Current Biology *20*, 1035–1041, June 8, 2010 ©2010 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2010.04.019

Report

Individual Differences Reveal the Basis of Consonance

Josh H. McDermott,[1,*] Andriana J. Lehr,[2] and Andrew J. Oxenham[2]
[1]Center for Neural Science, New York University, New York, NY 10003, USA
[2]Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA

Summary

Some combinations of musical notes are consonant (pleasant), whereas others are dissonant (unpleasant), a distinction central to music. Explanations of consonance in terms of acoustics, auditory neuroscience, and enculturation have been debated for centuries [1–12]. We utilized individual differences to distinguish the candidate theories. We measured preferences for musical chords as well as nonmusical sounds that isolated particular acoustic factors—specifically, the beating and the harmonic relationships between frequency components, two factors that have long been thought to potentially underlie consonance [2, 3, 10, 13–20]. Listeners preferred stimuli without beats and with harmonic spectra, but across more than 250 subjects, only the preference for harmonic spectra was consistently correlated with preferences for consonant over dissonant chords. Harmonicity preferences were also correlated with the number of years subjects had spent playing a musical instrument, suggesting that exposure to music amplifies preferences for harmonic frequencies because of their musical importance. Harmonic spectra are prominent features of natural sounds, and our results indicate that they also underlie the perception of consonance.

Results

Figure 1A shows the pleasantness ratings given by a group of subjects to different combinations of notes. Some combinations were consistently rated higher than others, irrespective of the instrument playing the notes. This is the phenomenon of consonance, the origins of which have remained controversial throughout history [1–12].

Ancient thinkers viewed consonance as determined by ratios (Figure 1B), but in modern times it has been linked to acoustic properties thought to be important to the auditory system [10]. The dominant contemporary theory posits that dissonance is due to beating between frequency components [2, 13–15]. Beating occurs whenever two sinusoids of differing frequency are combined (Figure 1C, top left). Over time, the components drift in and out of phase, and the combined waveform waxes and wanes in amplitude. This modulation produces a sound quality, known as roughness, that listeners typically describe as unpleasant [21, 22] and that has been thought to be prevalent in dissonant, but not consonant, musical chords [13–15].

Figure 1C (bottom two rows) shows spectra and waveforms for two musical intervals (chords with two notes). The minor second, a dissonant interval, contains many pairs of frequency components that are close but not identical in frequency and that produce beating, visible as amplitude fluctuations in the waveform. The (consonant) fifth presents a different picture, containing frequencies that are widely spaced or exactly coincident and that thus produce little beating.

However, the intervals differ in another respect. The fifth contains frequencies that are approximately harmonically related—they are all multiples of a common fundamental frequency (F0) (Figure 1C, top right). Not every component of the harmonic series is present, but each frequency corresponds to a harmonic. In this respect the fifth bears some resemblance to an individual musical note, whose frequencies are generally a series of harmonics, the F0 of which corresponds to the pitch of the note. The resemblance does not hold for the minor second, whose frequencies are inharmonic. This contrast exemplifies an alternative view—that consonant chords derive their pleasantness not from the absence of beating, but rather from their similarity to single notes with harmonic spectra [3, 17–20].

It has also seemed plausible that consonance might not be rooted in acoustics at all and is instead the arbitrary product of enculturation [23]—listeners might simply learn to like specific chords that are prevalent in the music of their culture. This notion is fueled in part by the use of the equal-tempered scale in modern music, in which consonant intervals only approximate integer ratios (Figure 1B) and are thus somewhat less harmonic, and less devoid of beating, than they would be otherwise. Of course, enculturation and acoustic-based explanations are not mutually exclusive. If a particular acoustic property were to underlie the distinction between consonance and dissonance, listeners could potentially learn an aesthetic association with that property by hearing it repeatedly in music.

In our efforts to address these issues, we took advantage of the fact that some listeners showed stronger consonance preferences than others. We investigated whether intersubject variability in consonance preferences could be explained by variation in preferences for particular acoustic factors. We measured acoustic preferences by asking subjects to rate the pleasantness of nonmusical stimuli designed to independently vary in beating and harmonic content. To isolate the aesthetic contribution of a particular factor, we formed preference measures by subtracting the ratings of stimuli possessing that factor from those that did not. If beating or harmonic spectra underlie consonance, the associated acoustic preference measures should be correlated with our consonance measures. To ensure robustness and replicability, we separately examined these correlations for chords made from different instrument sounds and separately tested two large cohorts of subjects (n = 142, 123).

Consonance Preferences

We measured consonance preferences with chord rating tests (Figure 1A). Two summary measures of this preference were computed for each instrument sound (timbre), one for

*Correspondence: jhm@cns.nyu.edu

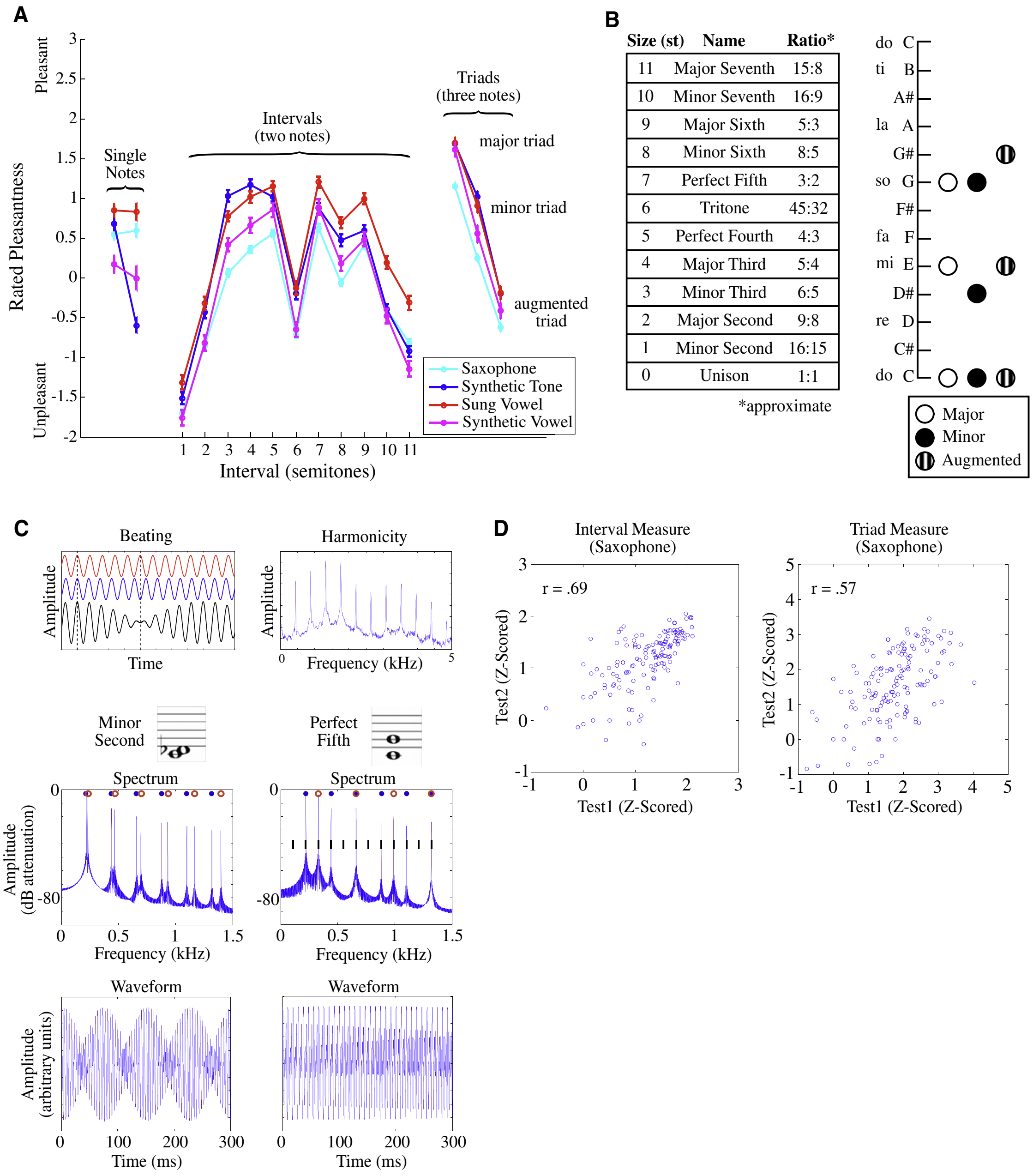

Figure 1. Consonance Preferences and Their Possible Acoustic Basis

(A) Mean pleasantness ratings of individual notes and chords, for cohort 1. The two single-note conditions differed in pitch (lower pitch on left). Error bars denote standard errors (SEs).

(B) Chords used in experiments, with diatonic scale as reference. Ratios in stimuli approximated those listed, because the equal-tempered scale was used.

(C) Beating and harmonicity in consonant and dissonant intervals. Top left: two sinusoids of different frequencies are plotted in red and blue; their superposition (in black) contains amplitude modulation known as "beating." Top right: amplitude spectrum for the note A440 played on an oboe. The frequencies in the note are all integer multiples of the fundamental frequency of 440 Hz and as a result are regularly spaced along the frequency axis. Bottom rows: spectra and waveforms for the minor second and perfect fifth, generated by combining two synthetic complex tones with different fundamental frequencies. Red and blue circles denote the frequencies belonging to each note. The frequencies of the fifth are approximately harmonically related (black lines denote harmonic series). Amplitude modulation (from beating) is evident in the waveform of the minor second, but not the fifth.

(D) Scatter plots of consonance preference measures computed from z-scored ratings of cohort 1 (saxophone notes) on two successive tests. The interval consonance measure was formed by subtracting the mean rating of the five lowest-rated intervals from that of the five highest-rated intervals. The triad consonance measure was formed by subtracting the ratings for the augmented triad from that of the major triad. Each circle denotes the scores of a single subject. Here and elsewhere, r is the Spearman correlation coefficient.

two-note chords (intervals), and one for three-note chords (triads). Each measure was formed from the difference between the ratings of consonant and dissonant chords. Large values of these measures indicate strong preferences, and individual subjects produced consistently different values, indicated by correlations in their scores from two successive tests. These correlations were not simply due to differences in how subjects used the rating scale, for they persisted once the ratings were z-scored to equalize the range used by each subject. Figure 1D shows representative test-retest scatter plots (for the saxophone consonance measures for cohort 1); test-retest correlations for the different note timbres and subject cohorts ranged from 0.60 to 0.75 (interval measure) and 0.46 to 0.63 (triad measure), all statistically significant ($p < 0.0001$).

Acoustic Preferences

Beating preferences were assessed by comparing ratings of pairs of pure tones (single frequencies) presented to either the same or different ears (diotic and dichotic presentation, respectively). Dichotic presentation of two tones is known to greatly attenuate perceived beats [24] but leaves the spectrum (and its harmonicity and pitch) unchanged [25]. In a pilot experiment, we found that pure-tone pairs were rated more highly when presented dichotically than diotically, but only when the tones were sufficiently close in frequency to fall within the same cochlear filter (Figure 2A). Beating is known to be audible only for frequency differences small enough to be registered by the cochlear filter bank [2]; our results therefore suggest that the dichotic-diotic rating difference reflects the extent to which audible beats are judged to be objectionable. To form a measure of preference for stimuli lacking beats (B1), we obtained ratings for narrowly spaced tone pairs in three frequency ranges (Figure 2B) and subtracted the ratings of all the diotic from all the dichotic tone pairs.

To assess preferences for harmonicity, we compared pleasantness ratings for harmonic and inharmonic complex tones. The harmonic stimuli contained a subset of the frequencies of a normal harmonic tone, spaced widely enough apart to avoid substantial beating (Figure 2C). The inharmonic stimuli were generated by perturbing the frequencies of the harmonic tones. The main harmonicity preference measure (H1) was the difference between the mean ratings of the harmonic and inharmonic stimuli.

Because the beating test stimuli, having but two frequency components, might be considered less similar to musical chords than the harmonic and inharmonic test stimuli, we also used a second measure of harmonicity preference. For this measure (H2), we subtracted the ratings of the low-frequency dichotic tone pairs (from the beating test stimuli; Figure 2B) from those of single pure tones (the simplest case of a harmonic stimulus). The frequencies of the tone pairs were not harmonically related and produced minimal beating due to the dichotic presentation; they allowed us to use some of the stimuli from the beating measures to probe harmonicity. This measure also served as a control for the possibility that distortion products might have produced beating in the other harmonicity test stimuli [26].

Subjects on average preferred harmonic over inharmonic spectra and stimuli without beats over those with beats (Figure 2D), but individual differences were evident in all the acoustic preference measures (Figures 2E and 2F: B1 and H1, cohort 1; test-retest correlations for the acoustic preferences in each cohort ranged from 0.41 to 0.76, all $p < 0.0001$). Notably, the beating and harmonicity effects were not significantly correlated across subjects, suggesting that our tests isolated two largely independent effects (Figure 2G: B1 and H1, cohort 1; correlations between the beating and harmonicity measures of each cohort ranged from -0.09 to 0.17; $p > 0.05$ in all cases).

Correlations between Acoustic and Consonance Preferences

Although the reliability, average effect size, and variance of the beating and harmonicity preferences were comparable (Figure S1), we found large differences in their correlations with consonance preferences. These correlations for the beating measures (Figure 3A, top) were weak and inconsistent (see also Figure S2). In contrast, both harmonicity measures correlated strongly with both consonance measures for synthetic as well as natural note sounds (Figure 3A, lower two rows). Subjects with stronger preferences for harmonic spectra thus had stronger preferences for consonant over dissonant chords. Although one might imagine that a listener's preference for one chord over another would be subject to many different influences (their mood, the musical genre most recently heard, etc.), our measures of their preference for harmonic spectra explain a sizeable portion of the variance in consonance preferences, whereas our beating measure explains little (Figure 3B).

To gain insight into these effects, we examined correlations between the acoustic preference measures and ratings of individual chords, averaging across note timbres to increase reliability (Figure 3C). The beating measure yielded modest negative correlations with the minor and major second (the two leftmost intervals in the plots), but not with other dissonant intervals. The harmonicity measures, in contrast, were negatively correlated with every dissonant chord that we tested. We note that the similarity in correlation patterns for the two harmonicity measures is nontrivial, because the measures were derived from nonoverlapping sets of stimuli that physically had little in common.

Effects of Musical Experience

When our acoustic preference measures were correlated with the number of years each subject had spent playing a musical instrument, another distinction emerged: both harmonicity measures were positively correlated with musical experience, whereas our beating measure was not (Figure 4A). Subjects with more musical experience thus had stronger preferences for harmonic over inharmonic spectra. A priori there was little reason to expect this—none of the acoustic test stimuli had musical connotations, and in fact were designed to avoid physical similarity to musical stimuli. This result is strong evidence for the importance of harmonicity in music, and it suggests that the aesthetic response to harmonic frequency relations is at least partially learned from musical experience.

Musical experience was also correlated with the strength of consonance preferences (Figure 4B), further consistent with a role for learning. Given that our measure of musical experience is only a crude estimate of the degree to which subjects had internalized the structure of Western music, it seems likely that the musical experience correlations are underestimates, perhaps substantially so (Figure S4).

Discussion

We used individual differences to explore the basis of consonance and dissonance. Our findings suggest that consonance is due to harmonic frequency relations and that dissonance results from note combinations that produce inharmonic frequencies. Moreover, preferences for harmonic spectra and consonant chords appear to be heavily influenced by musical experience. Our results thus support a strong role for enculturation in consonance but indicate that rather than learning to

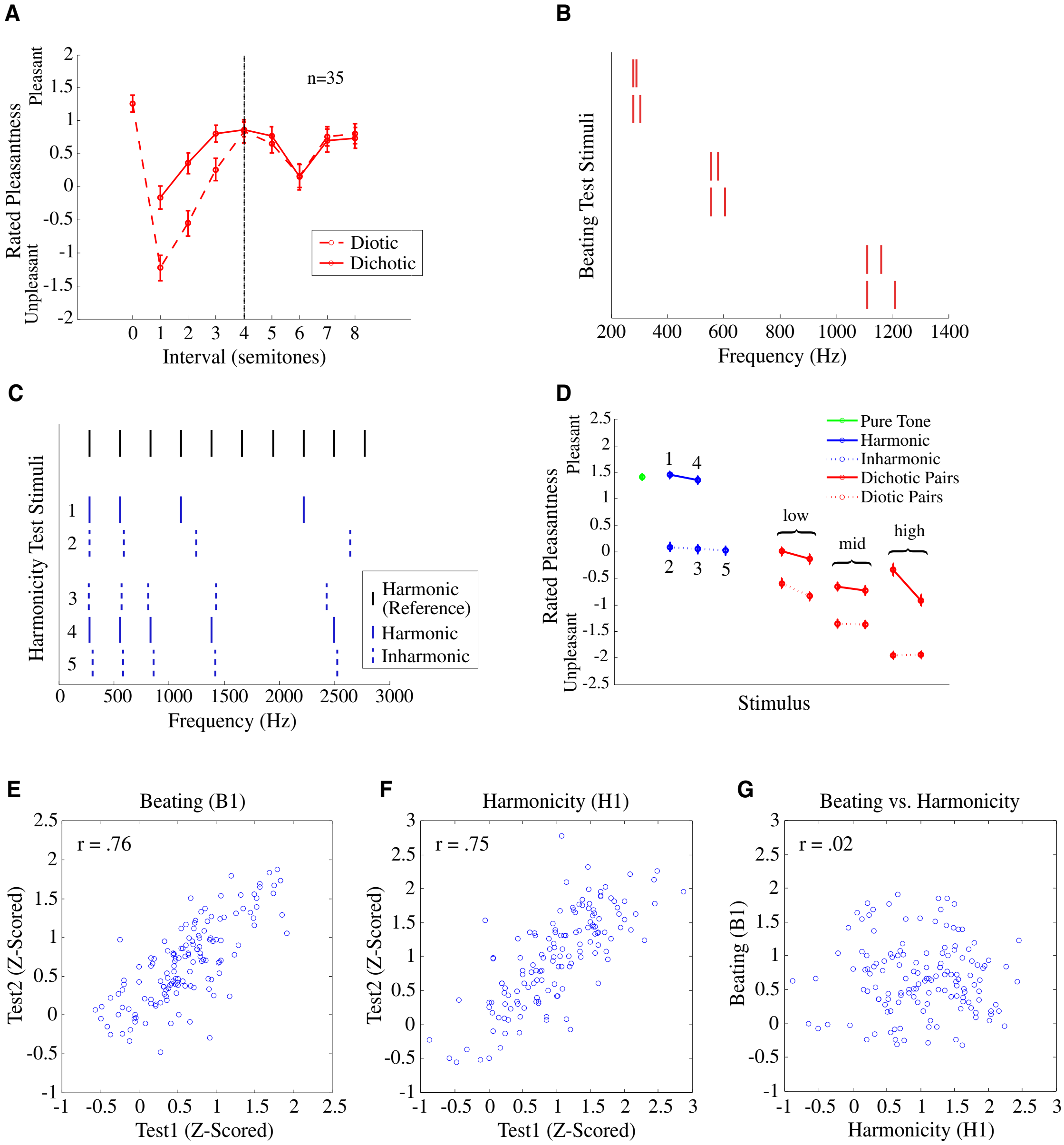

Figure 2. Diagnostic Measures of Beating and Harmonicity

(A) Mean pleasantness ratings of 35 subjects for pairs of pure tones, diotically or dichotically presented. Error bars denote SEs. The unison (0 semitone separation) could only be presented diotically. The dashed line represents the approximate frequency separation (derived from estimated cochlear filter bandwidths) at which beats become inaudible.

(B) Schematic spectra of beating test stimuli. Tone pairs were separated by either 0.75 or 1.5 semitones, such that considerable beating was heard when presented diotically.

(C) Schematic spectra of harmonicity test stimuli. Inharmonic complex tones were generated via small perturbations to the frequencies of each harmonic component, ensuring that all components were separated widely enough to avoid substantial beating. All other aspects of the harmonic and inharmonic test stimuli were identical. Numbers to left of spectra are given to enable comparison with (D).

(D) Mean pleasantness ratings of acoustic test stimuli, cohort 1. Error bars denote SEs.

(E and F) Scatter plots of B1 and H1 measures computed from z-scored ratings of cohort 1 on two successive tests. See also Figures S1 and S5.

(G) Scatter plot of B1 and H1 measures, averaged over the two tests.

find specific arbitrary chords pleasing, listeners learn to like a general acoustic property, that of harmonicity. Harmonic structure has broad importance in the auditory system [27, 28], and chord perception may simply involve the assignment of valence to the output of mechanisms that analyze harmonicity for pitch perception [29] or sound segregation [30].

Audible beating, or roughness, often evokes strong unpleasant reactions in listeners and is routinely used to modulate tension in music [31–33]. However, its aesthetic association does not appear to be learned from music-related experience, and we find little evidence for a relation to consonance. This is probably because dissonant chords do not always produce large degrees of beating, whereas consonant chords sometimes do. Because the beating of two frequencies becomes weaker as their amplitudes become more different [34], the beating produced by two notes depends on the notes'

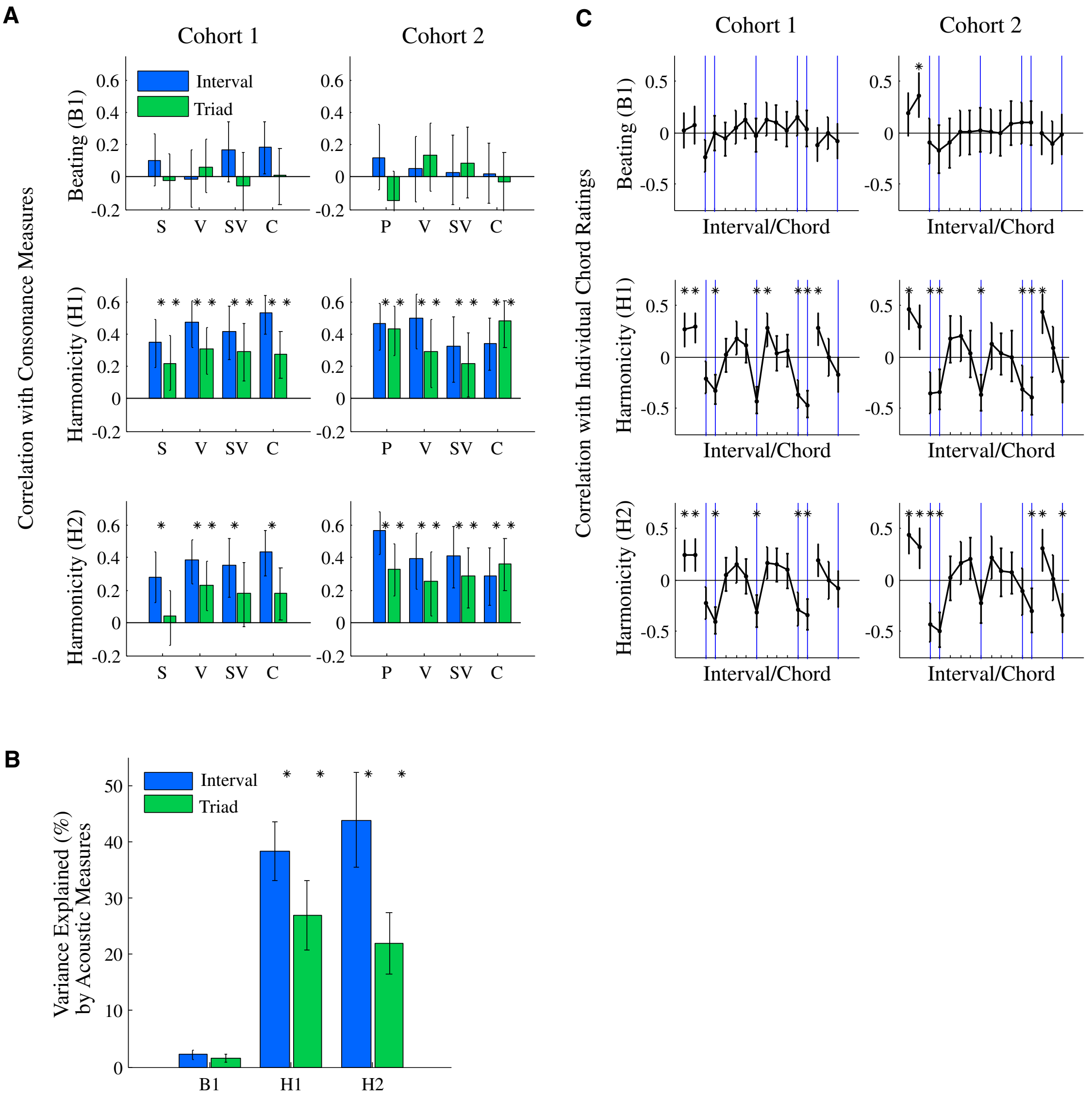

Figure 3. Correlations of Beating and Harmonicity Preferences with Consonance Preferences

(A) Correlations with interval and triad consonance measures. Letters on the x axis denote note timbre (saxophone, S; sung vowels, V; synthetic sung vowels, SV; synthetic complex tone, C; pure tone, P). Here and in (C), error bars denote 95% confidence intervals, and asterisks denote significance (0.05 criterion).

(B) Variance of consonance measures explained by acoustic preferences. Error bars denote SEs. Asterisks indicate that the variance explained by the harmonicity preference measure was significantly greater than that for the beating preference measure (0.05 criterion, sign test).

(C) Correlations with ratings of individual chords, averaged across note timbre. Interval and chord arrangement within subpanels follows the conventions of Figure 1A. Blue vertical lines denote dissonant intervals and triads included in consonance measures. See also Figures S2 and S3.

spectra; it can thus vary considerably across instruments [35]. For this reason beating may not reliably indicate chord character, instead functioning as a largely orthogonal aesthetic influence. The perception of harmonic frequency relations, by contrast, is much less dependent on the exact frequency amplitudes [36] and thus may be more invariantly related to musical structure. At present we lack perceptually calibrated methods to confirm this intuition with measurements of harmonic content (see Supplemental Experimental Procedures) and instead used correlations with unambiguously harmonic or inharmonic stimuli to test the role of harmonicity.

Consonance has long been a battleground for nature/nurture debates of music. We provide some support for nurture in showing a role for musical experience, but our results also indicate that the debate should perhaps be reframed in terms of acoustic properties. Previous studies of consonance perception in infants [7, 9] and non-Western adults [5, 11] have generally used stimuli that varied in both harmonicity and beating. It could be fruitful to separately examine their effects, given that we found only harmonicity preferences to be related to musical experience. It remains possible that the effect of musical experience reflects enhancement of an initial innate bias for harmonic sounds rather than a purely learned effect. Indeed, this notion derives plausibility from the prominence of harmonicity in mammalian vocalizations, where it may provide a signal of health and

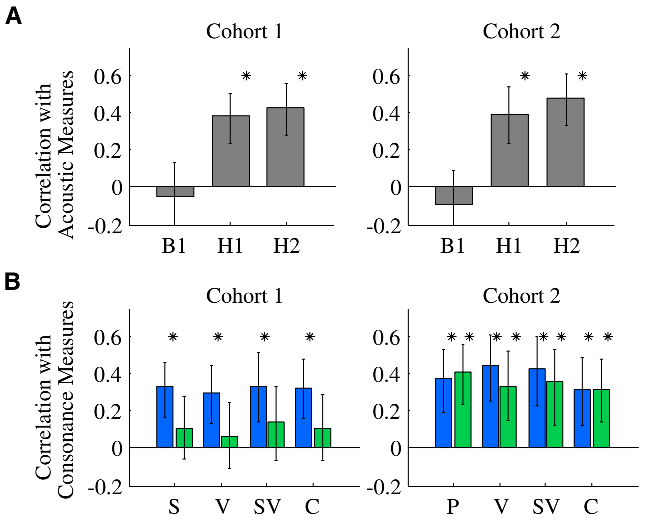

Figure 4. Effect of Musical Experience

(A) Correlation between the number of years a subject had spent playing an instrument and his or her acoustic preferences. Here and in (B), error bars denote 95% confidence intervals, and asterisks denote significance with 0.05 criterion, corrected for multiple comparisons. See also Figures S2 and S4.

(B) Correlations between musical experience and consonance preference measures (interval measure in blue, triad measure in green).

attractiveness [37], but a definitive resolution will require further study.

The idea that consonance derives from beating was fueled by reports that dissonance ratings of pure-tone pairs could predict the dissonance of intervals formed from complex tones (notes with multiple frequency components) [13, 14]. These studies argued that the dissonance of pure-tone pairs was due to beating and that their predictive value revealed the role of beating in consonance. However, we find that the dissonance of pure-tone pairs is a function both of their beating and of their harmonicity. Two narrowly separated frequencies are consistent only with an implausibly low fundamental frequency, and this seems to contribute as much to their unpleasantness as does their beating. Our H2 harmonicity measure was correlated with consonance preferences even though the tone pairs from which it was constructed were dichotically presented and thus produced minimal beating (Figure 3). This indicates that effects previously ascribed to beating probably had large, and unnoted, contributions from harmonicity.

Harmonicity preferences predicted chord ratings even though we used chords from the equal-tempered scale that were thus not perfectly harmonic. This suggests that the mechanisms for detecting harmonicity are somewhat coarsely tuned, perhaps because some natural sounds also deviate slightly from perfect harmonicity [38]. It remains to be seen whether harmonicity contributes to aesthetic responses to chord progressions, or to melodies, for instance via integration of frequency information over time [39]. Musical context critically influences whether a chord in a piece of music sounds pleasing [40], and the role of acoustic factors in such effects is an open issue.

Our study applies a new approach to old issues in music perception. Debates over consonance have remained unresolved because the candidate theories often make similar predictions [10] and because models of the candidate mechanisms [13–15, 19, 20, 22] hinge on assumptions and parameters that are difficult to verify. We have utilized individual differences to circumvent these difficulties and find evidence that harmonicity plays a key role in the perception of consonance.

Experimental Procedures

Methods are described in more detail in the Supplemental Information.

Participants and Method

All subjects (Minnesota undergraduates) completed a pair of acoustic tests (containing both the beating and harmonicity test stimuli) followed by paired chord-rating tests, each pair with chords generated with different note timbres (paired tests permitted test-retest reliability estimates). The two cohorts had similar demographic characteristics and differed only in taking slightly different versions of the tests (e.g., cohort 2 was tested on pure tones instead of saxophone notes). In each test, subjects were presented with stimuli in random order, with multiple repetitions of each stimulus (three for the acoustic tests; four for the chord tests), each time with a different root pitch.

Chords were derived from the equal-tempered scale. Chord root pitches were drawn from a fixed set in the octave above middle C, except for the sung vowel and saxophone stimuli, the root notes of which were drawn from G#3 upward to accommodate the ranges of the singer and instrument. After each trial, subjects entered a rating between –3 and 3, denoting the range from very unpleasant to very pleasant. Subjects were instructed to use the full rating scale. Before starting the tests, subjects were given a short practice test to familiarize them with the range of stimuli they would encounter.

Tone pairs used for the beating measures (Figure 2B) were separated by either 0.75 or 1.5 semitones, such that considerable beating was heard when presented diotically. Inharmonic complex tones were generated either by multiplying the frequencies of each harmonic component by a small factor or by adding a small constant offset to each frequency. All other aspects of the harmonic and inharmonic test stimuli were identical.

Analysis

The interval consonance measure was formed by subtracting the mean rating of the five lowest-rated intervals from that of the five highest-rated intervals. The triad consonance measure was formed by subtracting the ratings for the augmented triad from that of the major triad.

Spearman correlation coefficients and two-tailed significance tests were used throughout. Correction for multiple comparisons (modified Bonferroni) was performed on all sets of multiple statistical tests. Correlations between diagnostic measures and individual chords were computed with chord ratings averaged across timbre.

For computation of the variance in consonance preferences explained by acoustic preferences, the correlation for a given cohort and note sound was corrected for attenuation by using the test-retest correlations for both preferences, and then squared. These estimates of explained variance were averaged across note sound and cohort.

Supplemental Information

Supplemental Information includes five figures and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.cub.2010.04.019.

Acknowledgments

The authors thank Tom Bouchard, Peter Cariani, Ray Meddis, and Niels Waller for helpful discussions and Pascal Belin, Roland Fleming, Philip Johnson-Laird, Chris Plack, John Spiro, and Jonathan Winawer for comments on earlier drafts of the paper. This work was supported by the National Institutes of Health (NIH R01 DC 05216).

Received: March 5, 2010
Revised: April 7, 2010
Accepted: April 8, 2010
Published online: May 20, 2010

References

1. Rameau, J.P. (1971). Treatise on Harmony (New York: Dover Publications)., Originally published in 1722.
2. v Helmholtz, H. (1863). Die Lehre von den Tonempfindungen als Physiologische Grundlage für die Theorie der Musik (Braunschweig: F. Vieweg und Sohn).
3. Stumpf, C. (1890). Tonpsychologie (Leipzig: Verlag S. Hirzel).
4. Guernsey, M. (1928). The role of consonance and dissonance in music. Am. J. Psychol. *15*, 173–204.
5. Butler, J.W., and Daston, P.G. (1968). Musical consonance as musical preference: A cross-cultural study. J. Gen. Psychol. *79*(*1st Half*), 129–142.
6. Mathews, M.V., and Pierce, J.R. (1980). Harmony and nonharmonic partials. J. Acoust. Soc. Am. *68*, 1252–1257.
7. Zentner, M.R., and Kagan, J. (1996). Perception of music by infants. Nature *383*, 29.
8. Fishman, Y.I., Volkov, I.O., Noh, M.D., Garell, P.C., Bakken, H., Arezzo, J.C., Howard, M.A., and Steinschneider, M. (2001). Consonance and dissonance of musical chords: Neural correlates in auditory cortex of monkeys and humans. J. Neurophysiol. *86*, 2761–2788.
9. Trainor, L.J., Tsang, C.D., and Cheung, V.H.W. (2002). Preference for sensory consonance in 2- and 4-month-old infants. Music Percept. *20*, 187–194.
10. McDermott, J.H., and Oxenham, A.J. (2008). Music perception, pitch, and the auditory system. Curr. Opin. Neurobiol. *18*, 452–463.
11. Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A.D., and Koelsch, S. (2009). Universal recognition of three basic emotions in music. Curr. Biol. *19*, 573–576.
12. Bidelman, G.M., and Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. J. Neurosci. *29*, 13165–13171.
13. Plomp, R., and Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. J. Acoust. Soc. Am. *38*, 548–560.
14. Kameoka, A., and Kuriyagawa, M. (1969). Consonance theory. J. Acoust. Soc. Am. *45*, 1451–1469.
15. Hutchinson, W., and Knopoff, L. (1978). The acoustical component of western consonance. Interface *7*, 1–29.
16. Sethares, W.A. (1999). Tuning, Timbre, Spectrum, Scale (Berlin: Springer).
17. Terhardt, E. (1974). Pitch, consonance, and harmony. J. Acoust. Soc. Am. *55*, 1061–1069.
18. Tramo, M.J., Cariani, P.A., Delgutte, B., and Braida, L.D. (2001). Neurobiological foundations for the theory of harmony in western tonal music. Ann. N. Y. Acad. Sci. *930*, 92–116.
19. Cariani, P.A. (2001). Temporal codes, timing nets, and music perception. J. New Music Res. *30*, 107–135.
20. Ebeling, M. (2008). Neuronal periodicity detection as a basis for the perception of consonance: A mathematical model of tonal fusion. J. Acoust. Soc. Am. *124*, 2320–2329.
21. Terhardt, E. (1974). On the perception of periodic sound fluctuations (roughness). Acustica *30*, 201–213.
22. Daniel, P., and Weber, R. (1997). Psychoacoustical roughness: Implementation of an optimized model. Acustica *83*, 113–123.
23. Lundin, R.W. (1947). Toward a cultural theory of consonance. J. Psychol. *23*, 45–49.
24. Rutschmann, J., and Rubinstein, L. (1965). Binaural beats and binaural amplitude-modulated tones: Successive comparison of loudness fluctuations. J. Acoust. Soc. Am. *38*, 759–768.
25. Bernstein, J.G., and Oxenham, A.J. (2003). Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number? J. Acoust. Soc. Am. *113*, 3323–3334.
26. Feeney, M.P. (1997). Dichotic beats of mistuned consonances. J. Acoust. Soc. Am. *102*, 2333–2342.
27. Kadia, S.C., and Wang, X. (2003). Spectral integration in A1 of awake primates: Neurons with single- and multipeaked tuning characteristics. J. Neurophysiol. *89*, 1603–1622.
28. Lewis, J.W., Talkington, W.J., Walker, N.A., Spirou, G.A., Jajosky, A., Frum, C., and Brefczynski-Lewis, J.A. (2009). Human cortical organization for processing vocalizations indicates representation of harmonic structure as a signal attribute. J. Neurosci. *29*, 2283–2296.
29. Plack C.J., Oxenham A.J., Popper A.J., and Fay R.R., eds. (2005). Pitch: Neural Coding and Perception (New York: Springer).
30. Darwin, C.J. (1997). Auditory grouping. Trends Cogn. Sci. *1*, 327–333.
31. Pressnitzer, D., McAdams, S., Winsberg, S., and Fineberg, J. (2000). Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. Percept. Psychophys. *62*, 66–80.
32. Vassilakis, P. (2005). Auditory roughness as a means of musical expression. Selected Reports in Ethnomusicology *12*, 119–144.
33. Bigand, E., Parncutt, R., and Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. Percept. Psychophys. *58*, 124–141.
34. Kohlrausch, A., Fassel, R., and Dau, T. (2000). The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. J. Acoust. Soc. Am. *108*, 723–734.
35. Cook, N.D. (2009). Harmony perception: Harmoniousness is more than the sum of interval consonance. Music Percept. *27*, 25–41.
36. Ritsma, R.J. (1967). Frequencies dominant in the perception of the pitch of complex sounds. J. Acoust. Soc. Am. *42*, 191–198.
37. Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G.A., Kawahara, H., and Belin, P. (2010). Vocal attractiveness increases by averaging. Curr. Biol. *20*, 116–120.
38. Fletcher, H., Blackham, E.D., and Stratton, R. (1962). Quality of piano tones. J. Acoust. Soc. Am. *34*, 749–761.
39. Parncutt, R. (1989). Harmony: A Psychoacoustical Approach (Berlin: Springer-Verlag).
40. Steinbeis, N., Koelsch, S., and Sloboda, J.A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. J. Cogn. Neurosci. *18*, 1380–1393.

Current Biology, Volume *20*

Supplemental Information

Individual Differences

Reveal the Basis of Consonance

Josh H. McDermott, Andriana J. Lehr, and Andrew J. Oxenham

SUPPLEMENTAL DATA

Reliability and variance of acoustic test measures.

Fig. S1 presents the reliability and variance of each acoustic preference measures for each cohort. The reliability is the correlation of each measure across two successive tests. The variance is the variance of each measure across individuals, with each measure computed from the combined ratings of the two tests (it indicates the extent of individual differences in each measure). Both were computed on the raw ratings, as these were used to compute correlations with the consonance measures (z-scored ratings were used only to calculate the reliabilities reported in the main text, to ensure that the correlations did not reflect rating-scale effects). The beating (B1 and B2; see Supp. Fig. 2) and harmonicity (H1 and H2) measures were roughly comparable in both respects.

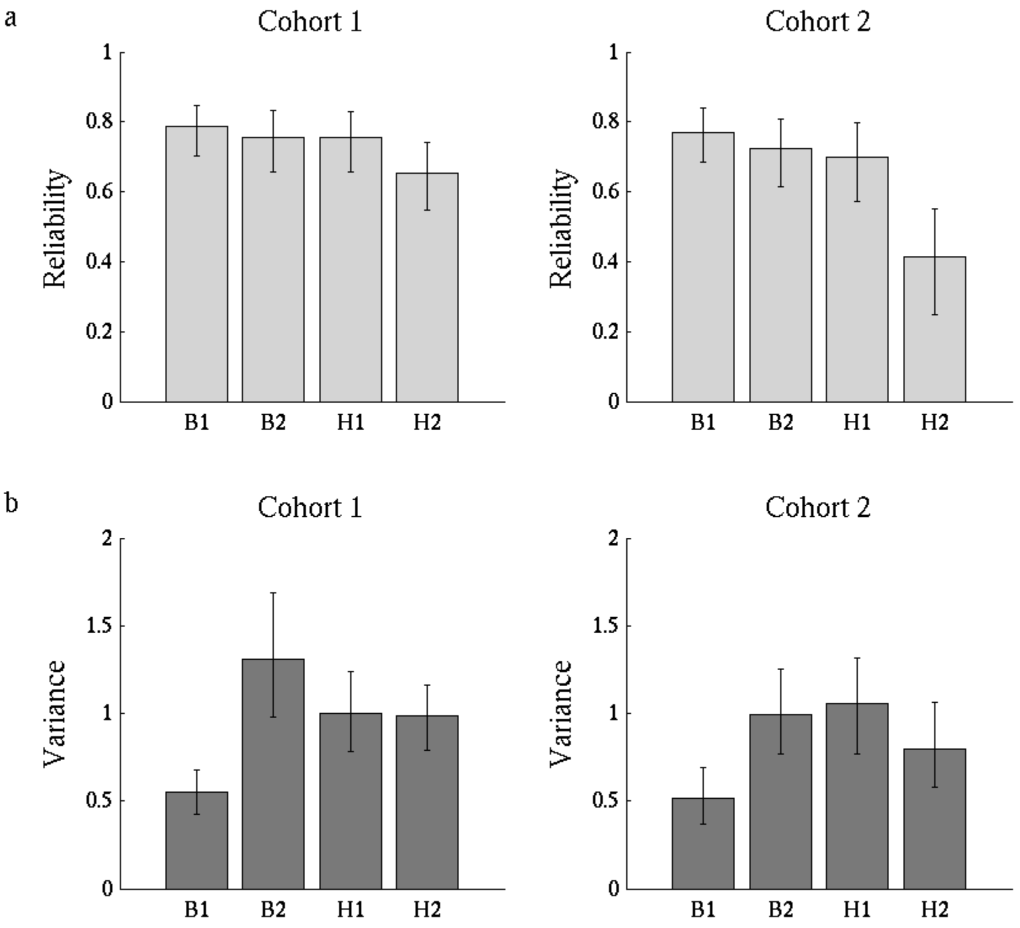

Figure S1. Reliability and variance of acoustic test measures. **(A)** Reliability of acoustic test measures for each cohort. Error bars here and in (B) are 95% confidence intervals. **(B)** Variance of acoustic test measures. See Fig. S4 for description of B2 measure, and Fig. S2 for analyses involving B2 measure.

Alternative measure of beating preference

Because the dichotic manipulation of the high frequency pure tone pairs (see Fig. 2b) produced the largest effect (see Fig. 2d), we constructed an alternative measure of beating preference just using the high frequency tones. This measure was just the difference between the average ratings of the dichotic and diotic presentations of these tone pairs. As shown in Fig. S2, this alternative measure again showed weak and inconsistent correlations with consonance preferences. Like the B1 measure, it explained little of the variance in consonance preferences (compare to Fig. 3b). There was also no significant correlation with musical experience (compare to Fig. 4a).

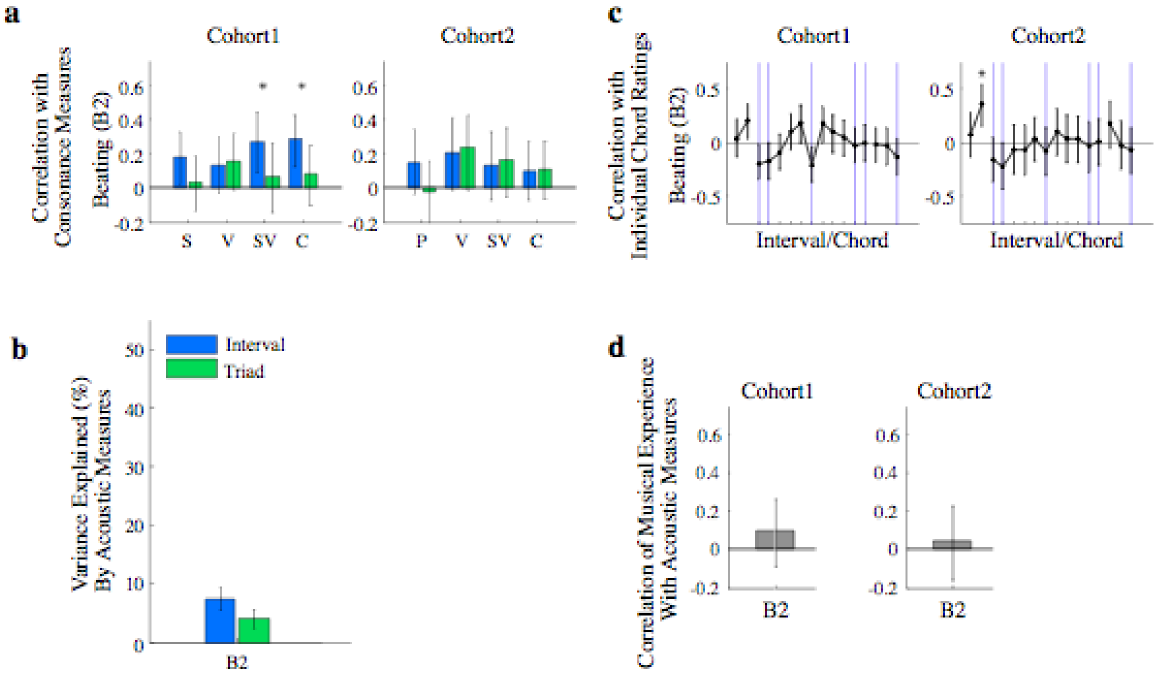

Figure S2. Correlations of alternative measure of beating preference (B2) with consonance preferences. **(A)** Correlations with interval (blue) and triad (green) consonance measures. Letters on x-axis denote note timbre (saxophone (S), sung vowels (V), synthetic sung vowels (SV), synthetic complex tone (C), pure tone (P)). Here and in (C), error bars denote 95% confidence intervals, and asterisks denote significance (.05 criterion). **(B)** Variance of consonance measures explained by acoustic preferences. Error bars denote standard errors. **(C)** Correlations with ratings of individual chords, averaged across note timbre. Interval/chord arrangement within subpanels follows conventions of Figure 1a. Blue vertical lines denote dissonant intervals and triads included in consonance measures. **(D)** Correlations with the number of years spent playing an instrument.

Analyses of Figure 3 repeated without z-scoring chord ratings

Fig. S3 is the result of regenerating Figure 3 using chord ratings that were not z-scored. All the main effects of interest remain visible. The correlations are somewhat larger, indicating a contribution from variation in the use of the rating scale (underscoring the importance of using z-scored ratings to ensure that this is not driving our effects).

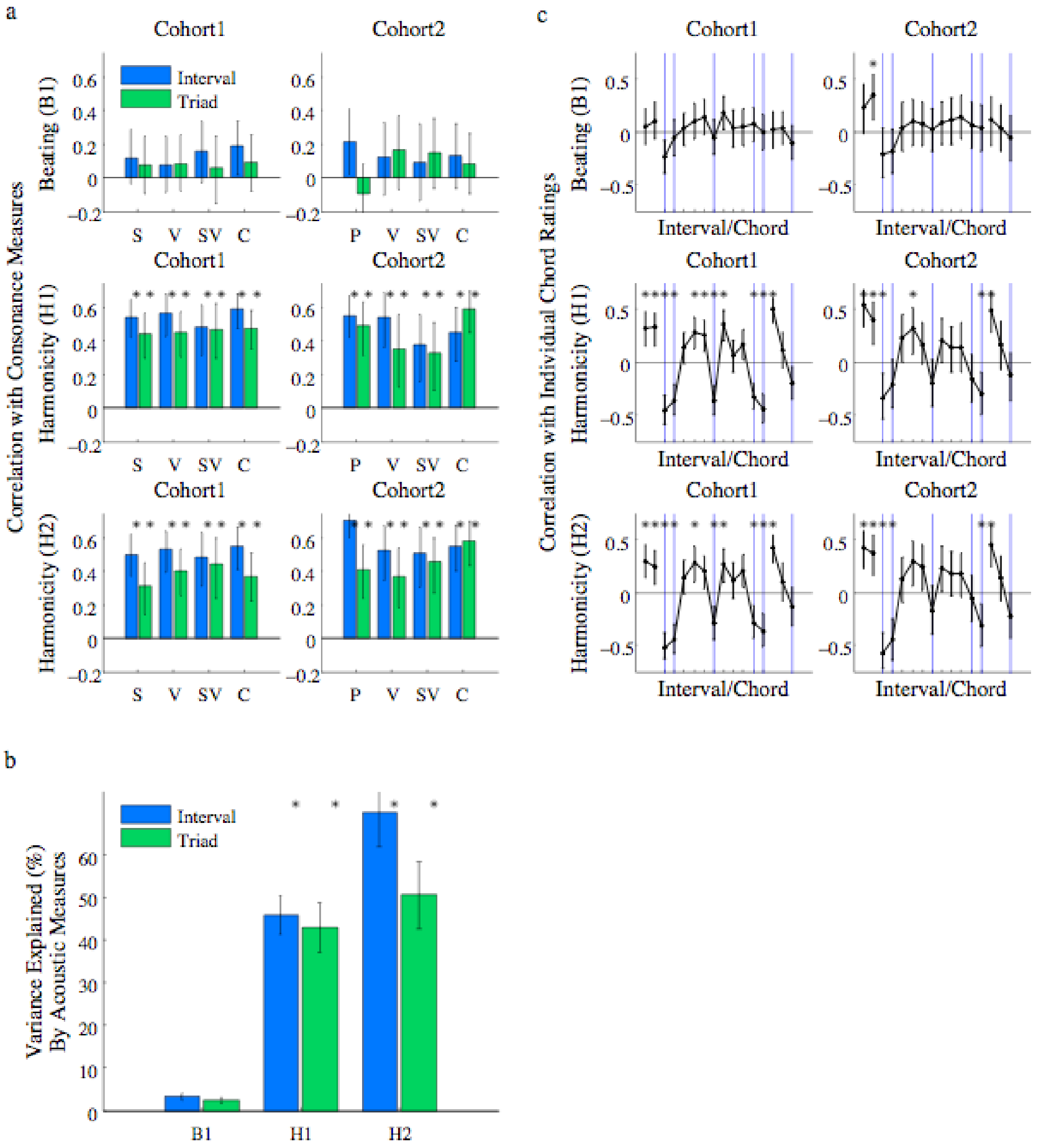

Figure S3. Analyses of Figure 3 repeated without z-scoring chord ratings. **(A)** Correlations of beating and harmonicity measures with interval and triad consonance measures computed from untransformed mean ratings. **(B)** Variance of consonance measures explained by acoustic measures. Note that y-axis is expanded relative to that of Figure 3b. **(C)** Correlations of beating and harmonicity measures with untransformed ratings of individual chords, averaged across note sounds. All other conventions are as in Figure 3.

Partial correlation analysis of harmonicity and musical experience.

Fig. S4 presents the effect of partialing out either harmonicity preferences or musical experience. Each decreases but does not eliminate the correlation of the other with consonance preferences. This suggests that part of the influence of harmonicity on consonance is mediated by musical experience, and vice versa. The residual correlations after musical experience is partialed out are consistent with a contribution of harmonicity to consonance that is independent of musical experience, but could also simply reflect the crude and imperfect nature of our measure of musical experience.

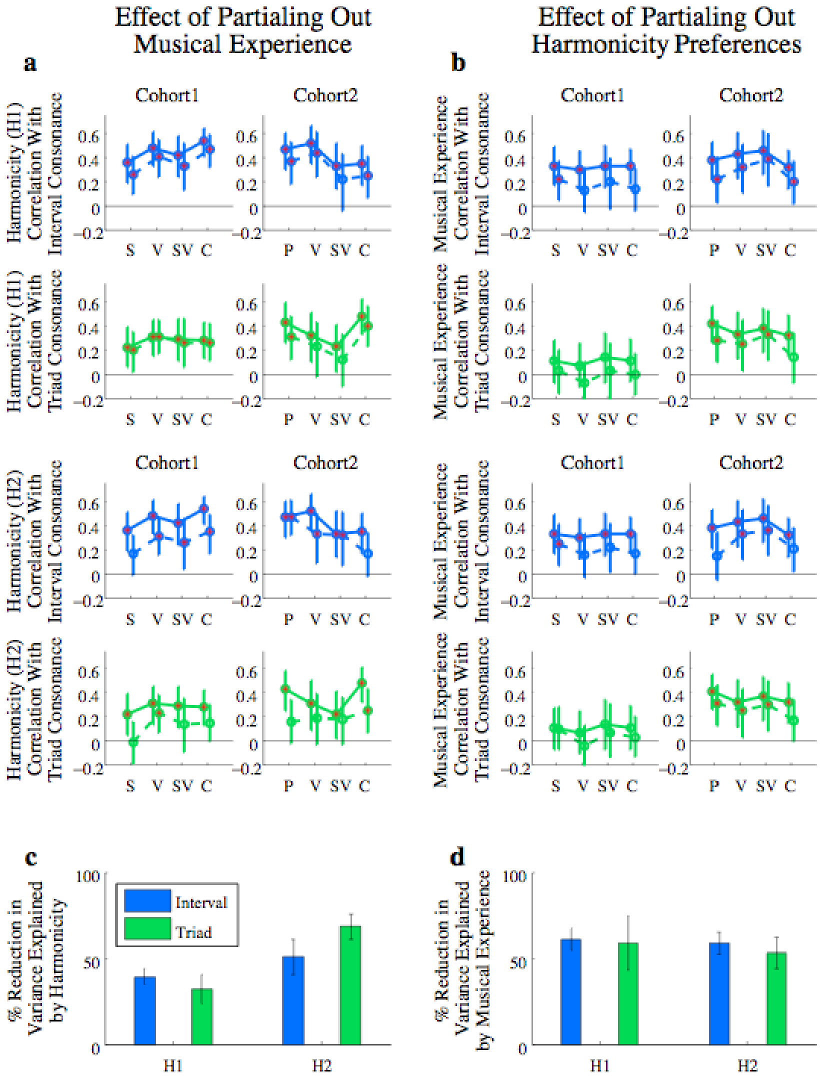

Figure S4. Partial correlation analysis of harmonicity and musical experience. **(A)** Correlations of the harmonicity preference measures with consonance preference measures, with (dashed lines) and without (solid lines) musical experience partialed out. Here and in (B), error bars denote 95% confidence intervals, and red filled symbols indicate significant correlations (.05 criterion). **(B)** Correlations of musical experience with consonance preference measures, with (dotted lines) and without (solid lines) the H1 or H2 measures partialed out. **(C)** Amount by which the variance of the consonance measures explained by the harmonicity preference measures was decreased when musical experience was partialed out. Here and in (D), error bars are standard errors. **(D)** Amount by which the variance of the consonance measures explained by musical experience was decreased when the harmonicity preference measures were partialed out.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Participants

Both cohorts averaged 20.2 years of age. Cohort 1 (143 subjects; 95 females) and Cohort 2 (122 subjects; 79 females) had spent an average of 5.4 and 5.7 years, respectively, playing an instrument or singing.

Cohort 1 subjects completed the saxophone, sung vowel, and complex tone tests in random order, followed by the synthetic sung vowel test if time permitted (95 out of 143 subjects). Cohort 2 first completed one pair of tests with pure and complex tone chords, and then did a pair with natural and synthetic sung vowels if time permitted (85 out of 122 subjects); some came back for a second session to complete this latter pair of tests.

Stimuli

The three types of inharmonic test stimuli were modifications of the harmonic test stimuli, as shown in Figure 2c. In the first, components were separated by 13 semitones instead of 12 as in the first harmonic complex tone. In the second, the frequencies of the second harmonic tone were perturbed up (even components) or down (odd components) by .5 semitones. In the third (Cohort 1 only), the frequencies of all the components in second harmonic tone were increased by 30 Hz. All harmonic and inharmonic test stimuli had components whose amplitudes decreased by 14 dB per octave to resemble naturally occurring sounds.

Inharmonicity detection thresholds for a single mistuned frequency component in a harmonic complex have been previously estimated to be about 1% of the harmonic's frequency [1, 2]. Our inharmonic perturbations were well in excess of this (between ~6-19%, depending on the harmonic, for the stretched condition, ~3% for the jittered condition, and between ~1.5-10%, depending on the harmonic, for the shifted condition). Each of the perturbations thus produced clearly audible inharmonicity. We also computed harmonic-to-noise ratios (HNRs) for the stimuli. Unfortunately, most methods for doing this are designed to distinguish tonal frequency components from noise (e.g. in voiced vs unvoiced speech, or vocal hoarseness [3]), and are poorly suited to comparing harmonic and inharmonic tones. For instance, the well-known HNR algorithm in PRAAT [3] uses the largest local maximum in the signal autocorrelation function, which is high for periodic signals. For harmonic signals with added noise, such as speech, this method works well, but inharmonic tones often have at least one peak near 1 despite being aperiodic, producing artifactually high HNRs with standard methods. Inharmonic tones are more clearly distinguished from harmonic tones by the absence of regularly spaced peaks of consistent height in the autocorrelation, and thus methods using a harmonic sieve may be more appropriate [4]. In the absence of an accepted standard alternative, we nonetheless computed HNRs using the standard PRAAT algorithm [3]. Despite its questionable appropriateness for our purposes, this measure showed a clear contrast between the harmonic and inharmonic test stimuli that we used, with HNR differences in excess of 20 dB. The inharmonic stimuli yielded HNRs of 15.71, 21.32, and 20.85 dB for the stretched, jittered, and shifted conditions, respectively, vs. 44.69 and 44.25 dB for the two harmonic complex tones. The stimuli used in the H2 harmonicity measure produced similar results (20.59 and 27.27 dB for the .75 and 1.5 semitone intervals, respectively, vs. 47.6 dB for the pure tone). The development of methods better suited to differentiating harmonic and inharmonic tones, and that take into account the properties of biological auditory systems (for instance, the different effects of resolved and unresolved harmonics), will be a useful future research project. Such methods would permit precise comparisons of perceptual judgments of consonance and the perceptually relevant harmonic content of chords and other stimuli.

The base frequency of each acoustic test stimulus was roved across trials, drawn without replacement from the 6-semitone range above middle C, with the exception of the tone pairs for the beating test measures – the base frequencies of the mid- and high-range pairs were 13-18 and 25-30 semitones above middle C, respectively (one and two octaves higher than the low-frequency pairs).

Chords and intervals were derived from the equal-tempered scale. All chords had a root pitch drawn from the 8-semitone range above middle C (G#3 for the saxophone and sung notes, to accommodate the range of the singer/instrument), except for the high-pitched single notes from the chord experiments (13-20 semitones above middle C or G#3). All synthetic stimuli (acoustic test stimuli and chords) had envelopes that were the product of a half-Hanning window (10 ms) and a decaying exponential (decay constant of 2.5 sec^{-1}) that was truncated at 2 sec. All stimuli were presented at 65 dB SPL.

The saxophone chords were generated from soprano saxophone notes, recordings of which were downloaded from the University of Iowa Musical Instrument Samples website (http://theremin.music.uiowa.edu/MIS.html). These were edited and in a few cases pitch-corrected using PRAAT (http://www.fon.hum.uva.nl/praat/).

Sung notes (aah /a/ and ooh /u/ vowel sounds) were generated and recorded by the second author (a professional singer), in the same pitch range as the notes used in the other experiments. They were high-pass filtered with a cutoff frequency of 150 Hz (4th order Butterworth filter) to eliminate handling noise from the microphone. For Cohort 1, we flattened the pitch contours of the recorded sung notes using PRAAT, and used only the aah sounds; the stimuli for Cohort 2 were unaltered, and included both ooh and ahh sounds (half the trials used the ooh sounds, and half the ahh), the ratings for which were averaged and treated as a single condition in the analysis.

For both the saxophone and sung vowels, the amplitude envelope of the note onset was not altered apart from applying a half-Hanning window (10 ms). A linear ramp was applied over the last half of each note to fade them smoothly down to zero, such that the duration was the same as the synthetic notes.

Synthetic sung vowels were generated by applying an average spectral envelope for each vowel type (derived from our sung notes) to a synthetic harmonic complex tone. The smoothed average spectral amplitude envelope of the sung vowels was applied to this stimulus. In practice this produced high-frequency amplitudes that were higher in the synthetic vowels than in the sung vowels, due to the presence of breath noise in the sung vowels that elevated the high-frequency portion of the spectral envelope. A linear ramp was applied over the last half of each note, to give them the same temporal envelope as the sung notes.

Procedure

Stimuli were delivered using headphones (Sennheiser HD 580) in a double-walled sound-attenuating chamber (Industrial Acoustics Corp). All trials of all experiments began with a 2 sec burst of white noise (60 dB SPL), intended to reduce the influence of successive trials on each other. In addition to the chords described in Figure 1, the chords tests also included the major seventh, dominant seventh, and minor seventh chords; these were not included in the analysis. The acoustic tests also included several other test stimuli that were not analyzed in this study.

In each test, subjects were presented with stimuli in random order, with multiple repetitions of each stimulus (3 for the acoustic tests, 4 for the chord tests), each time with a different root pitch. All subjects completed two acoustic tests, containing both beating and harmonicity test stimuli, followed by pairs of the chord rating tests, each pair featuring a different instrument sound.

Following each trial subjects entered a rating between –3 and 3, denoting the range from very unpleasant to very pleasant. Subjects were instructed to use the full rating scale. Before starting the tests, subjects were given a short practice test to familiarize them with the range of stimuli they would encounter.

Pilot Experiment

The procedure for the pilot tone-pair experiment (Figure 2a) was identical to that of the acoustic experiments except that there was a single test, and the base frequency of the pure tone pairs was either 523 or 740 Hz. Subjects (N=35) heard 3 repetitions of each stimulus, in random order.

Analysis

The test measures were computed by subtracting the average of the ratings for one set of stimuli from the average of the ratings of a second set. The stimuli contributing to each measure are shown in Fig. S5.

Spearman correlation coefficients and two-tailed significance tests were used throughout. Correction for multiple comparisons (modified Bonferroni) was performed on all sets of multiple statistical tests. 95% confidence intervals were derived from bootstrap analysis with 1000 repetitions.

For the analysis of Figures 3a, 3b, and 4b, correlations were computed from the scores of all the subjects who had participated in each chord experiment. For the analyses of Figures 3c, in which chord ratings were averaged across timbre prior to computing correlations, only those subjects who had completed all who gave ratings for all four timbres were included.

We wanted to ensure that the correlations we measured were not merely reflections of variations in how different subjects used the rating scale (for instance, some subjects might have used the full scale, while others might have restricted their ratings to a narrow range; this would introduce test-retest correlations in our test measures). Accordingly, we z-scored the mean ratings given to each chord within each subject, such that the variance of the chord ratings was the same for each subject. We also used this z-scoring procedure on the acoustic test ratings to confirm that the individual differences we observed were not merely due to differences in the use of the rating scale (the r values reported in Figure 2e-g and in the text were computed from z-scored mean ratings). However, we did not z-score the acoustic ratings prior to computing correlations with the consonance measures and chord ratings, as z-scoring one of the two was sufficient to remove correlations in rating variance, and it seemed desirable to manipulate the data as little as possible. We also did not use z-scored acoustic ratings when calculating the test-retest reliabilities used in the variance estimates of Figure 3c, to remain consistent with how the correlations were computed. To ensure that this choice was not critical to our results, we reran the same analysis omitting the z-scoring procedure, after applying the z-scoring to both diagnostic and chord ratings, and after z-scoring the diagnostic ratings but not the chord ratings. Although there were various subtle differences in the correlation magnitudes, none of the paper's conclusions would have been different had we used one of these alternate procedures. We have included the results of the correlation analysis on the ratings without z-scoring in the Supplementary Results (Fig. S3). The main effect of omitting the z-scoring procedure is that all the correlations are slightly higher. We attribute this to a correlation between the rating range used by subjects in the diagnostic and chord rating experiments. It is apparent from the figure that all the trends of interest are evident whether or not z-scoring is used.

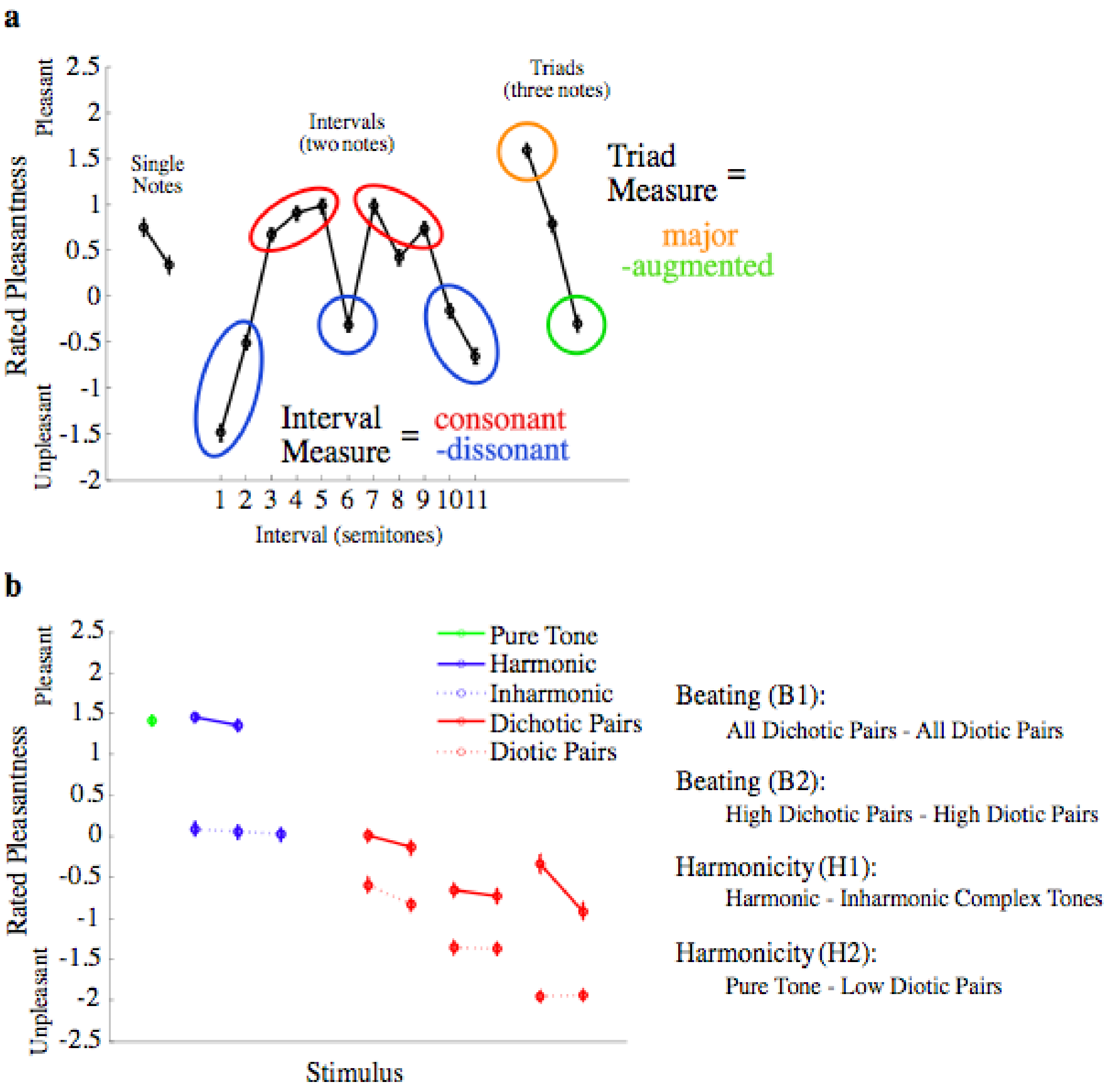

Figure S5. Test measure definitions **(A)** Definition of consonance measures. Solid black line is the mean across note timbres of the chord ratings for Cohort 1. **(B)** Definition of acoustic measures. Ratings data for Cohort 1 is included for reference.

Supplemental References

1. Moore, B.C.J., Glasberg, B.R., and Peters, R.W. (1986). Thresholds for hearing mistuned partials as separate tones in harmonic complexes. Journal of the Acoustical Society of America *80*, 479-483.
2. Grube, M., von Cramon, D.Y., and Rubsamen, R. (2003). Inharmonicity detection: Effects of age and contralateral distractor sounds. Experimental Brain Research *153*, 637-642.
3. Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. IFA Proceedings *17*, 97-110.
4. Tramo, M.J., Cariani, P.A., Delgutte, B., and Braida, L.D. (2001). Neurobiological foundations for the theory of harmony in Western tonal music. Annals of the New York Academy of Science *930*, 92-116.