Current Biology, Volume *19*

Supplemental Data

Universal Recognition

of Three Basic Emotions in Music

Thomas Fritz, Sebastian Jentschke, Nathalie Gosselin, Daniela Sammler, Isabelle Peretz, Robert Turner, Angela D. Friederici, and Stefan Koelsch

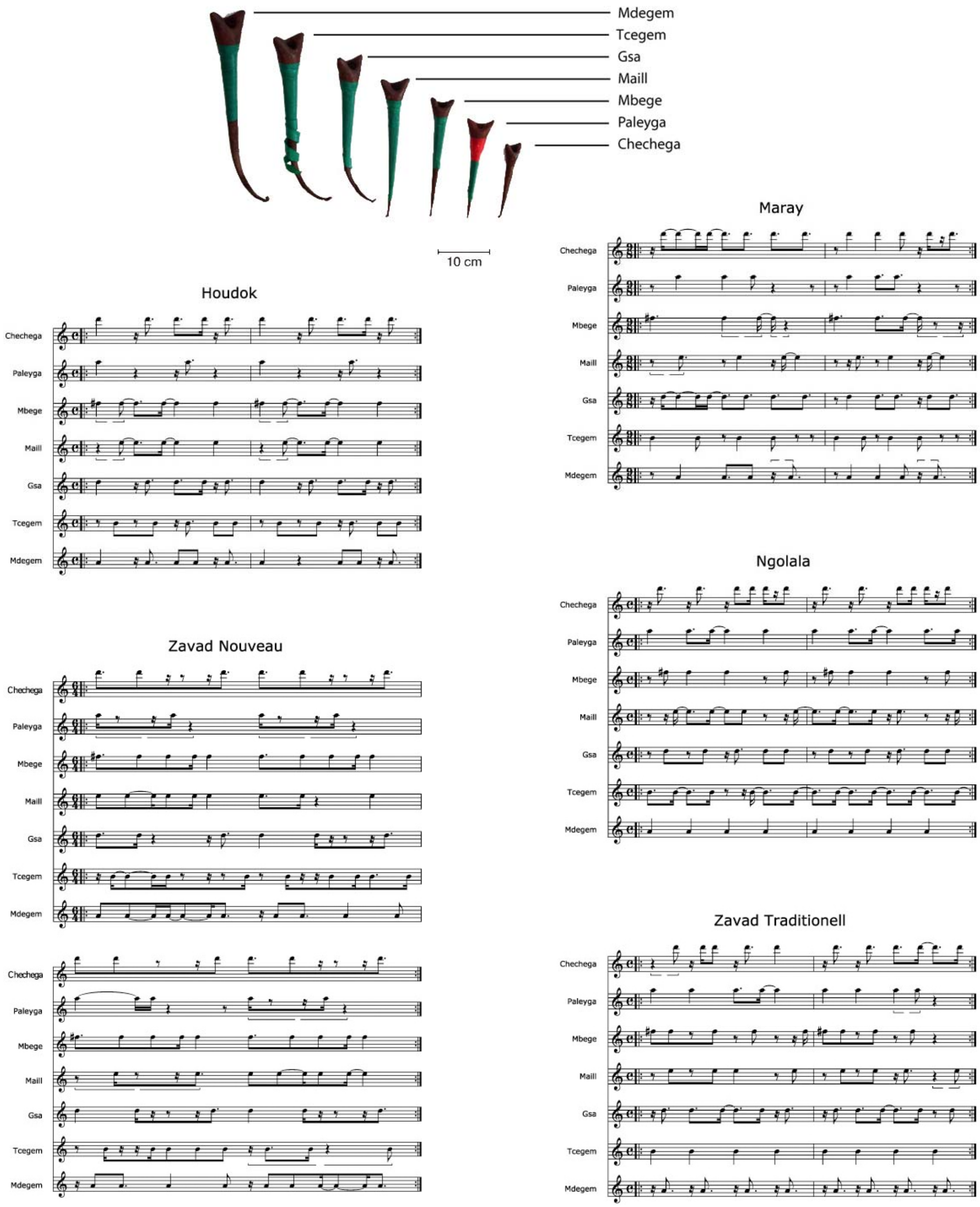


Figure S1. Scores of the recorded Mafa music played with the instruments depicted above and in Figure 1, with each pattern presented twice (Mafa players perform these patterns repetitively). Making flute music is usually mandatorily associated with rituals, and the music performed at these occasions does not have a designation, but is closely associated with the corresponding ritual. The playing of the flute is accompanied by dancing or running of the players during strenuous flute playing involving rapid exhalation, and a good physical fitness is regarded as a precondition to engaging in this activity which can last for several hours. The culturally imprinted formal and structural dispositions to music perception, as established in many Western concert settings do not apply to the Mafa listeners, as they are not familiar with the concept of professional musicianship and the musical rituals that they engage in demand general active participation.

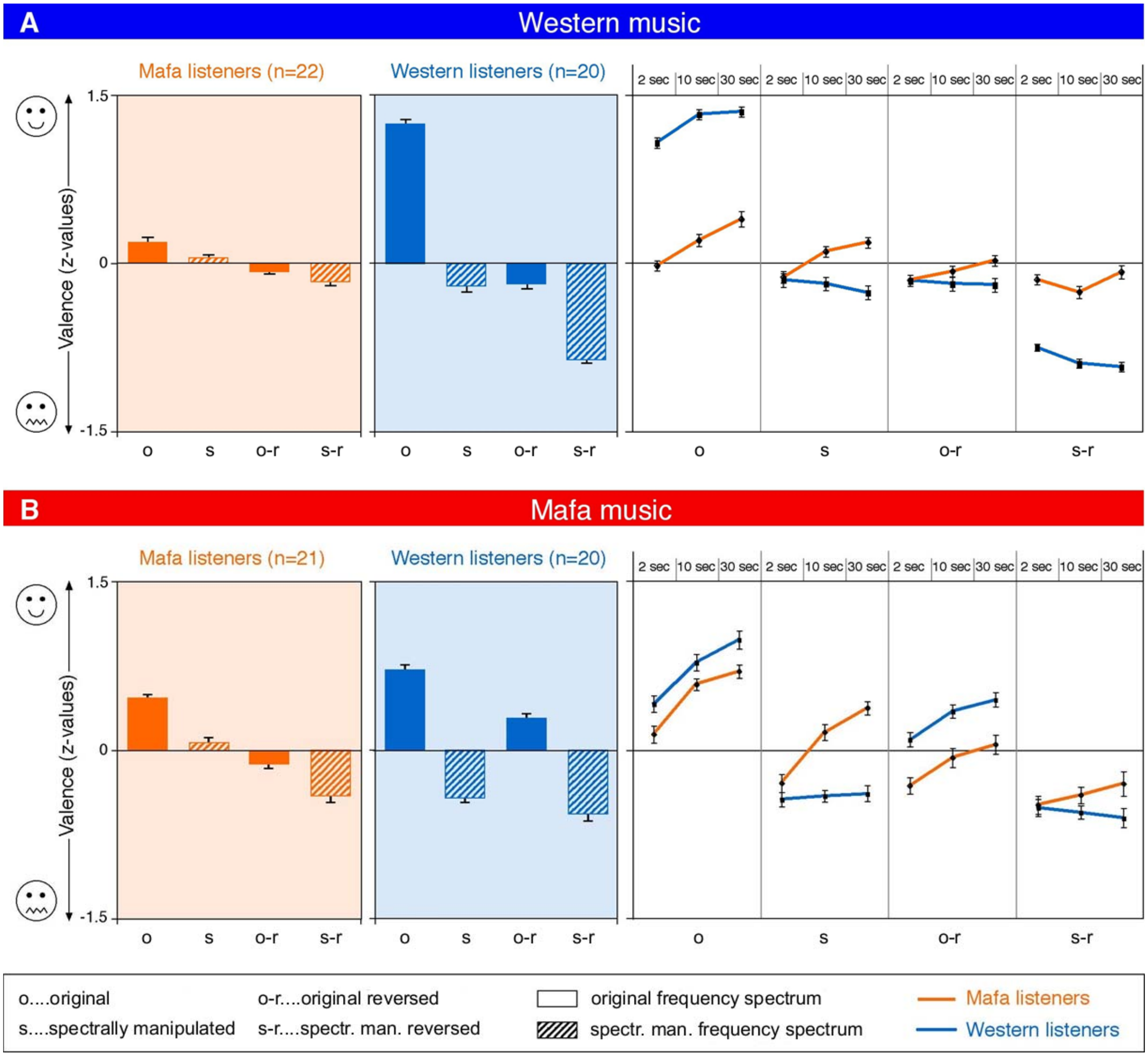


Figure S2. The figure shows *z*-values of the ratings from Experiment 2, which were standardized in relation to the mean value and the standard deviation of each individual to exclude individual assessment tendencies (see Supplementary Table 4 for statistical evaluation, Supplementary Figure 3 for non-standardized values). Mean values are depicted for each entire stimulus category (o, s, o-r, s-r), and each stimulus duration (2, 10, 30 sec) respectively. Error bars indicate SEM (for values see Supplementary Table 1).

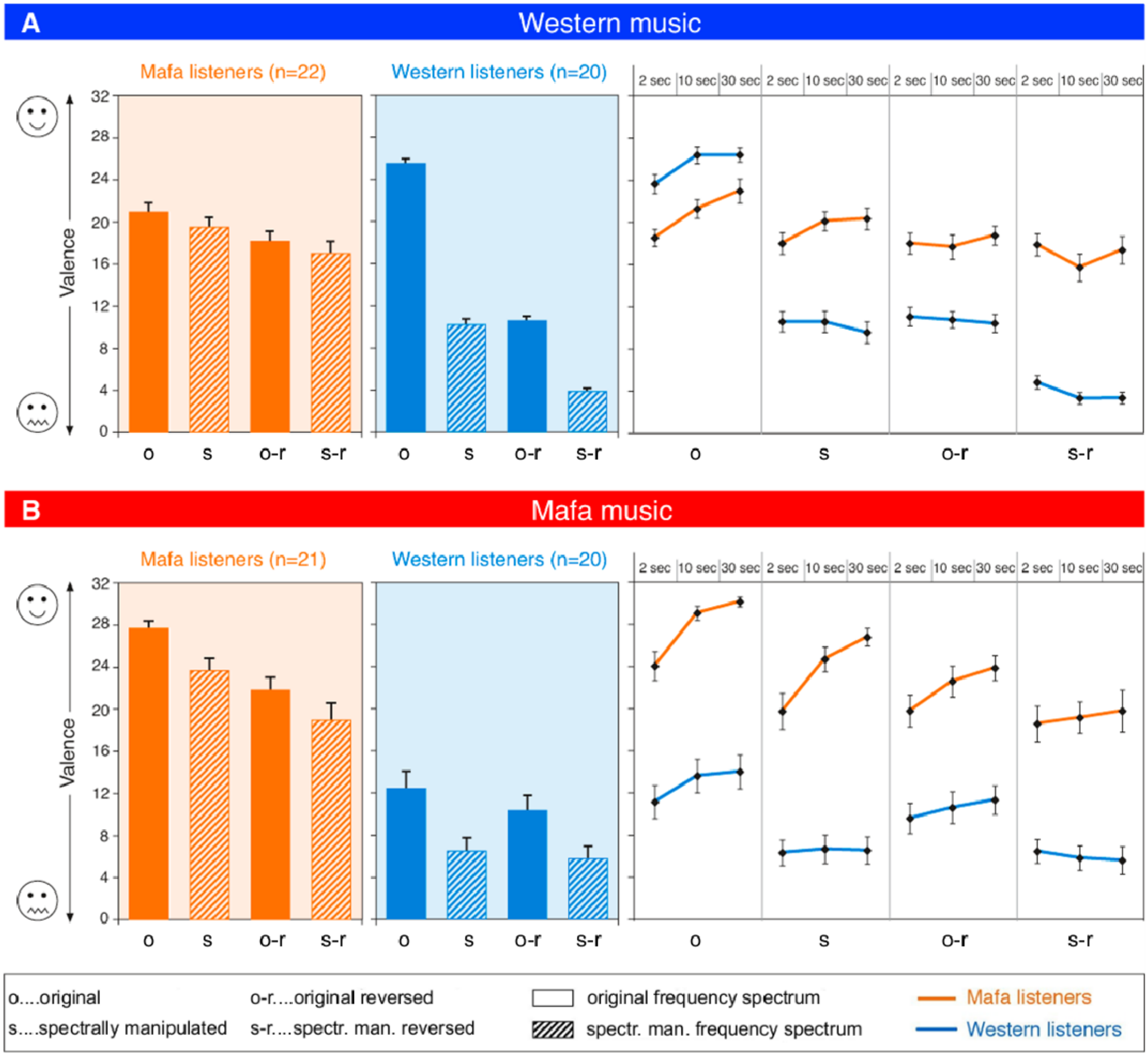


Figure S3. The figure shows non-transformed values of the ratings conducted in Experiment 2. Mean values are depicted for each entire stimulus category (o, s, o-r, s-r), and each stimulus duration (2, 10, 30 sec) respectively. Error bars indicate *SEM.*

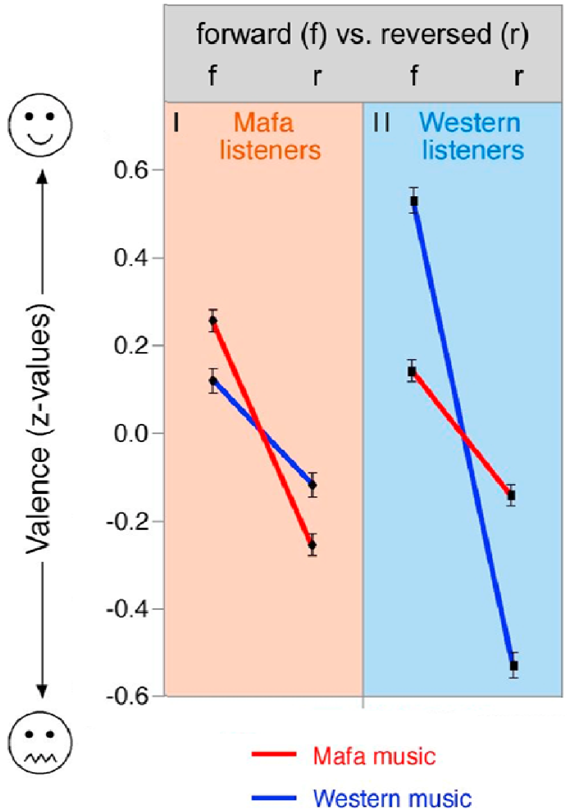
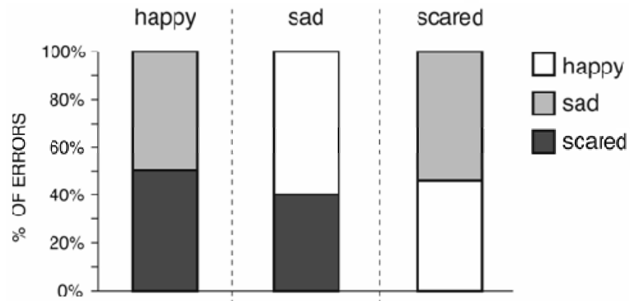


Figure S4. Experiment 2. Mean values of valence ratings for the main effects of *direction* (error bars indicate SEM), separately for Mafas (light red panel) and Westerners (light blue panel). Each panel shows the mean ratings for excerpts played in a forward direction (*f*; pieces with original and manipulated spectrum pooled) and pieces played backwards (*r*; pieces with original and manipulated spectrum pooled). Ratings for Mafa music are connected by red lines, ratings for Western music by blue lines.



	sad > scared	happy>scared	happy > sad
mean	-0.32%	10.06%	-4.22%
std. error	3.92%	6.33%	5.61%
t(21)	0.08	1.59	-0.75
p	0.935	0.127	0.460

Figure S5. Mafa error rates (i.e. non-target choices) in percentage for the recognition of each emotional expression. The table shows mean differences in percentages of choices of non-target categories (*M*), standard error (*SEM*), *t*-values, and *p*-values. Note that the error rates of the Mafa participants show that for none of the three categories 'happy', 'sad' and 'scared' a significant error tendency for another emotional expression was observed (the error rate of the Western listeners was too low to reveal significant error tendencies due to a floor effect). Instead, the expression was attributed to another emotion on a random basis. This speaks against the possibility that the Mafa were conceptually confusing certain emotional expressions, which would have resulted in a more systematic pattern of errors, such as confusing sad and scared if they made their decisions based on major/minor or tempo.

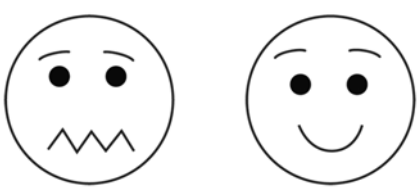


Figure S6. The figure shows the Self-Assessment-Smileys at the poles of the valence dimension as depicted on the left (unpleasant) and right (pleasant) of the slider on the rating interface. During the familiarization with their task the depictions of the smileys were paired with a corresponding smile or frown of the experimenter. The Mafa easily understood these associations.

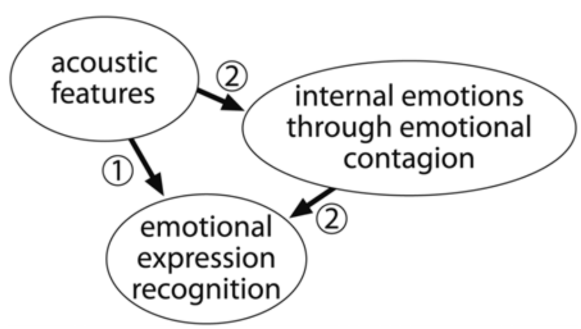


Figure S7. Depiction of two mechanisms by which emotional expression recognition may occur.

		Mafa music				Western music			
	category	c	d	r-c	r-d	c	d	r-c	r-d
Mafa listeners	value	0.46	0.07	-0.12	-0.41	0.19	0.05	-0.07	-0.16
	SEM	0.03	0.04	0.04	0.06	0.04	0.02	0.02	0.03
Western listeners	value	0.71	-0.42	0.28	-0.57	1.24	-0.20	-0.18	-0.86
	SEM	0.05	0.04	0.04	0.06	0.04	0.06	0.05	0.03

c ....consonant d ....dissonant r-c....reversed consonant r-d ...reversed dissonant

Table S1. *Z*-values depicted in Supplementary Figure 2.

General Linear Model

	Mafa music		Western music	
direction	$F_{(1,39)} = 128.88$	$p < 0.001$	$F_{(1,40)} = 259.21$	$p < 0.001$
direction × subgroup	$F_{(1,39)} = 10.47$	$p = 0.002$	$F_{(1,40)} = 103.46$	$p < 0.001$
spectrum	$F_{(1,39)} = 190.39$	$p < 0.001$	$F_{(1,40)} = 261.78$	$p < 0.001$
spectrum × subgroup	$F_{(1,39)} = 44.93$	$p < 0.001$	$F_{(1,40)} = 175.00$	$p < 0.001$
length	$F_{(2,38)} = 19.60$	$p < 0.001$	$F_{(2,39)} = 2.54$	$p = 0.092$
length × subgroup	$F_{(2,38)} = 2.44$	$p = 0.101$	$F_{(2,39)} = 4.89$	$p = 0.013$
direction × spectrum	$F_{(1,39)} = 10.54$	$p = 0.002$	$F_{(1,40)} = 40.29$	$p < 0.001$
direction × spectrum × subgroup	$F_{(1,39)} = 2.30$	$p = 0.138$	$F_{(1,40)} = 37.63$	$p < 0.001$
direction × length	$F_{(2,38)} = 7.00$	$p = 0.003$	$F_{(2,39)} = 12.41$	$p < 0.001$
direction × length × subgroup	$F_{(2,38)} = 1.30$	$p = 0.286$	$F_{(2,39)} = 0.55$	$p = 0.583$
spectrum × length	$F_{(2,38)} = 8.50$	$p = 0.001$	$F_{(2,39)} = 11.02$	$p < 0.001$
spectrum × length × subgroup	$F_{(2,38)} = 6.55$	$p = 0.004$	$F_{(2,39)} = 0.90$	$p = 0.416$
direction × spectrum × length	$F_{(2,38)} = 0.21$	$p = 0.812$	$F_{(2,39)} = 2.38$	$p = 0.106$
direction × spectrum × length × subgroup	$F_{(2,38)} = 0.69$	$p = 0.507$	$F_{(2,39)} = 1.72$	$p = 0.193$

Factors: direction (forward vs. reversed)................(c + d) vs. (r-c + r-d) length..........2 vs. 10 vs. 30 sec
spectrum (consonant vs. dissonant)..... (c + r-c) vs. (d + r-d) subgroup.....Mafa vs. Western listeners

Table S2. Overview of the results from the MANOVA calculated with *z*-values for Mafa and Western music; significant main effects are printed in black.

General Linear Model

	Mafa music		Western music	
subgroup	$F_{(1,39)} = 71.48$	$p < 0.001$	$F_{(1,40)} = 30.54$	$p < 0.001$
direction	$F_{(1,39)} = 58.11$	$p < 0.001$	$F_{(1,40)} = 157.85$	$p < 0.001$
direction × subgroup	$F_{(1,39)} = 17.67$	$p < 0.001$	$F_{(1,40)} = 54.41$	$p < 0.001$
spectrum	$F_{(1,39)} = 71.56$	$p < 0.001$	$F_{(1,40)} = 165.91$	$p < 0.001$
spectrum × subgroup	$F_{(1,39)} = 3.63$	$p = 0.064$	$F_{(1,40)} = 97.56$	$p < 0.001$
length	$F_{(2,38)} = 10.18$	$p < 0.001$	$F_{(2,39)} = 2.11$	$p = 0.134$
length × subgroup	$F_{(2,38)} = 4.36$	$p = 0.020$	$F_{(2,39)} = 5.25$	$p = 0.010$
direction × spectrum	$F_{(1,39)} = 5.65$	$p = 0.023$	$F_{(1,40)} = 31.51$	$p < 0.001$
direction × spectrum × subgroup	$F_{(1,39)} = 0.38$	$p = 0.542$	$F_{(1,40)} = 27.49$	$p < 0.001$
direction × length	$F_{(2,38)} = 8.84$	$p = 0.001$	$F_{(2,39)} = 11.17$	$p < 0.001$
direction × length × subgroup	$F_{(2,38)} = 2.25$	$p = 0.120$	$F_{(2,39)} = 0.58$	$p = 0.567$
spectrum × length	$F_{(2,38)} = 6.00$	$p = 0.005$	$F_{(2,39)} = 9.87$	$p < 0.001$
spectrum × length × subgroup	$F_{(2,38)} = 1.00$	$p = 0.379$	$F_{(2,39)} = 0.52$	$p = 0.598$
direction × spectrum × length	$F_{(2,38)} = 1.74$	$p = 0.190$	$F_{(2,39)} = 3.85$	$p = 0.030$
direction × spectrum × length × subgroup	$F_{(2,38)} = 2.03$	$p = 0.146$	$F_{(2,39)} = 2.58$	$p = 0.089$

Factors: direction (forward vs. reversed)................(c + d) vs. (r-c + r-d) length..........2 vs. 10 vs. 30 sec
spectrum (consonant vs. dissonant)..... (c + r-c) vs. (d + r-d) subgroup.....Mafa vs. Western listeners

Table S3. Overview of the results from the MANOVA calculated with non-transformed values for Mafa and Western music; significant main effects are printed in black.

	Mafa listeners				Western listeners			
	Mafa music		Western music		Mafa music		Western music	
direction	$F_{(1,20)} = 41.29$	$p < 0.001$	$F_{(1,21)} = 19.32$	$p < 0.001$	$F_{(1,19)} = 25.35$	$p < 0.001$	$F_{(1,19)} = 313.85$	$p < 0.001$
spectrum	$F_{(1,20)} = 27.13$	$p < 0.001$	$F_{(1,21)} = 6.14$	$p = 0.022$	$F_{(1,19)} = 43.75$	$p < 0.001$	$F_{(1,19)} = 322.01$	$p < 0.001$
length	$F_{(2,19)} = 7.56$	$p = 0.004$	$F_{(2,20)} = 6.48$	$p = 0.007$	$F_{(2,18)} = 4.96$	$p = 0.019$	$F_{(2,18)} = 0.47$	$p = 0.633$
direction × spectrum	$F_{(1,20)} = 1.00$	$p = 0.329$	$F_{(1,21)} = 0.06$	$p = 0.806$	$F_{(1,19)} = 11.50$	$p = 0.003$	$F_{(1,19)} = 44.27$	$p < 0.001$
direction × length	$F_{(2,19)} = 6.35$	$p = 0.008$	$F_{(2,20)} = 5.65$	$p = 0.011$	$F_{(2,18)} = 2.89$	$p = 0.082$	$F_{(2,18)} = 11.79$	$p = 0.001$
spectrum × length	$F_{(2,19)} = 1.17$	$p = 0.332$	$F_{(2,20)} = 2.84$	$p = 0.082$	$F_{(2,18)} = 5.77$	$p = 0.012$	$F_{(2,18)} = 10.27$	$p = 0.001$
direction × spectrum × length	$F_{(2,19)} = 2.57$	$p = 0.103$	$F_{(2,20)} = 0.80$	$p = 0.461$	$F_{(2,18)} = 0.25$	$p = 0.785$	$F_{(2,18)} = 3.17$	$p = 0.066$

Table S4. MANOVAs with the within-subject factors *direction* (forward vs. reversed), *spectrum* (original vs. spectrally manipulated) and *length* (2, 10, and 30 sec) that were computed separately for each subgroup and each musical culture; standardized values, significant main effects are printed in black.

		happy	sad	scared
Mafa	happy	**64.61%**	24.03%	30.52%
	sad	17.86%	**47.73%**	20.45%
	scared	17.53%	28.25%	**49.03%**
West.	happy	**98.93%**	4.29%	5.36%
	sad	0.36%	**81.43%**	2.14%
	scared	0.71%	14.29%	**92.50%**

Table S5. Confusion matrix with the percentage of responses for each category, separately for Mafa (upper part) and for Western listeners (bottom part). The classification of the pieces as happy, sad, or scared is given in the columns, and the actual responses of the particpants in the rows. For the sad and the scared pieces the ranking of the misclassifications is similar. That is in both groups sad pieces are misclassified as scared rather than as happy; and scared pieces are misclassified as happy rather than as sad.

		Major	Indefinite	Minor	Tempo
Mafa	happy	**64.61%**	13.64%	28.91%	**0.570**
	sad	17.86%	**59.09%**	31.09%	-0.076
	scared	17.53%	27.27%	**40.00%**	**-0.625**
West.	happy	**98.93%**	0.00%	5.40%	**0.675**
	sad	0.36%	**86.67%**	36.40%	0.076
	scared	0.71%	13.33%	**58.20%**	**-0.800**

Table S6. Percentage of responses for each category according to the mode of the piece (major, indefinite tonality or minor; first columns) and correlations of the piece's tempo with the number of responses for each category (last column). Values are given separately for Mafa (upper part) and for Western listeners (bottom part). The emotional category that received the most responses (first columns) as well as the significant correlations with the tempo of the pieces (last column) are written in bold.

- Supplemental Results and Discussion

Intercultural comparisons and globalization

Unfortunately, opportunities for intercultural comparisons between individuals exposed to completely incongruent music cultures are becoming increasingly rare, due to globalization. Western music culture mainly spreads with electricity supply (and thus the possibility to operate radios) and Christianization (through Western Christian song). Since the efforts of early pioneers such as Erich M. von Hornbostel at the beginning of the last century, the present study is the first to empirically investigate human individuals who were completely naïve to Western music.

The study of musical universals in infants

Besides ethnomusicological investigation, developmental study is another important means for the investigation of music universals. Similar to ethnomusicological investigation this methodology employs strategic participant selection. By monitoring orienting reactions to musical stimuli varying in their consonance/dissonance, several researchers have accumulated evidence that may suggest a preference for consonance over dissonance in Western infants [1-4]. However, these studies have the methodological constraint that the fetus is sensitive to music stimulation (given that music perception begins in utero during the third trimester of pregnancy) [5]. Hence, preference in infants could have thus been influenced by mere exposure [6, 7] or a form of evaluative conditioning in the womb where the infant associates its auditory perception with bodily states of the mother [8]. Insights about music universals can also arise from comparative animal research [9], but evidence is as of yet sparse because methodological constraints are a great challenge, and it remains elusive how these results are applicable to humans [7].

Experiment 1: Distribution of variables

In the data from Experiment 1 (the emotional expression recognition experiment) all (but two) variables conformed to a standard normal distribution (as tested by means of Kolmogorov-Smirnov tests). The deviations resulted from ceiling effects: Western listeners recognized happy and sad pieces in almost all cases (98.93 and 92.50 percent correct recognition).

Experiment 1: Performance differences between groups

In this experiment, Western listeners recognized the intended emotional expressions of the music more accurately (Figure 1A). This corresponds to previous data showing that music listeners perform better on same-culture than on other-culture materials [10, 11]. Thse results could either reflect the fact that the Westerners discerned a set of culturally determined cues not perceived by the Mafa listeners or simply the unfamiliarity of Mafas with the abstract nature of the face and music presentations without a social context of music production and perception (e.g. headphones) which is not customary in Mafa culture.

Experiment 1: Acoustic and musical analysis of stimuli

To understand whether the excerpts in Experiment 1 may have been systematically classified based on their acoustic features, we calculated correlation and contingency analyses. Actual tempo values were used for the correlation with the acoustic features. In the Mafa we observed that pieces with higher tempo are more likely to be classified as happy by a large number of participants ($r_h = 0.570$, $p = 0.003$) and scared/fearful with lower tempo ($r_f = -0.625$, $p < 0.001$), whereas for sad pieces no correlation with tempo was observed ($r_s = -0.076$, $p = 0.368$). Interestingly, the Westerners showed the same profile with slightly stronger correlations ($r_h = 0.675$, $p < 0.001$; $r_s = 0.076$, $p = 0.375$; $r_f = -0.800$, $p < 0.001$). Hence, both groups seem to have used tempo cues in a similar manner to determine the emotional content of a piece.

Similarly, we explored the influence of the mode on the classification of the emotional connotation of a particular piece. We observed that the majority of major pieces are classified as happy, the majority of pieces with indefinite mode as sad and most of the pieces in minor as scared/fearful (see Supplementary Table 6). Again, in both groups the general response pattern was found to be similar. To test whether the mode has a significant influence on the classification of the pieces, we used a chi-square-test (calculated with the total number of pieces in major, indefinite, or minor mode that were rated either as happy, sad or scared, separately for Mafa and Western listeners). It revealed that the frequency with which the participants classified a piece into a category varied significantly with the mode of the piece in both groups, even though this response pattern was more marked in the Western listeners ($\chi^2_{(4)} = 794.781$, $p < 0.001$) compared to the Mafa listeners ($\chi^2_{(4)} = 142.467$, $p < 0.001$).

Experiment 1: Discussion of mechanisms underlying the recognition of emotional expression in music

Two possible processes may account for the successful recognition of an emotional expression in music (see Supplementary Figure 7). One process by which a listener can recognize an emotional expression from the music appears to occur at the level of the acoustic features of the musical phrase, such as tempo and loudness (depicted as process 1 in the model in Supplementary Figure 7). A second process by which a listener may recognize an emotional expression in the music is by a process of emotional contagion, by which for example a musical phrase expressing sadness makes the listener feel sad, and a musical phrase expressing happiness makes the listener feel happy. The listener then categorizes the emotional expression of the music in relation to his own emotional experience.

Both processes can be labelled emotional expression recognition, whereas the recognition process involving emotional contagion involves a stronger engagement of the listener during the recognition process. The correlation analyses performed show that in both Mafa and Western listeners, tempo and mode influence the emotional expression categorization. However, our data does not show how much the second process comes into play (this is due to the fact that the methodology to gain experimental evidence of internal emotional states has not yet been developed).

Experiment 1: An argument against music as a universal language of emotions

Although some of the data aquired in Experiment 1 may be interpreted as corroborating the idea of music as a medium to universally mediate emotion, a possible absence of a variety of emotional expressions in Mafa music (as reported in the section Experimental Procedures, Stimuli) suggests a different interpretation. If music were in its essence indeed a universal language of emotions, why does Mafa music seems to not express a comparable variety of emotions as Western music? The appropriate answer to this is that although emotional expressions in music are perceived universally, this may not be the principal function of music (as already pointed out by Hanslick in his 1854 essay [12]). Despite the observed universals of emotional expression recognition one should thus be cautious to conjure the idea of music as a universal language of emotion, which is partly a legacy of the period of romanticism.

Experiment 2: Sensory and musical dissonance

In the literature two types of dissonance are distinguished: the first, sensory or psychoacoustical, arising from inharmonicity in the frequency spectrum of the acoustic signal (deriving from multiple frequency components interacting within critical bandwidths; [13-19]), and thus bearing a principled relationship to the operation of the auditory system. The second, usually referred to as musical dissonance, arises from cultural musical practice [20-22], wherein certain tones or simultaneous combinations of tones are perceived to be syntactically 'unstable', requiring resolution by succeeding 'stable' tones or tone-combinations [21, 23-25].

Experiment 2: Group differences of ratings for original vs. spectrally manipulated music

As briefly mentioned in the main text, the difference in perceived pleasantness (between original and spectrally manipulated music) was greater in Western than in Mafa listeners (Figure 1B and Supplementary Figure 2,see also Supplementary Table 4, and interaction *spectrum* × *subgroup* in Supplementary Table 2). Several reasons might account for this finding:

(1) To the Mafa, the spectrally manipulated music (where more frequencies were simultaneously audible) may have sounded like more players were involved in the musical performance. Because Mafa flutes are exhausting to play, the Mafa appreciate powerful music performances, which are characterised by a great number of performers and also by the liveliness and duration of the musical pieces (which is also why longer pieces of music and their counterparts were probably more appreciated by the Mafa, see Supplementary Table 4).

However, the spectrally manipulated excerpts were not preferred over the original performance, so this is therefore probably a secondary, cultural effect adding to a sensory factor (see main text) working in the opposite direction. Such secondary cultural effects are also observable in many other non-Western cultures (see e.g. [26]) where musical sounds that appear strongly dissonant to most Western listeners may be perceived as musically acceptable.

(2) Western listeners are more accustomed to listening to variable and long polyphonic music sequences (notably, Mafa music is extremely repetitive, usually based on two-measure patterns). This provides more opportunity for probabilistic learning of harmonic structures, so Western listeners may therefore be more trained to perceive harmonies, and to differentiate between consonance and dissonance. It should be noted that whereas an alternation of consonance and dissonance may possibly also apply as a syntactical measure in Western music, this does not necessarily pertain to music in non-Western cultures.

(3) On average (not for individuals) the Mafa may have tended to not judge sounds as being highly unpleasant, thereby compressing the range.

Experiment 2: Forward vs. backward music

For Western listeners, the valence percept (their percept of pleasantness/unpleasantness) evoked by a musical signal can also be purposefully altered by a manipulation of its temporal order [27]. We therefore also investigated the appreciation of music played forward and backward with Mafa and Western participants. In this section of Experiment 2, which is not discussed in the main text, we included Western music pieces and Mafa music recordings and also their reversed counterparts.

Both Mafa and Western listeners preferred forward over reversed music stimuli when assessing the music of their own culture, and interestingly also when evaluating music of the respective other culture (Supplementary Figure 4, see also results of MANOVAs in Supplementary Table 4). The latter finding is striking considering the fact that neither the Western listeners nor the Mafa, had ever before listened to music of the respective other culture before, and did not know that the stimulus sample contained manipulated music pieces.

Time reversed envelopes of natural sounds seem to generally make sounds less pleasant to listen to, presumably because many natural sounds (including speech sounds) have rapid onsets and gradual decays. It is thus possible that the listeners responded to an unnaturalness of the reversed stimulus material. However, the Mafa listeners tended to attribute the manipulated Mafa music excerpts to real people (children or inhabitants of other villages). One of the reasons for this may be that the Mafa have no concept of technically manipulated natural sounds. Another reason may be that the on- and offsets of the Mafa flute tones are rather prompt and symmetric (~ 30 ms), so that the single flute tones sound quite similar in either direction. An alternative explanation (besides that the versions played backward sounded unnatural) may be that the listeners preferred the originally composed sequence of sounds over this sequence played backwards, both for the music of their own culture, as well as for the unknown music of the respective other culture (to which they were naïve).

Experiment 2: For Mafa but not for Western listeners the effect of spectral manipulation is additive

Only Western listeners showed a significant interaction of *spectrum* × *direction*, see Table 1), indicating that they responded more strongly to a spectral manipulation of the forward stimuli than to a spectral manipulation of the stimuli played backwards. For the Mafa listeners the effect of the spectral manipulation is merely additive, i.e., similar when forward and reversed stimuli were manipulated.

Experiment 2: Effects of length of stimuli

As reported in the section Stimuli and Experimental Design of the main text, the stimuli used in Experiment 2 varied in their duration (2, 10, 30 sec.). While the Mafa listeners showed a significant interaction between length and valence (the longer the music, the stronger the pleasantness ratings), Western listeners only show such an interaction for the Mafa (and not for Western) music. The finding of a significant interaction between length and valence in Mafa listeners may be explained by the observation that a flute performance of only a short duration is generally regarded as a bad performance among Mafas (the Mafa listeners would often attribute the short stimuli to unmotivated performers or children trying to play the flute), which is probably connected to the fact that playing the Mafa flutes is a vigorous, physically challenging activity, and in a natural setting a longer performance corresponds to higher skill and greater fitness [9, 28] of the players. The question of why Western listeners show a significant interaction between length and valence for the unknown Mafa music but not their own remains open.

Experiment 1 and 2: How the capability to recognize emotional expressions in Western music relates to its appreciation

Finally, we investigated whether the capability to recognize emotional expressions in Western music pieces modulates a listener's appreciation (valence percept) of the Western music. For this, the emotion expression recognition performance of Mafa individuals who had participated in both experiments was correlated with their valence assessment behaviour as determined in Experiment 2 (for Western listeners, no significant correlations were found (neither for the manipulations on direction, r = 0.17, nor on those of spectrum, r = -0.11), presumably due to the ceiling effects observed in Experiment 1). Interestingly, the Mafa showed a positive correlation between performance in the emotional expression recognition experiment and the extent to which they differentiated (in terms of valence) between forward and reversed Western music excerpts (r = 0.74; p = 0.015; N = 10). In contrast, no correlation was observed between the performance in the emotional expression recognition experiment and the extent to which they differentiated between music excerpts with original and manipulated spectrum (r = 0.12; p = 0.737; N = 10). This suggests that Mafa participants who better recognized the emotional expressions in Western music were more sensitive to a distortion of the temporal (but not spectral) order of the music excerpts, and showed a greater preference for the original Western music excerpts as opposed to the manipulated versions. In other words, a capacity to recognize emotional expressions in the Western music contributed to the Mafa perceiving the Western music as pleasant and the reversed music as unpleasant. One possibility that may account for this finding is that recognizing emotional expressions in the music makes the music more pleasant for the Mafa, and that playing the music backward distorts the emotional expression for the Mafa more strongly than its spectral manipulation.

- Supplemental Experimental Procedures

Participants

Only Mafas who had never before listened to radio, and had never been to or lived in the proximity of a church were accepted as participants for the experiments because exposure to Western music mainly spreads with electricity supply (and the resulting possibility to operate radios) and Christianization (through Western Christian song). The age of most Mafa participants had to be estimated because they do not have a concept of accumulated age. The Mafa participants could neither be categorized as musicians nor as nonmusicians, because they do not have a concept of professional musicianship even though many dedicate time to musical activity on an everyday basis.

Stimuli

Stimuli of the emotional expression recognition experiment (Experiment 1) were controlled for timbre, tone volume, tone attack, tone decay, tone release, and tempo fluctuation. No Mafa flute music was included in this experiment, as the Mafa music recorded locally was mandatorily associated with certain festive rituals, but not unambiguously assigned by its performers to any particular emotional expressions.

Not more than two versions of the same stimulus category occurred in direct succession, and an equivalent number of stimuli from each category occurred in each third of the experiment.

The Mafa flute music included in Experiment 2 plays a key role in the performance of a number of rituals. Its integrated and context-dependent character in Mafa society is nicely illustrated by the observations that the Mafa do not have a word for music at all, and that music performers labeled music patterns by reference to the rituals where they are traditionally performed. To the naïve Western ear, Mafa music sounds bizarre (supplementary audio examples, Figure 3, Supplementary Figure 1).

Experimental design

In Experiment 1, an emotional expression recognition experiment with computer-generated piano music excerpts was performed, where the music stimuli had to be classified with the corresponding depictions of facial expressions from the Ekman archive (happy, sad, scared/fearful) [29]. The experiment was conducted with a group of Mafa and German participants. Before the experiment all participants were asked to verbally identify the three emotional expressions depicted on the paper sheets and agreement was reached about which emotional expression corresponded to which depiction (three Mafa participants had difficulties recognizing facial expressions on the two-dimensional paper sheet presentations and had to be excluded from the test). Note that emotional expressions in music need not map directly onto the emotions that have traditionally been researched in the facial expression literature (for more information on musical emotions see e.g. [30, 31]).

During the experiment, the participants had to indicate which facial expression best fit with the expression of each music excerpt (forced choice). Additionally, they were asked to verbally identify the corresponding emotional expression in order to ensure that they remained attentive during the experiment and that their response did not arise from a faulty recognition of the facial expressions. Importantly, if the two responses (indication of facial expression and verbalization of the emotional expression) did not match, this was taken as indicative of low attentiveness of the participant and the trial had to be repeated (ensuring active and concentrated participation in the experiment). If the participants had problems recognizing facial expressions on the paper sheets, and accordingly had several mismatches in a row, they were excluded from the experiment.

To illustrate the extremes of the bipolar dimension (Experiment 2) on the slider interface in a universally comprehensible manner, we adapted an SAM (Self-Assessment-Manikin [32]) which was designed to signify valence (from displeasure to pleasure) but in the original rather depicts a scale from unhappiness to happiness (see Supplementary Figure 6).

To validate the valence scale ratings, participants were asked to verbally express their approximate valence judgement before pressing a button on the slider interface to confirm their rating. If the two responses (slider and articulation) did not match, this was taken as an indication of a low attentiveness of the participant, and in such cases the trial was repeated.

The instructions were presented in a standardized fashion with the Presentation software either in German or Mafa (translated and recorded before the field trip by a Mafa woman living in Europe). While auditory instructions were presented over headphones, the experimenter simultaneously demonstrated how to handle the interface. During the experiment only the participants could listen to the stimuli to avoid a response bias through responses of the experimenter.

The Mafa usually attributed the manipulated music to real people. Typical Mafa comments were: “you shouldn’t let children play the flutes, this is no good”, or “I know this, this is from the people of the Gouzda village, I really don’t like how they play the flutes”.

- Supplemental Statistical Evaluation

Test results were considered to be significant when the probability of error was lower than 0.05 (two-tailed). When necessary, a Bonferroni correction was applied. In Experiment 1 a *t*-test against chance (1/3) was calculated to test for significance of the recognition rates of emotional expression in Western piano music excerpts (Figure 1A). In addition, a binomial distribution test (with the assumption that above a percentile rank of 95 the result is not chance) was calculated to determine the percentage of Mafa and Western participants that performed above this criterion in the experiment. A MANOVA for the percentages of correct responses was computed with the within-subject factor “emotional expression” (happy vs. sad vs. scared) and the between-subjects factor “subgroup” (Mafa vs. Western listeners). A paired-samples *t*-test was conducted to test whether the choice of the non-target categories in attributing the emotional expression were biased towards one of these two emotions (i.e., it was determined if one of the two non-target categories was chosen significantly more often than the other) (table in Supplementary Figure 5).

For the statistical evaluation in Experiment 2, the rating values for each participant were *z*-transformed to eliminate individual assessment tendencies. With these values, four MANOVAs with the within-subject-factors *direction* (forward vs. reversed), *spectrum* (consonant vs. dissonant), and *length* (2, 10, and 30 sec) were computed for each subgroup and each musical culture separately (Supplementary Table 4).

We used these four models because our focus of interest was the investigation of the subgroup-specific responses to the music from both cultures. Furthermore, we observed interactions for *direction* (forward vs. reversed) × *subgroup* (Western vs. Mafa listeners) as well as for *spectrum* (consonant vs. dissonant) × *subgroup* for both Western and Mafa music (see additional online material, Supplementary Table 2 and 3). In the mixed-model MANOVAs used in these analyses, *direction*, *spectrum*, and *length* (2, 10, and 30 sec) served as within-subject-, and *subgroup* as between-subjects-factor. Because not all Mafa participants took part in both Experiments A and B, we computed two separate MANOVAs for Western and Mafa music, respectively. These MANOVAs were computed for the *z*-transformed (Supplementary Table 2) as well as for the non-transformed values (Supplementary Table 3). The latter analysis was necessary because the main effect of subgroup could not be evaluated for the *z*-transformed values (where the mean within each subject was zero, resulting in a mean of zero in the two subgroups). A significant main effect of *subgroup* was revealed in both MANOVAs (with the non-transformed values) for Western as well as for Mafa music. Apart from these differences, the MANOVAs (for *z*-transformed and non-transformed values) revealed comparable results.

Supplemental References

1. Masataka, N. (2006). Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. Developmental Science *9*, 46-50.
2. Zentner, M.R., and Kagan, J. (1998). Infants’perception of consonance and dissonance in music. Infant Behaviour & Development *21*, 483-492.
3. Trainor, L.J., and Heinmiller, B.M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. Infant Behavior and Development *21*, 77-88.

4. Trainor, L.J., Tsang, C.D., and Cheung, V.H.W. (2002). Preference for Sensory Consonance in 2- and 4-Month-Old Infants. Music Perception *20*, 187-194.
5. James, D.K., Spence, C.J., and Stepsis, B.W. (2002). Fetal learning: a prospective randomized controlled study. Ultrasound in Obstetric and Gynecology *20*, 431-438.
6. Zajonc, R.B. (1968). Attitudinal effects of mere exposure. American Psychological Association.
7. McDermott, J., and Hauser, M. (2005). The origins of music: Innateness, uniqueness, and evolution. Music Perception *23*, 29-59.
8. Parncutt, R. (in press). Prenatal conditioning, the mother schema, and the origins of music and religion. Musicae Scientiae.
9. Fitch, W.T. (2005). The Evolution of Music in Comparative Perspective. Annals of the New York Academy of Sciences *1060*, 29-49.
10. Balkwill, L.L., Thompson, W.F., and Matsunaga, R. (2004). Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. Japanese Psychological Research *46*, 337-349.
11. Balkwill, L.L., and Thompson, W.F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. Music Perception *17*, 43-64.
12. Hanslick, E. (1980). Vom Musikalisch Schönen. 20 Edition.
13. Braun, M. (1999). Auditory midbrain laminar structure appears adapted to f0 extraction: further evidence and implications of the double critical bandwidth. Hearing Research *129*, 71-82.
14. Greenwood, D.D. (1991). Critical Bandwidth and Consonance in Relation to Cochlear Frequency-Position Coordinates. Hearing Research *54*, 164-208.
15. Helmholtz, H.L.F. (1885/1954). On the Sensations of Tone as a Physiological Basis for the Theory of Music (Die Lehre von den Tonempfindungen), 2 Edition, (New York: Dover Publications).
16. Plomp, R., and Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. Journal of the Acoustical Society of America *38*, 548-560.
17. Kameoka, A., and Kuriyagawa, M. (1969). Consonance theory, part I: Consonance of dyads. Journal of the Acoustical Society of America *45*, 1451-1459.
18. Terhardt, E. (1974). On the perception of periodic sound fluctuations (roughness). Acustica *30*, 201-213.
19. Sethares, W.A. (1993). Local consonance and the relationship between timbre and scale. Journal of the Acoustical Society of America *94*, 1218-1228.
20. Terhardt, E. (1984). The concept of musical consonance: A link between music and psychoacoustics. Music Perception *1*, 276-295.
21. Cazden, N. (1945). Musical Consonance and Dissonance: A Cultural Criterion. The Journal of Aesthetics and Art Criticism *4*, 3-11.
22. Lundin, R.W. (1947). Towards a cultural theory of consonance. Journal of Psychology *23*, 45–49.
23. Apel, W. (1972). Harvard Dictionary of Music, 2 Edition, (Cambridge, MA: Harvard University Press).
24. Burns, E.W., and Ward, W.D. (1982). Intervals, scales, and tuning. In The Psychology of Music, D. Deutsch, ed. (New York: Academic), pp. 241-269.
25. Terhardt, E. (1983). Musikwahrnehmung und elementare Hörempfindungen. Audiological Acoustics *22*, 52-56.
26. Stobart, H.F. (1996). Tara and Q'iwa: Words of Sound and Meaning. In Cosmología y música en los Andes (Music and Cosmology in the Andes), Volume 55, M.P. Baumann, ed. (Madrid, Frankfurt: Biblioteca Iberoamericana), pp. 67-81.
27. Menon, V., and Levitin, D.J. (2005). The rewards of music listening: Response and physiological connectivity of the mesolimbic system. Neuroimage *28*, 175-184.
28. Huron, D. (2001). Is Music an Evolutionary Adaptation? Annals of the New York Academy of Sciences *930*, 43-61.

29. Ekman, P. (1976). Pictures of facial affect (Palo Alto, CA: Consulting Psychologists Press).
30. Scherer, K. (2004). Which Emotions Can be Induced by Music? What Are the Underlying Mechanisms? And How Can We Measure Them? Journal of New Music Research *33*, 239-251.
31. Scherer, K.R., and Zentner, M.R. (2001). Emotional effects of music: Production rules. In Music and emotion: theory and research., P.N. Juslin and J.A. Sloboda, eds. (Oxford, New York: Oxford University Press), pp. 361–392.
32. Mehrabian, A., and Russell, J.A. (1974). An approach to environmental psychology, (Cambridge, MA: MIT).