, $t(1179) = -1.23$, $P = 0.217$), and Culture ($\beta = -0.13$, $t(1218) = -0.77$, $P = 0.438$). The interaction between Chord and Culture was significant ($t(1179) = 3.30$, $P = 0.00098$) as was the interaction between Type and Culture ($t(1179) = -2.32$, $P = 0.0204$). To illustrate the pattern of responses, we have plotted the means across Chord, Culture, and Type in Figure 1.

We look at the valence ratings of only major and minor chords since this is where the Western and Northwest Pakistani musical cultures have a reversed prevalence of occurrence. We look at the mean ratings collapsed across Type and Timbre. As hinted in Figure 1, there are significant differences between major and minor chords in valence ratings within all three groups. For UK listeners, major receives higher positive valence (M = 3.86) than minor (M = 3.55, $\beta = 0.31$, $t(550) = 3.05$, $P < 0.001$). For the Kalash, this pattern is reversed; major is less positive in valence (M = 2.89) than minor (M = 3.64, $\beta = -0.75$, $t(550) = -5.17$, $P < 0.001$), and a similar pattern holds for the Khow (major M = 2.72 and minor M = 3.61, $\beta = -0.88$, $t(550) = -5.92$, $P < 0.001$). The data and all analyses can be accessed at https://github.com/tuomaseerola/pakistan-chords.

In the paired comparison task, each participant chose the preferred chord between the two presented options. There was no effect of Type (arpeggio or chord) across Culture, using the likelihood ratio with conditional independence ($\chi^2(9) = 5.32$, $P = 0.805$). Also, the instruments (piano and sitar) operated nearly identically across the chords ($\chi^2(9) = 1.84$, $P = 0.993$; see Supporting Information (online only) for a full breakdown of the chord choices across instruments and the type of delivery). Moreover, there was no three-way interaction between Instrument, Type, and Culture ($\chi^2(6) = 0.97$, $P = 0.986$), so we focus on the impact of Culture where significant differences emerged ($\chi^2(6) = 277.8$, $P < 0.001$) (see Table 2). It is readily apparent that Kalash and Khow participants preferred the minor triad (34.8% and 37.5%) over other chord types, and the second most preferred option was the augmented triad (30.0% and 28.1%). Also, the chromatic cluster was not the least chosen in the task, as was the case with UK participants (2.6%), but their least preferred choice was remarkably the major chord (16.2% and 11.6% for Kalash and Khow, respectively). The preference of chords across the Kalash and Khow participants is, in fact, alike ($\chi^2(3) = 6.04$, $P = 0.11$).

Discussion

Hypothesis 1: major/minor mode

The finding that the Western affective distinction of major-positive minor-negative is reversed for Northwest Pakistani listeners is striking; however, it is in line with our original hypothesis according to which this affective contrast is indeed a cultural convention and influenced by familiarity as per the mere exposure effect, which postulates that familiarity yields positive valence.[46] This is demonstrated in the results: for UK listeners, the major triad is more positive in valence (Fig. 2) and is overall preferred more than the minor triad and other chords, and this is corroborated by the observation that in Western popular music, major triads are about four times more common than minor triads.[23,57] For Northwest Pakistani listeners, this affective pattern is reversed (Fig. 2), arguably because of the notably higher frequency of minor over major thirds in their music (85% minor, 10% major, and 5% in-between, i.e., neutral thirds). Although the culture dependence of this mode distinction has remained somewhat unclear in previous empirical settings,[8,11] the current results imply that the Western major-positive minor-negative affective distinction could indeed be a cultural convention and its possible (and contentious) acoustic origins (e.g., see Ref. 57) might be overruled by familiarity through mere exposure.[46] Hence, the current results do not corroborate the theory[4] according to which this affective distinction is based on universal vocal characteristics of different affective states.

Hypothesis 2: consonance and dissonance

For UK listeners, the valence and preference ratings go according to the predictions (ranking: 1-major; 2-minor; 3-augmented; and 4-chromatic), and this is notably in line with familiarity and tonal dissonance rankings. The acoustic predictors (roughness and harmonicity), on the other hand, yield notable differences only between the chromatic cluster and three other chord types (see Table 1). This observation is in line with the tonal dissonance theory of Johnson-Laird *et al.*,[25] according to which the augmented triad is dissonant in Western music because it is less familiar than major and minor triads (it is not present in major key tonality, only

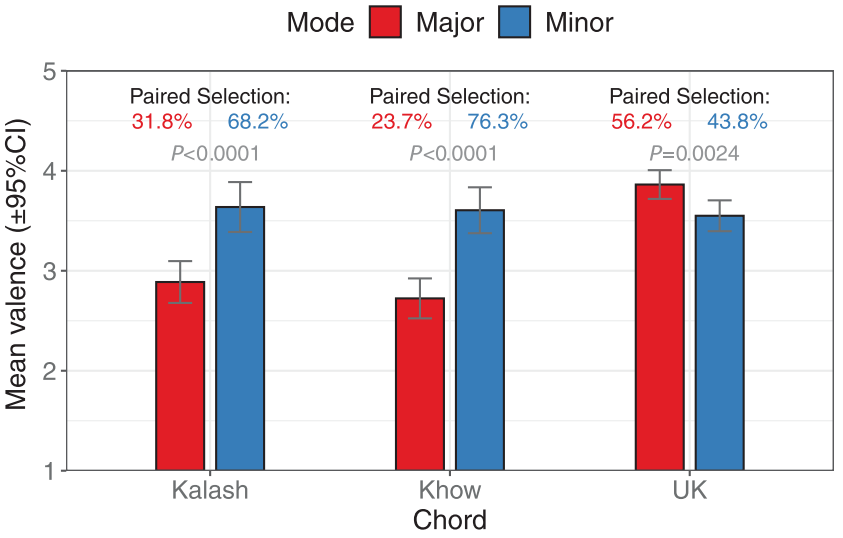

Figure 2. **Mean valence ratings for major and minor chords, as well as paired comparison preference choice task decisions for the three cultural groups.**

Table 2. Counts of the choices across the chords (collapsed across the type of delivery) and cultural background (percentages in parentheses)

	UK	Kalash	Khow
Major	399 (41.6)	78 (16.2)	53 (11.6)
Minor	311 (32.4)	167 (34.8)	171 (37.5)
Augmented	225 (23.4)	144 (30.0)	128 (28.1)
Chromatic cluster	25 (2.6)	91 (19.0)	104 (22.8)

in minor) despite a negligible difference in acoustic roughness and harmonicity between the triads (see Table 1). For Northwest Pakistani listeners, the valence and preference ratings retell the UK ranking of minor and augmented triads and the chromatic cluster chord. For Northwest Pakistani listeners, the minor triad is the most familiar (readily explaining its preference), followed by the chromatic cluster, the augmented triad, and finally, the major triad. Despite the chromatic cluster being somewhat more familiar than the augmented triad, the latter is preferred over the chromatic cluster chord. This finding corroborates the hypothesis according to which an aversion to a high amount of roughness is present cross-culturally: the acoustically smoother minor and augmented triads were clearly preferred over the maximally rough chromatic cluster, similarly as in the case of UK listeners. The most notable difference between UK and Northwest Pakistani listeners is in the perception of the major triad: for Northwest Pakistanis, it is virtually at the same level in relation to valence as the chromatic cluster and is the least liked chord. This remarkable pattern of perception is arguably not only due to its lowest familiarity among the four presented chords but also for additional cultural reasons that we will discuss next.

Resolving results conflicting with the hypotheses

The only finding that goes against the predefined hypotheses of the current study is the complete lack of preference for the consonance of the major triad among Northwest Pakistani listeners. While the negative valence of major compared with minor was an expected outcome owing to the low familiarity of the major third in the music of the Northwest Pakistani tribes, this lack of clear preference for the smoothness of the major triad over the roughness of the chromatic cluster was a surprise, in light of the previous literature, although notably in line with the findings of McDermott and colleagues.[14] In the West, the major triad has been historically put forward as a "perfect chord"[58] and as "natural" because of the major triad's closer similarity to the harmonic series as opposed to the

"artificialness" of the minor triad.[3] More recent research has been echoing this naturalistic notion: it has been suggested that the attractiveness of consonance may result from consonant chords' similarity to harmonic human vocalizations,[12,40] which would make it a cross-cultural universal. According to the current results, however, this may not be the case: not only was the major triad conspicuously unattractive for the Northwest Pakistani listeners, but the more inharmonic augmented triad (see Table 1) was also preferred over the major triad. In addition to the overall low frequency of major thirds in the music of the Northwest Pakistani tribes, post-experiment discussions involving additional concepts from a subset of the Kalash and Khow respondents (see Supporting Information, online only) make it clear that for them, the major triad is not only unattractive because of its low familiarity but also because it is specifically a "trope" for marked otherness specifically associated with the West. For example, the major triad was described as sounding "strange," as "European music," and as something that is "not our music." Accordingly, the "otherness" of the major triad for the Northwest Pakistani tribes is somewhat akin to the exoticism and otherness of the flat second degree (♭2) for Western listeners, which effectively conjures up an air of orientalism, readily used as a trope by Western composers.[59]

Timbre and the type of chord delivery

Somewhat surprisingly, the Northwest Pakistani tribes typically perceived the piano as more positive than the sitar, which is a more familiar instrument for them. However, this result is consistent with findings of acoustic cues signaling valence in past studies conducted on Western listeners—for instance, Eerola *et al.* established that lower brightness and lower subband flux lead to higher valence ratings of instrument sounds.[60] Here, the acoustic differences between the sitar and piano exhibit a similar pattern, where the piano has lower brightness ($M = 0.145$, $SD = 0.015$) compared with the sitar ($M = 0.203$, $SD = 0.038$) and lower subband flux ($M = 0.350$, $SD = 0.118$) than the sitar ($M = 0.108$, $SD = 0.042$) if we look at these two most significant predictors of valence of timbre calculated from all chords used in the experiment. It is worth noting that the piano chord stimuli were slightly louder in amplitude than the sitar stimuli, which might have partly driven this tendency; however, the chords were played in separate blocks according to timbre, so the direct comparison between the two sounds was arguably not influenced by this small amplitude difference.

With respect to the type of delivery (arpeggio versus vertical chord), the arpeggios were typically perceived as higher in valence than the vertical chords across all cultural groups (see Fig. 1). However, as the differences are overall relatively small according to the type of delivery, we interpret that the two separate types of delivery are analogous and comparable rather than the vertical type being exceptional and specific to Western music. This is notable in the light that the vertical harmonic orientation is often seen as one of the major differences between Western and much non-Western music.[42,43] In line with this observation, vertical harmonic alignment is not the prevalent norm among the Kalash and Khow tribes, nor in the region of Northwest Pakistan. However, it is present to a lesser extent in the form of heterophony between the leading voice and the leading melodic instrument (for Kho participants), and through homophony in one specific style of ritual music (for Kalash participants, in which the homophonic part is sung as chromatic drones).

Conclusions and future research

The findings of the current study empirically demonstrate that the Western affective mode distinction of major-positive minor-negative may not be a cross-cultural universal and is, in fact, reversed for Northwest Pakistani owing to an opposing distribution of major versus minor in their musical culture when compared with the West. Despite this remarkable difference with regard to the major/minor mode affective distinction, there are also cross-cultural commonalities in chord perception: the aversion to a high amount of acoustic roughness seems to be universal as both cultural groups did prefer nonrough (minor and augmented triads) over rough (chromatic cluster) sonorities, even if the major triad was perceived as unattractive because of low familiarity and marked cultural otherness among Northwest Pakistani listeners. While moderately small amounts of roughness in musical intervals (minor and major seconds) are promoted in some musical cultures for aesthetic ends (e.g., Balkan and Baltic folk music), the chromatic cluster chord arguably contains such a high amount

of roughness that the aversion to it may well be universal. This finding is also of interest in light of previous cross-cultural research on the perception of consonance/dissonance conducted by McDermott *et al.,* where a complete indifference to dissonance was found.[14] Notably, the authors of that study did not include objectively very rough/inharmonic stimuli that could have influenced these results. Of course, it could also be that the aversion to dissonance is not necessarily universal, but this remains to be investigated with a number of un-Westernized cultures with versatile stimuli that range from maximum dissonance to perfect consonance. The current study's finding of a cross-cultural aversion to maximum dissonance is, however, notably in line with previous research demonstrating an aversion to dissonance in 4-month-old infants.[15] Notably, low preference for the major triad among the Northwest Pakistani tribes is not corroborating the theory according to which the attractiveness of consonance is due to harmonic similarity to human vocalizations[12,40] and is in line with the historical observation according to which the major third became consonant only over time in the framework of Western music as well.[61,62] It is important to note, however, that the current results do not prove that there could not theoretically exist an innate preference for consonance; cultural learning can evidently override possibly biologically based inclinations, such as in the case of the acoustically rough and inharmonic Balinese gamelan music. Furthermore, it would be interesting to see, for example, whether preferences for harmonic major tonalities did emerge among Pakistanis who lacked the association of major key music with the West, and also whether consonance/dissonance and major/minor preferences change over the course of development among Pakistani youth. After all, both of these aesthetic responses seem to be culturally malleable in the West as well.[41,32,63] Also, a possible limitation to keep in mind with regard to the current study is the fact that the UK and Northwest Pakistani groups were not matched on gender and musical expertise.

Taken together, our current results imply that the aversion to a high amount of acoustic roughness in dissonant chords may be universal, but the attractiveness of highly consonant chords (the major triad) is prone to notable cultural variation. To our best knowledge, the current study is the first to contain a versatile set of stimuli in relation to consonance and dissonance (wide variation in acoustic roughness and harmonicity), and also open-ended responses from non-Western participants about what qualities they perceive in musical chords. Although we applied a direct selection task and pictorial representations to circumvent some of the semantic confounds that have been posing challenges for cross-cultural research into the perception of consonance/dissonance, we encourage future research to come up with further ways to mitigate these confounds, such as affective priming methods or neural measures (e.g., electroencephalography), to tap into automatic responses to bypass any semantic limitations.

Acknowledgments

This research was carried out with a grant from the Ella and Georg Ehrnrooth Foundation awarded to the first author, a COFUND/Marie Curie grant awarded to the second author, and with additional funds provided by Durham University's Faculty of Arts and Humanities Pro-Vice-Chancellor's award. We acknowledge the essential contribution of Mr. Taleem Khan and his team (Ekbal Shah, Haroon Khan, Saif Ul Islam, and Zahir Shah) who worked effortlessly in the field as translators and for their assistance in recruitment, and to Inspector Maikal Shahrakat for liaising with the local authorities. Special thanks go to Ishfaq Ahmed for his most valuable information regarding the Chitrali sitar and the modal system, the Kalash and Khow tribes for their participation, and to the Islamic Republic of Pakistan for providing security during fieldwork. Finally, we are grateful to the Durham University International Office for their assistance and Simone Tarsitani from Durham University's Music Department for providing technical equipment and support.

Author contributions

I.L. contributed to the experimental design, stimulus selection/production, and interpretation of the obtained data; I.L. also played the main role in writing the manuscript. G.A. collected the data and contributed to the experiment design, formal analysis, and writing. T.E. played the main role in formally analyzing the data, preparing the figures, and contributing to the writing.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1. Choices across the chords and instruments for the three groups.

Table S2. Frequency of chord choices across chord types and delivery types for the three groups.

Table S3. Roughness across timbre, chord, and the type of delivery.

Table S4. Chord familiarity summaries across the groups.

Table S5. Textual analysis from postexperiment interviews on major chord stimulus.

Table S6. Textual analysis from postexperiment interviews on minor chord stimulus.

Table S7. Textual analysis from postexperiment interviews on augmented chord stimulus.

Table S8. Textual analysis from postexperiment interviews on chromatic chord stimulus.

Figure S1. An example mode estimation process with an example (4B TRACK 4).

Competing interests

The authors declare no competing interests.

References

1. Mehr, S.A., M. Singh, D. Knox, *et al.* 2019. Universality and diversity in human song. *Science* **366:** eaax0868.
2. Jacoby, N., E.H. Margulis, M. Clayton, *et al.* 2020. Cross-cultural work in music cognition. *Music Percept.* **37:** 185–195.
3. Schenker, H. 1906. *Harmonielehre (Neue musikalische Theorien und Phantasien*. Stuttgart: J. G. Cotta'sche Buchhandlung Nachfolger.
4. Bowling, D.L., J. Sundararajan, S. Han, *et al.* 2012. Expression of emotion in Eastern and Western music mirrors vocalization. *PLoS One* **7:** e31942.
5. Curtis, M.E. & J.J. Bharucha. 2010. The minor third communicates sadness in speech, mirroring its use in music. *Emotion*, **10:** 335–348.
6. Huron, D., G. Yim & P. Chordia. 2010. The effect of pitch exposure on sadness judgments: an association between sadness and lower than normal pitch. In Proceedings of the 11th International Conference on Music Perception and Cognition. 63–66. Seattle, WA: Causal Productions.
7. Friedman, R.S. 2018. Reexploring the effects of relative pitch cues on perceived sadness in an unconventionally tuned musical scale. *Psychomusicol: Music Mind Brain* **28:** 108–116.
8. Fritz, T., S. Jentschke, N. Gosselin, *et al.* 2009. Universal recognition of three basic emotions in music. *Curr. Biol.* **19:** 573–576.
9. Post, O. & D. Huron. 2009. Music in minor modes is slower (except in the Romantic period). *Empir. Musicol. Rev.* **4:** 1–9.
10. Horn, K. & D. Huron. 2015. On the changing use of the major and minor modes 1750–1900. *Music Theory Online* **21:** 11.
11. Fang, L., J. Shang & N. Chen. 2017. Perception of Western musical modes: a Chinese study. *Front. Psychol.* **8:** 1905.
12. Bowling, D.L. & D. Purves. 2015. A biological rationale for musical consonance. *Proc. Natl. Acad. Sci. USA* **112:** 11155–11160.
13. Lundin, R.W. 1947. Toward a cultural theory of consonance. *J. Psychol.* **23:** 45–49.
14. McDermott, J.H., A.F. Schultz, E.A. Undurraga, *et al.* 2016. Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature* **535:** 547–550.
15. Zentner, M.R. & J. Kagan. 1998. Infants' perception of consonance and dissonance in music. *Infant Behav. Dev.* **21:** 483–492.
16. Trainor, L.J., C.D. Tsang & V.H. Cheung. 2002. Preference for sensory consonance in 2-and 4-month-old infants. *Music Percept.* **20:** 187–194.
17. Plantinga, J. & S.E. Trehub. 2014. Revisiting the innate preference for consonance. *J. Exp. Psychol. Hum. Percept. Perform.* **40:** 40–49.
18. Butler, J.W. & P.G. Daston. 1968. Musical consonance as musical preference: a cross-cultural study. *J. Gen. Psychol.* **79:** 129–142.
19. Maher, T.F. 1976. "Need for resolution" ratings for harmonic musical intervals: a comparison between Indians and Canadians. *J. Cross-Cult Psychol.* **7:** 259–276.
20. Prete, G., M. Fabri, N. Foschi, *et al.* 2015. The "consonance effect" and the hemispheres: a study on a split-brain patient. *Lateral.: Asymm. Body Brain Cogn.* **20:** 257–269.
21. Lahdelma, I., J. Armitage & T. Eerola. 2020. Affective priming with musical chords is influenced by pitch numerosity. *Music. Sci*. https://doi.org/10.1177/1029864920911127
22. Bowling, D.L., M. Hoeschele, K.Z. Gill & W.T. Fitch. 2017. The nature and nurture of musical consonance. *Music Percept*. **35:** 118–121.
23. Harrison, P. & M. Pearce. 2020. Simultaneous consonance in music perception and composition. *Psychol. Rev.* **127:** 216–244.
24. Hutchinson, W. & L. Knopoff. 1978. The acoustic component of Western consonance. *Interface* **7:** 1–29.
25. Johnson-Laird, P.N., O.E. Kang & Y.C. Leong. 2012. On musical dissonance. *Music Percept.* **30:** 19–35.
26. Ambrazevičius, R. 2017. Dissonance/roughness and tonality perception in Lithuanian traditional Schwebungsdiaphonie. *J. Interdisc. Music Stud.* **8:** 39–53.
27. Messner, G.F. 1981. The two-part vocal style on Baluan Island Manus province, Papua New Guinea. *Ethnomusicology* **25:** 433–446.
28. Parncutt, R. 1989. *Harmony: A Psychoacoustical Approach*. Berlin: Springer-Verlag.
29. McDermott, J.H., A.J. Lehr & A.J. Oxenham. 2010. Individual differences reveal the basis of consonance. *Curr. Biol.* **20:** 1035–1041.

30. Cazden, N. 1980. The definition of consonance and dissonance. *Int. Rev. Aesthet. Sociol. Music* **2:** 123–168.
31. McLachlan, N., D. Marco, M. Light, *et al.* 2013. Consonance and pitch. *J. Exp. Psychol. Gen.* **142:** 1142–1158.
32. Lahdelma, I. & T. Eerola. 2020. Cultural familiarity and musical expertise impact the pleasantness of consonance/dissonance but not its perceived tension. *Sci. Rep.* **10:** 8693.
33. Smit, E.A., A.J. Milne, R.T. Dean & G. Weidemann. 2019. Perception of affect in unfamiliar musical chords. *PLoS One* **14:** e0218570.
34. Friedman, R.S., D.A. Kowalewski, D.T. Vuvan & W.T. Neill. 2021. Consonance preferences within an unconventional tuning system. *Music Percept.* **38:** 313–330.
35. Zatorre, R.J. & A.R. Halpern. 1979. Identification, discrimination, and selective adaptation of simultaneous musical intervals. *Percept. Psychophys.* **26:** 384–395.
36. Rakowski, A. 1990. Intonation variants of musical intervals in isolation and in musical contexts. *Psychol. Music* **18:** 60–72.
37. Athanasopoulos, G., T. Eerola, I. Lahdelma, *et al.* 2021. Harmonic organization conveys both universal and culture-specific cues for emotional expression in music. *PLoS One* **16:** 1–17.
38. Pallesen, K.J., E. Brattico, C. Bailey, *et al.* 2005. Emotion processing of major, minor, and dissonant chords: a functional magnetic resonance imaging study. *Ann. N.Y. Acad. Sci.* **1060:** 450–453.
39. Lahdelma, I. & T. Eerola. 2016. Single chords convey distinct emotional qualities to both naive and expert listeners. *Psychol. Music* **44:** 37–54.
40. Bowling, D.L., D. Purves & K.Z. Gill. 2018. Vocal similarity predicts the relative attraction of musical chords. *Proc. Natl. Acad. Sci. USA* **115:** 216–221.
41. Dalla Bella, S., I. Peretz, L. Rousseau, *et al.* 2001. A developmental study of the affective value of tempo and mode in music. *Cognition* **80:** B1–B10.
42. Malm, W.P. 1996. *Music Cultures of the Pacific, the Near East, and Asia*. Pearson College Division.
43. Tymoczko, D. 2010. *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice*. Oxford University Press.
44. Wang, Y., G. Shen, H. Guo, *et al.* 2013. Roughness modelling based on human auditory perception for sound quality evaluation of vehicle interior noise. *J. Sound Vib.* **332:** 3893–3904.
45. Harrison, P. & M. Pearce. 2018. An energy-based generative sequence model for testing sensory theories of Western harmony. In Proceedings of the 19th International Society for Music Information Retrieval Conference. 160–167. Paris, France.
46. Zajonc, R. 2001. Mere exposure: a gateway to the subliminal. *Curr. Dir. Psychol. Sci.* **10:** 224–228.
47. Roberts, L. 1986. Consonance judgements of musical chords by musicians and untrained listeners. *Acustica* **62:** 163–171.
48. Meyer, L.B. 1956. *Emotion and Meaning in Music*. Chicago, IL: Chicago University Press.
49. Bradley, M.M. & P.J. Lang. 1994. Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* **25:** 49–59.
50. Morris, J.D. 1995. Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. *J. Advert. Res.* **35:** 63–68.
51. Redondo, J., I. Fraga, I. Padrón & A. Piñeiro. 2008. Affective ratings of sound stimuli. *Behav. Res. Methods* **40:** 784–790.
52. Gomez, P. & B. Danuser. 2004. Affective and physiological responses to environmental noises and music. *Int. J. Psychophysiol.* **53:** 91–103.
53. Witvliet, C.V. & S.R. Vrana. 2007. Play it again Sam: repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cogn. Emot.* **21:** 3–25.
54. Marin, M.M., B. Gingras & J. Bhattacharya. 2012. Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain. *Emotion* **12:** 618.
55. Kuznetsova, A., P.B. Brockhoff, R.H. Christensen, *et al*. 2017. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82:** 1–26.
56. Luke, S.G. 2017. Evaluating significance in linear mixed-effects models in R. *Behav. Res. Methods* **49:** 1494–1502.
57. Parncutt, R. 2014. The emotional connotations of major versus minor tonality: one or more origins? *Music Sci.* **18:** 324–353.
58. Rameau, J.-P. 1971. *Treatise on Harmony*. New York: Dover Publications. (Original work published 1722).
59. Moore, S. 2014. The other leading note: a comparative study of the flat second pitch degree in North Indian classical, Ottoman or Arabian influenced, Western, heavy metal and film musics. Doctoral dissertation. University of Sheffield.
60. Eerola, T., R. Ferrer & V. Alluri. 2012. Timbre and affect dimensions: evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds. *Music Percept.* **30:** 49–70.
61. Hindemith, P. 1942. *The Craft of Musical Composition*. New York: Belwin-Mills.
62. Tenney, J. 1988. *A History of 'Consonance' and 'Dissonance'*. New York: Excelsior Music Publishing Company.
63. Weiss, M.W., L.K. Cirelli, J.H. McDermott & S.E. Trehub. 2020. Development of consonance preferences in Western listeners. *J. Exp. Psychol. Gen.* **149:** 634–649.