

Losing the Music: Aging Affects the Perception and Subcortical Neural Representation of Musical Harmony

 Oliver Bones and Christopher J. Plack

School of Psychological Sciences, University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom M13 9PL

When two musical notes with simple frequency ratios are played simultaneously, the resulting musical chord is pleasing and evokes a sense of resolution or “consonance”. Complex frequency ratios, on the other hand, evoke feelings of tension or “dissonance”. Consonance and dissonance form the basis of harmony, a central component of Western music. In earlier work, we provided evidence that consonance perception is based on neural temporal coding in the brainstem (Bones et al., 2014). Here, we show that for listeners with clinically normal hearing, aging is associated with a decline in both the perceptual distinction and the distinctiveness of the neural representations of different categories of two-note chords. Compared with younger listeners, older listeners rated consonant chords as less pleasant and dissonant chords as more pleasant. Older listeners also had less distinct neural representations of consonant and dissonant chords as measured using a Neural Consonance Index derived from the electrophysiological “frequency-following response.” The results withstood a control for the effect of age on general affect, suggesting that different mechanisms are responsible for the perceived pleasantness of musical chords and affective voices and that, for listeners with clinically normal hearing, age-related differences in consonance perception are likely to be related to differences in neural temporal coding.

Key words: aging; auditory brainstem; frequency-following response; musical consonance

Introduction

A musical note is an example of a complex tone. A complex tone consists of a fundamental frequency (F0) and a number of harmonics with frequencies at integer multiples of the F0. When a two-note chord (a “dyad”) is created using notes with simple F0 ratios, the resulting perception of resolution and stability that occurs is described as “consonance.” The perception of consonance is likely driven by the extent to which the combined harmonics of a consonant dyad share a common implied F0 and the extent therefore that the combined spectrum resembles that of a single complex tone (McDermott et al., 2010; Fig. 1A,B). The harmonic structure of a dyad is temporally coded in the inferior colliculus (IC), the principal auditory midbrain nucleus, by the tendency for neurons to fire at a particular phase of the individual low-numbered harmonics and the temporal envelope produced by the interaction of harmonics in the cochlea (Brugge et al., 1969; Bidelman and Krishnan, 2009; Lee et al., 2009; Bones et al., 2014). This neural temporal coding of harmonic structure pre-

dicts individual differences in the perception of consonance in young, normal hearing listeners (Bones et al., 2014).

Aging has a detrimental effect on temporal processing that is independent of age-related clinical hearing loss (Strouse et al., 1998; Clinard et al., 2010; Hopkins and Moore, 2011; Neher et al., 2012). This may be partly due to age-related changes in the response properties of neurons in the IC. Single units in old mice spike more frequently and code temporal fluctuations less accurately than young mice (Walton et al., 2002; Frisina and Walton, 2006) and units that respond with rapid onset in young mice respond with greater latency (Walton et al., 1997) and with a longer recovery period (Walton et al., 1998). This age-related loss of temporal precision is consistent with a decline in neural inhibition (Caspary et al., 2005) and may be the consequence of downregulation of the inhibitory neurotransmitter GABA (Milbrandt et al., 1994). GABA receptor agonists broaden the frequency response of IC neurons in chinchillas (Caspary et al., 2002) and have a deleterious effect on phase locking of IC neurons in bats (Koch and Grothe, 1998; Klug et al., 2002), suggesting that normal spectrotemporal neural response properties may be dependent on the sharpening provided by GABAergic inhibition.

Age-related deficits in temporal processing may also be a consequence of deafferentation. Afferent auditory nerve spiral-ganglion cells become decoupled from the inner hair cells of the cochlea with normal aging (Makary et al., 2011; Sergeev et al., 2013) and after glutamate excitotoxicity resulting from noise trauma (Puel et al., 1994). Elevation of hearing threshold after noise trauma is usually temporary, with a time course corresponding to synaptic recovery time (Puel et al., 1998). However, animal models demonstrate that deafferentation persists even when auditory thresholds return to prenoise levels (Kujawa and

Received Aug. 4, 2014; revised Dec. 10, 2014; accepted Dec. 15, 2014.

Author contributions: O.B. and C.J.P. designed research; O.B. performed research; O.B. and C.J.P. analyzed data; O.B. and C.J.P. wrote the paper.

This work was supported by the Economic and Social Research Council (Grant ES/J500094/1).

The authors declare no competing financial interests.

This article is freely available online through the *J Neurosci* Author Open Choice option.

Correspondence should be addressed to Dr. Oliver Bones, G36, Leung Kau Kui Building, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong. E-mail: olly@cuhk.edu.hk.

DOI:10.1523/JNEUROSCI.3214-14.2015

Copyright © 2015 Bones and Plack

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

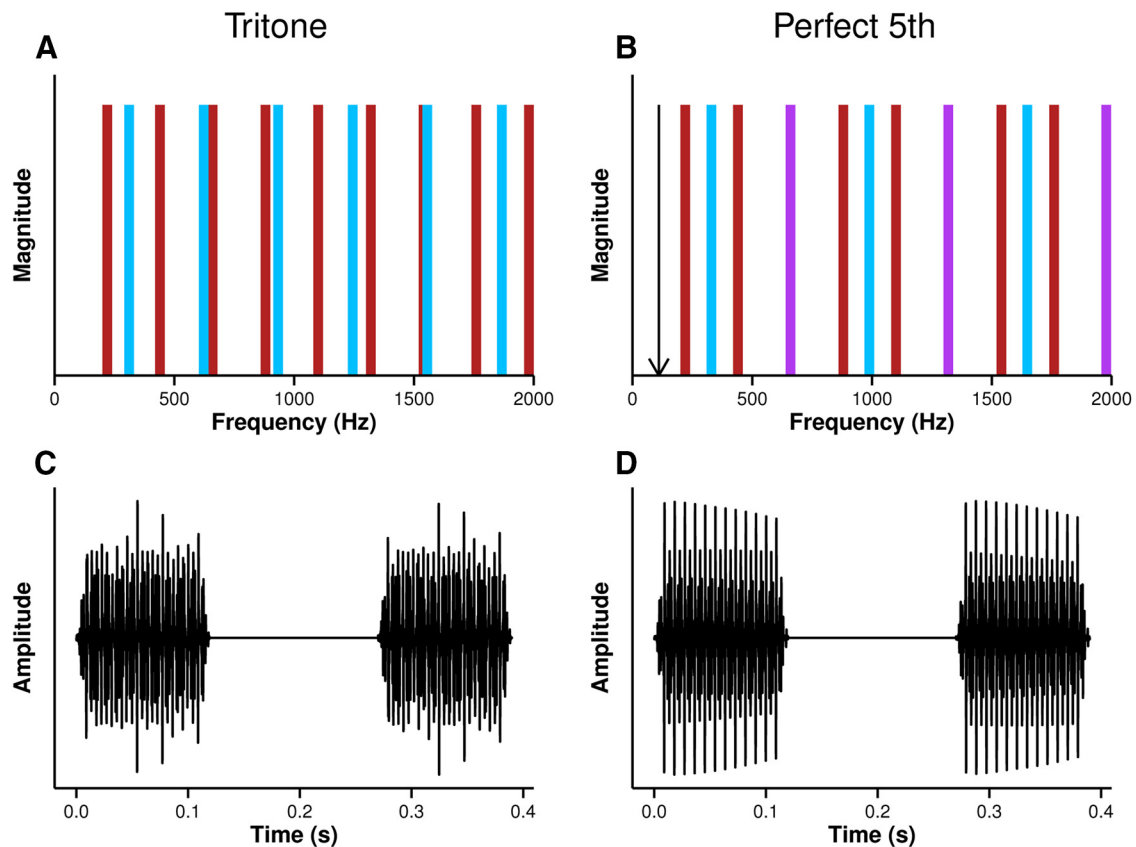


Figure 1. Each dyad is created by combining a low note (red bars) and a high note (blue bars). Each note is a complex tone with an F0 and harmonics at integer multiples of the F0. In the case of the dissonant tritone (**A**), the harmonics of the combined spectrum are irregularly spaced and therefore inharmonic. In the case of the consonant perfect fifth, the harmonics of the combined spectrum are regularly spaced and in some cases perfectly coincide (violet bars). The combined harmonics form a harmonic series such that the combined notes resemble a single note. The arrow indicates the implied F0 of the harmonic series (**B**). During recording of the FFR, each stimulus presentation consisted of two dyads separated by 150 ms silence (**C,D**). The waveform of the tritone (**C**) is irregular, whereas the harmonic relation of the frequencies of the perfect fifth results in a waveform with regular peaks (**D**).

Liberman, 2009; Furman et al., 2013), suggesting that deafferentation may go undetected and occur over the lifespan.

The present study expands upon previous work by Bones et al. (2014) by testing the hypothesis that age is associated with a decline in the distinction between different musical dyads in terms of the neural temporal representation at the level of the IC and that this deficit is associated with a decline in the perception of consonance.

Materials and Methods

Participants. All participants had clinically normal hearing thresholds (≤ 20 dB HL) at octave frequencies between 250 and 2000 Hz. No participants reported having “perfect pitch” (“can you recreate a given musical note without an external reference?”) or having a history of speech or language difficulties. Participants were nonmusicians, having not received formal music training within the past 5 years. When participants had received music training, an estimate of their total training in hours was logged and controlled.

Forty-four (27 female) participants 18–81 years of age (mean = 38.0, SD = 20.5) participated in the behavioral section of the study. The distribution of age is displayed in Figure 2. Group comparisons were made between 28 participants <40 years of age (mean = 23.7 years, SD = 6.1) and 16 participants >40 years of age (mean = 63.1 years, SD = 9.1). The mean hearing threshold between 250 and 2000 Hz in the younger group was 4.66 dB HL (SD = 4.42 dB), the mean of the older group was 10.26 dB HL (SD = 3.62 dB). The mean music experience of the younger group was 434 h (SD = 702), the mean of the older group was 260 h (SD = 413). One participant withdrew from the older group before electrophysiological testing.

Behavioral stimuli. Dyads were created by combining each of eight low notes with 11 high notes in the octave above, making 88 dyads in total. Each dyad is named after the ratio of the F0s of the low and the high notes (the “interval”; Table 1). Low notes were D3 and seven notes above (Table 2).

Dyads were diotic (both notes were presented to both ears). Each note consisted of equal amplitude sinusoidal harmonics with their levels set so that the overall level of each note was 72 dB SPL when low-pass filtered at 2200 Hz and the overall level of each dyad was 75 dB SPL. Dyads had a duration of 2000 ms, including 10 ms raised-cosine onset and offset ramps. Each dyad was preceded by Gaussian noise with the same duration and filtering. A 500 ms silence separated the noise and the dyad. The purpose of the noise was to prevent randomly generated melodic sequences from influencing pleasantness judgments (McDermott et al., 2010; Bones et al., 2014).

To test for the possibility that older listeners rate dyads differently due to having a general lack of affect or an inability to perform the task rather than due to a decline in consonance perception, participants also rated diotic affective voices for their perceived pleasantness (Cousineau et al., 2012). Voices were a subset of the Montreal Affective Voices (MAV) battery (Belin et al., 2008; <http://vnl.psy.gla.ac.uk/>). The MAV battery is a collection of recordings of actors expressing different emotions without using spoken words. Four affective voices (two male and two female) from the categories “fear,” “happiness,” “sadness,” and “surprise” were used.

Behavioral procedure. Behavioral methodology reported previously (McDermott et al., 2010; Cousineau et al., 2012; Bones et al., 2014) was used to record individual pleasantness ratings for dyads and affective voices. Participants rated each sound for its pleasantness from −3 to 3,

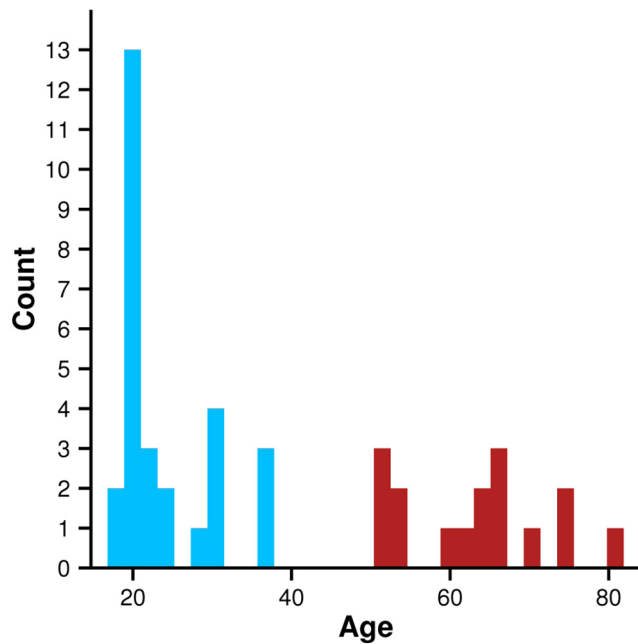


Figure 2. Distribution of participant ages.

where -3 indicated very unpleasant and 3 indicated very pleasant. Individual differences in the perceptual distinction of different musical dyads were measured by calculating consonance preference scores. This was calculated by first z-scoring each participant's rating for each dyad to account for individual use of the ratings scale and then subtracting the average z-scored rating of the five most theoretically dissonant intervals from the average z-scored rating of the five most theoretically consonant dyads (McDermott et al., 2010; Cousineau et al., 2012) determined a priori by Western musical tradition (Rameau, 1971; Mickelson and Riemann, 1977). The ranking of dyads by consonance is displayed in Table 1. Due to an outlier in the consonance preference scores, correlations are reported with 95% confidence intervals (CIs) from 2000 bootstrapped samples (Wright and Field, 2009). A happiness preference score was calculated for comparison with consonance preference by subtracting the mean of each participant's z-scored rating of "fear" and "sadness" (the two lowest rated affective voices) from their z-scored rating of "happiness" (the highest rated affective voice).

Participants sat in a sound-attenuating booth and made ratings via a keyboard with a computer monitor outside the booth visible through a window. Stimuli were presented via Sennheiser HD 650 supra-aural headphones. Before testing, participants performed one practice block of one of each dyad and each affective voice. After the practice block, participants performed two runs, each consisting of the 88 dyads and two of each affective voice category presented in a random sequence. The two runs were performed consecutively on the same day. Ratings for each dyad were averaged across root notes and runs so that each participant's pleasantness rating of each dyad was the average of 16 responses and each affective voice pleasantness rating was the average of eight responses.

Frequency-following response: recording protocol and stimuli. The frequency-following response (FFR) is a scalp-recorded event-related potential representing the sustained, phase-locked response of populations of neurons in the auditory brainstem to a sustained stimulus (Moushegian et al., 1973). Periodic peaks in the waveform correspond to phase locking to individual frequency components in the stimulus and to the temporal envelope of the basilar membrane response resulting from the interaction of harmonics in the cochlea (for review, see Krishnan, 2007). In the present study, a vertical electrode montage was used (Krishnan et al., 2005; Bidelman and Krishnan, 2009; Gockel et al., 2011; Krishnan and Plack, 2011; Gockel et al., 2012; Bones et al., 2014): an active electrode was positioned at the high forehead hairline, a reference electrode was positioned at the seventh cervical vertebra, and a ground elec-

Table 1. Dyads created by combining a low note with each of the 11 notes above it

Rank	F0 ratio	Name	Abbreviation	Consonance
1	1.498307	Perfect 5th	P5	Perfectly consonant
2	1.33484	Perfect 4th	P4	
3	1.681793	Major 6th	M6	Imperfectly consonant
4	1.259921	Major 3rd	M3	
5	1.587401	Minor 6th	m6	
6	1.189207	Minor 3rd	m3	Dissonant
7	1.781797	Minor 7th	m7	
8	1.414214	Tritone	Tri	
9	1.887749	Major 7th	M7	
10	1.122462	Major 2nd	M2	
11	1.059463	Minor 2nd	m2	

Combining a low note with each of the 11 notes above it creates a different dyad, named after the ratio ("interval") of the two notes. Intervals can be categorized as "perfectly consonant," "imperfectly consonant," and "dissonant." The hierarchical ranking of intervals contributes to a sense of musical "key." Intervals are ranked here by consonance in accordance with Western music theory, and with the pleasantness ratings of the younger group.

Table 2. The eight low notes used to create dyads

Note	F0 (Hz)
D3	146.83
D#3	155.56
E3	164.81
F3	174.61
F#3	185.00
G3	196.00
G#3	207.65
A3	220.00

Each low note was combined with 11 high notes to create the 11 intervals in Table 1.

trode was positioned at Fpz. Impedences were checked at regular intervals and maintained <3 k Ω throughout.

Participants were seated in a comfortable, reclining armchair within a sound-attenuating booth. They were instructed to keep their arms, legs, and neck straight and to keep their eyes closed and to sleep if possible. Stimuli were delivered via a TDT RP2.1 Enhanced Real Time Processor and HB7 Headphone Driver, and Etymotic ER30 insert headphones. The length of the headphone tubing made it possible to position the transducers outside of the booth, preventing stimulus artifacts from contaminating the recording.

Two diotic dyads were used to record the FFR. A low note A3 was combined with the high notes D#4 and E4 to create a tritone and a perfect fifth dyad, respectively. Dyads were 120 ms in duration, including 10 ms onset and offset raised-cosine ramps, and were filtered and had the same level as in behavioral testing.

The FFR to each dyad was collected separately. Presentations consisted of two repetitions of the dyad separated by 150 ms silence (Fig. 1C,D). As in previous studies, the starting phase of the second dyad in each presentation was inverted by 180 degrees relative to the first dyad (Gockel et al., 2011; Bones et al., 2014). This method allows for selectively enhancing the neural response to either the temporal fine structure or the envelope of the BM response by subtracting or adding the response to the two stimulus polarities, respectively (Goblick and Pfeiffer, 1969). However, in the present study, this technique was not used: the spectra that are analyzed are the average of the spectra of the responses to the two stimulus polarities and are therefore calculated in the same way as the FFR_{RAW} spectra reported by Bones et al. (2014).

Presentations were made at a rate of 1.82 Hz; that is, successive pairs of dyads were separated by 160 ms silence. Responses were recorded during an acquisition window with 447 ms duration from the presentation onset. Responses were filtered online between 50 and 3000 Hz. Four thousand responses to each dyad were collected from each participant, compiled online in subaverages of 10 acquisitions. Any subaverage in which the peak amplitude exceeded ± 30 μ V was assumed to contain an artifact and rejected offline.

FFR analysis. To determine how well the harmonic structure of each dyad was represented by neural temporal coding, we used a procedure

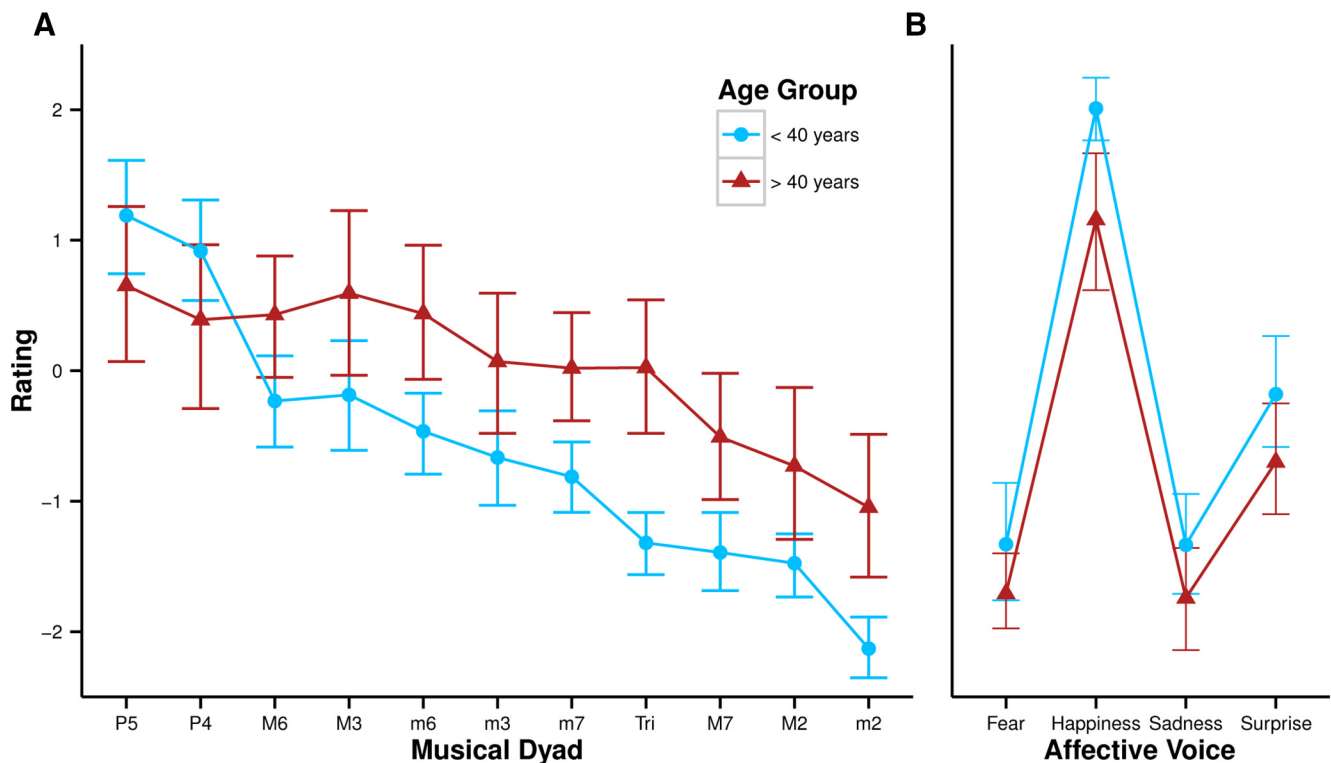


Figure 3. Pleasantness ratings for musical dyads (**A**) and affective voices (**B**) by each age group. Error bars indicate 95% confidence intervals. Dyads on the x-axis are arranged by theoretical consonance from most consonant (P5) to least consonant (m2; Table 1).

described previously (Bones et al., 2014). First, the F0s of the best-fitting harmonic series to the power spectra of the two dyads were calculated. The salience of harmonic series with F0s ranging from 30 to 1000 Hz were analyzed by measuring the power inside 15 Hz wide bins placed at integer multiples of the F0. The fit of each harmonic series was calculated as the ratio between the sum of power inside the bins to the sum of power outside the bins. The F0s of the best-fitting harmonic series were 44.25 and 55 Hz for the tritone and the perfect fifth, respectively. Second, the representation of the harmonic series with these F0s in the power spectrum of the FFR (hereafter referred to as “neural harmonic salience”) were calculated in the same way: the ratio between power inside and outside 15 Hz wide bins placed at integer multiples of the F0 of the best-fitting harmonic series to the dyad.

The Neural Consonance Index (NCI; Bones et al., 2014) is a measure of the harmonic salience of the FFR to consonant dyads relative to dissonant dyads and is a neural analog of the consonance preference measure. Each participant’s NCI score was calculated by subtracting the neural harmonic salience of the FFR to the tritone (a dissonant dyad) from the neural harmonic salience of the perfect fifth (a consonant dyad).

Results

Effect of age on consonance perception

Pleasantness ratings by the two groups for each dyad are displayed in Figure 3A. The average SD of each participant’s rating for each dyad was not significantly different between groups ($t = 0.799$, $p = 0.43$). Musical dyad had a large effect on pleasantness ratings ($\epsilon = 0.32$, $F_{(3.2,134.4)} = 91.11$, $p < 0.001$, $\eta^2 = 0.388$). Pleasantness ratings by the younger group were consistent with data reported previously (Plomp and Levelt, 1965; Tufts et al., 2005; Bidelman and Krishnan, 2009; McDermott et al., 2010; Bones et al., 2014) and were in agreement with theoretical denominations (Table 1): the perfectly consonant dyads (perfect fourth and perfect fifth) were rated as being most pleasant, followed by the imperfectly consonant dyads (minor third, major

third, minor sixth, and major sixth), followed by the dissonant dyads (minor second, major second, tritone, minor seventh, and major seventh). However, age was also an effect ($F_{(1,42)} = 5.58$, $p = 0.02$, $\eta^2 = 0.086$) and interacted with the effect of musical dyad ($\epsilon = 0.32$, $F_{(3.2,134.4)} = 11.35$, $p < 0.001$, $\eta^2 = 0.388$) in that the relative ratings of the intervals depended on age group. The plot of pleasantness ratings by the younger group as a function of theoretical consonance is steep, with a sharp transition from the consonant dyads to the imperfectly consonant dyads to the dissonant dyads. The plot for the older group is shallow in comparison, with less distinction between dyads. Consider, for example, the difference in pleasantness rating of the perfectly consonant perfect fourth and the dissonant tritone in the older group compared with the younger group. In Bonferroni-corrected ($\alpha = 0.005$) Welch two-sample t tests and Wilcoxon rank-sum tests, the older group rated the dissonant dyads the minor seventh ($t_{(27.4)} = 3.08$, $p = 0.004$), tritone ($t_{(21.7)} = 4.63$, $p < 0.001$), and minor second ($t_{(20.3)} = 3.47$, $p = 0.002$) as significantly more pleasant than did the younger group. Other pairwise comparisons were not significant with the corrected α .

Dyad category (perfectly consonant; imperfectly consonant; dissonant; Table 1) was also a significant effect ($\epsilon = 0.85$, $F_{(1.7,71.4)} = 114.13$, $p < 0.001$, $\eta^2 = 0.384$) that interacted with the effect of age ($\epsilon = 0.85$, $F_{(1.7,71.4)} = 18.97$, $p < 0.001$, $\eta^2 = 0.094$). In Bonferroni-corrected ($\alpha = 0.017$) Welch two-sample t tests and Wilcoxon rank-sum tests, older listeners rated dissonant dyads as more pleasant than did younger listeners ($t_{(21.41)} = 3.47$, $p = 0.002$). Other pairwise comparisons between category were not significant with the corrected α .

The interaction between age and dyad and between age and dyad category was reflected in consonance preference scores (the difference in mean z-scored pleasantness ratings of the five most consonant and the five most dissonant dyads). The younger

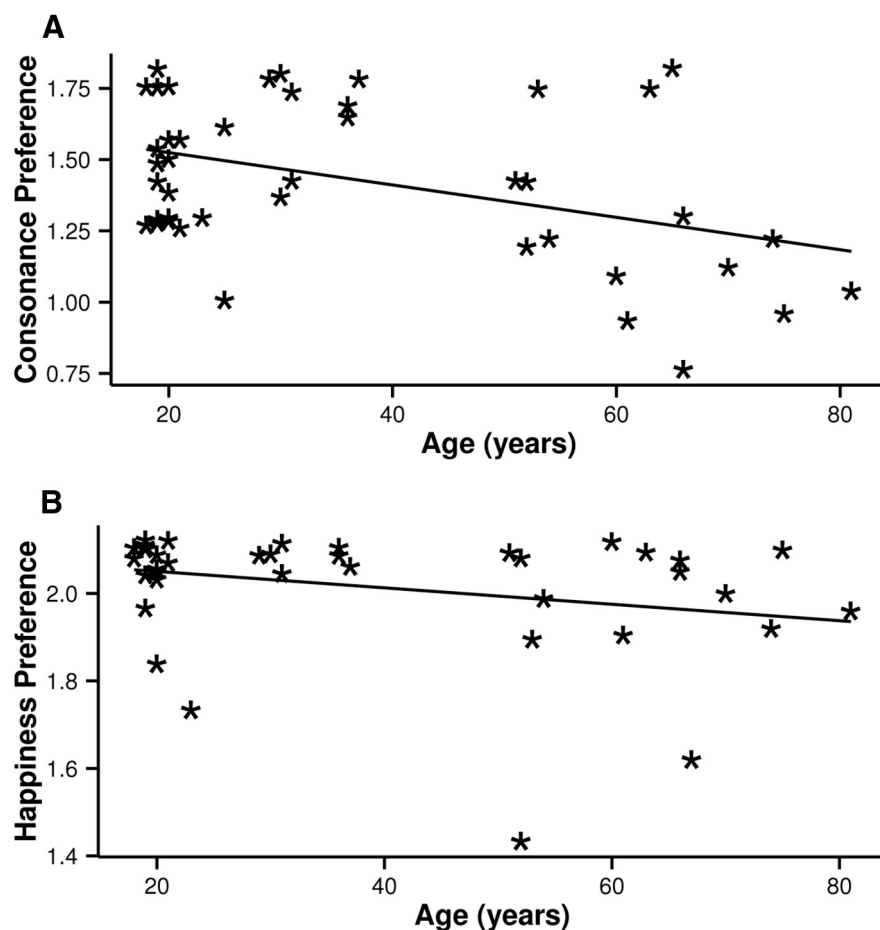


Figure 4. Age in years predicted a decline in consonance preference (**A**) and happiness preference (**B**).

group had greater consonance preference scores (mean = 1.51) than the older group (mean = 1.20; $t_{(20.1)} = 2.86$, $p = 0.007$). To check for reliability, consonance preference scores of the two groups calculated from the two test runs were compared. Scores for both groups significantly correlated across runs (younger, $r_{s(26)} = 0.59$, $p < 0.001$; older, $r_{s(14)} = 0.39$, $p = 0.001$).

Affective voice data were missing from four participants, two from the older group and two from the younger group. Affective voice category ($\varepsilon = 0.62$, $F_{(1.9,70.3)} = 198.73$, $p < 0.001$, $\eta^2 = 0.666$; Fig. 3B) and age ($F_{(1,38)} = 4.98$, $p = 0.03$, $\eta^2 = 0.075$) had a significant effect on ratings of affective voice, but there was no interaction ($\varepsilon = 0.62$, $F_{(1.9,70.3)} = 0.98$, $p = 0.37$). The younger group also had greater happiness preference scores (the difference in z-scored pleasantness ratings of the “happiness” affective voice and the mean z-scored pleasantness rating of the “fear” and “sadness” affective voices; mean = 2.05) than the older group (mean = 1.95; $W = 261$, $p = 0.04$, $r = -0.32$).

To test the effect of age on both sets of preference ratings further, an ANOVA of preference scores was performed with factors age and type (consonance and happiness). Age ($F_{(1,38)} = 17.7$, $p < 0.001$, $\eta^2 = 0.220$) and type ($F_{(1,38)} = 197.91$, $p < 0.001$, $\eta^2 = 0.673$) were significant effects. The two effects interacted ($F_{(1,38)} = 9.21$, $p = 0.004$, $\eta^2 = 0.087$) in that the effect of age was greater for consonance preference scores than for happiness preference scores.

Consonance preference as a function of age is displayed in Figure 4A. Age was significantly negatively correlated with consonance preference ($r_{s(42)} = -0.36$, $p = 0.02$, bootstrapped 95%

CI = -0.66 to -0.07). Although all participants were nonmusicians, the estimate of the total time spent playing a musical instrument in hours was used to control for effects of musical experience (McDermott et al., 2010; Bones et al., 2014). The partial correlation controlling for musical experience was also significant ($r_{s(41)} = -0.44$, $p = 0.003$).

Happiness preference as a function of age is displayed in Figure 4B. Age was also negatively correlated with happiness preference ($r_{s(38)} = -0.34$, $p = 0.035$). To test whether the relation between age and consonance preference was driven by a decline in general affect, consonance preference was partially correlated with age, controlling for happiness preference. This correlation was also significant ($r_{s(37)} = -0.49$, $p = 0.002$), indicating that the effect of age on consonance preference could not be accounted for by the effect of age on affect.

Effect of age on neural harmonic salience

Figure 5 displays the average FFR waveforms (Fig. 5A,B) and spectra (Fig. 5C,D) to the tritone and the perfect fifth. Note that the regularity of the perfect fifth stimulus waveform (Fig. 1D) is reflected in the FFR. The regular periodicity of the waveform is also reflected in the regularly spaced harmonics in the spectrum that form a harmonic series (Fig. 5D). As reported previously (Lee et al., 2009; Bones

et al., 2014), the spectrum of the FFR to the perfect fifth contains frequency components that are not present in the stimulus, but serve to enhance the harmonicity of the response, including the implied 110 Hz F0 of the harmonic series. These distortion products are a consequence of the regularly spaced frequency components of the stimulus and the nonlinearities that exist at all stages of the auditory pathway (Lins et al., 1995; Krishnan, 1999; Lee et al., 2009; Gockel et al., 2012; Smalt et al., 2012; Bones et al., 2014).

Note that 110 Hz is one octave below the F0 of the perfect fifth root note (220 Hz) and corresponds to the second harmonic of a 55 Hz harmonic series. As in a previous study (Bones et al., 2014), the note one octave below the F0 of the root note was found to be marginally less salient than 55 Hz due to the greater number of bins in sieves with low-frequency F0s. Note, however, that the 55 Hz harmonic series is a close match to the spectrum of the perfect fifth because the frequencies coincide with the 110 Hz harmonic series.

For both dyads, the FFR of the younger group is greater in amplitude than that of the older group. The frequency content of the FFR of each dyad is the same for both groups, but the peaks in the spectra in the younger group are greater in amplitude, including the peaks corresponding to distortion products in the case of the perfect fifth.

The neural harmonic salience of the FFR is plotted in Figure 6A. Musical dyad ($F_{(1,41)} = 64.24$, $p < 0.001$, $\eta^2 = 0.282$) was a significant effect, with the neural harmonic salience of the FFR to the perfect fifth being greater than that to the tritone. Age was also a significant effect ($F_{(1,41)} = 10.73$, $p = 0.002$, $\eta^2 = 0.164$), and

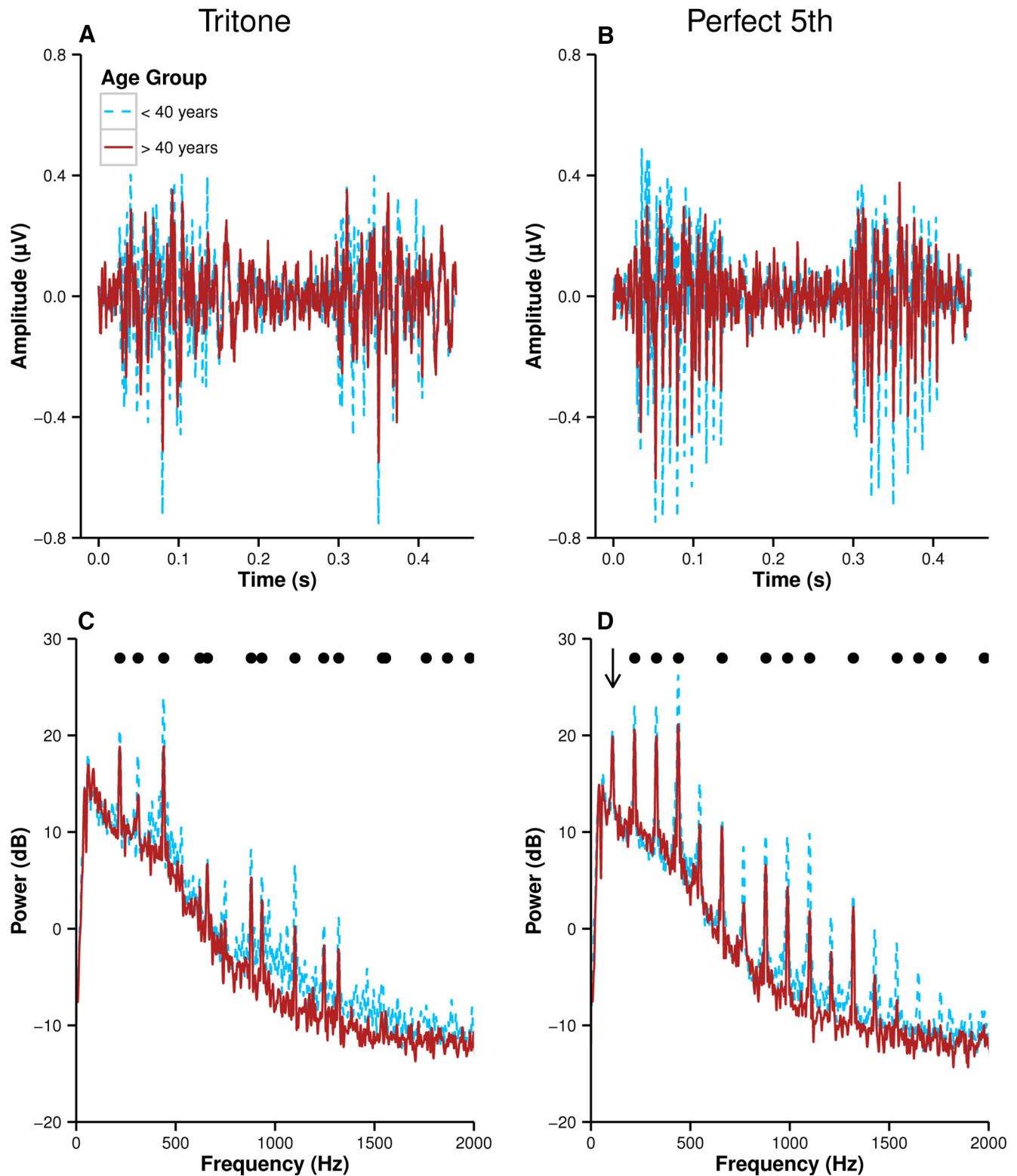


Figure 5. Averaged waveforms (**A**, **B**) and spectra (**C**, **D**) of the FFR to the tritone (left column) and the perfect fifth (right column). The arrow indicates the implied F0 of the harmonic series of the FFR to the perfect fifth (**D**). Power spectra are scaled in dB referenced to 10^{-16}V^2 . Black dots indicate frequencies of the stimulus.

the two effects interacted ($F_{(1,41)} = 11.81$, $p = 0.001$, $\eta^2 = 0.067$). This can be understood by considering that the lower amplitude of the FFR in the older group corresponded to a relatively greater reduction in neural harmonic salience of the FFR to the perfect fifth than the tritone due to the more harmonic structure of the spectrum. Note however that unlike the behavioral ratings, the

effect of age and dyad on neural harmonic salience was not a crossover interaction. Were there a simple linear relation between neural harmonic salience and behavioral ratings, it might be predicted from the FFR data that the older group would rate both the tritone and the perfect fifth as less pleasant than the younger group.

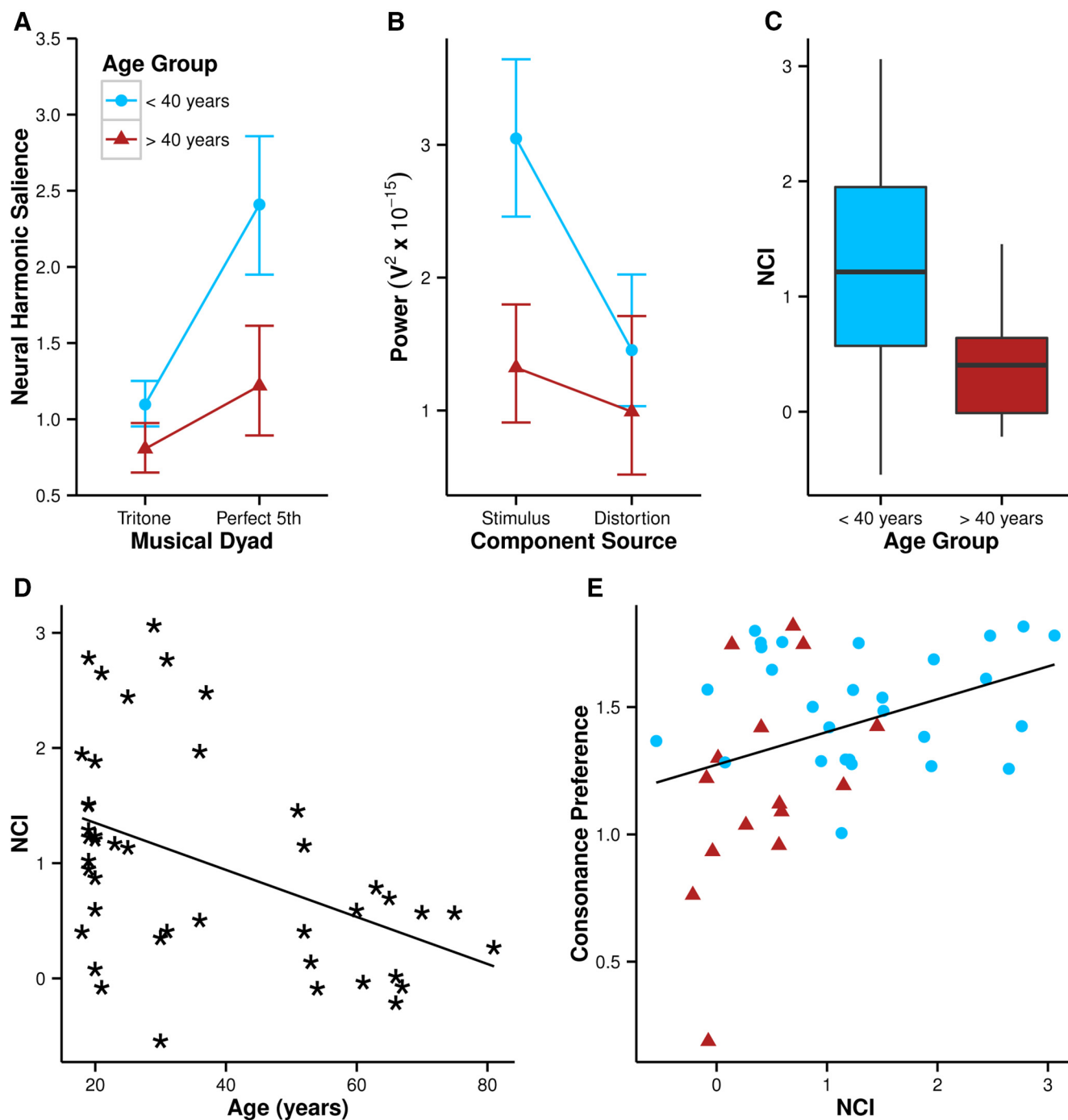


Figure 6. Neural harmonic saliency of the FFR to each dyad (**A**) for both age groups. Error bars indicate 95% confidence intervals. Shown is the mean power of peaks in the FFR to the perfect fifth for each age group that correspond to frequency components in the stimulus and to distortion products (**B**). The younger age group had greater NCI scores (**C**) and age in years predicted a decline in NCI (**D**). NCI predicted greater consonance preference over the whole sample (**E**).

Previous work suggests that the auditory system's nonlinear processing of consonant dyads produces distortion products which serve to enhance both the harmonicity of the FFR and the perceived consonance of musical dyads (Bones et al., 2014). It is possible that reduced cochlear nonlinearity in the older group in the present study could have led to a reduction in the strength of distortion products, leading to a reduction in consonance preference. To test this, an ANOVA of the mean power of the harmonic frequency components in the FFR for the perfect fifth was performed with factors of age group and component source (stimulus or distortion; Fig. 6B). Frequency components was a significant main effect ($F_{(1,41)} = 23.19$,

$p < 0.001$, $\eta^2 = 0.15$), stimulus frequencies having greater magnitude than distortion products. As already shown, age group was a significant main effect ($F_{(1,41)} = 8.46$, $p = 0.008$, $\eta^2 = 0.12$), with the harmonic components having greater amplitude in the younger group. Importantly, the effect of age group and the effect of frequency component interacted ($F_{(1,41)} = 6.31$, $p = 0.016$, $\eta^2 = 0.04$), with the effect of age group being greater on the neural representation of stimulus components compared with distortion products. This analysis suggests that the effect of age on consonance preference was not due primarily to a reduction in the strength of distortion products in the older listeners.

The difference in relative harmonic salience of the two dyads between groups meant that the younger group had significantly greater NCI scores (mean = 1.31) than the older group (mean = 0.41; $t_{(41.0)} = 4.13$, $p < 0.001$; Fig. 6C). NCI as a function of age is plotted in Figure 6D. Age was significantly negatively correlated with NCI ($r_{s(41)} = -0.46$, $p = 0.002$). Partial correlation controlling for musical experience was also significant ($r_{s(40)} = -0.45$, $p = 0.003$).

Figure 6D displays consonance preference as a function of NCI for the whole sample. Crucially, NCI—a measure of neural harmonic salience of consonant relative to dissonant dyads—predicted consonance preference across the age range ($r_{s(41)} = 0.32$, $p = 0.036$, bootstrapped 95% CIs = 0.04–0.62). This correlation remained significant when controlling for musical experience ($r_{s(40)} = 0.34$, $p = 0.029$).

Discussion

Age is associated with a decline in the distinction between categories of musical dyad

In the present study, 44 participants rated musical dyads and affective voices for their pleasantness. The purpose of the study was to test the hypothesis that age would be associated with a decline in the perceptual distinction between consonant and dissonant musical dyads due to a decline in temporal coding. An affective voice-rating task was used to control for differences between the two age groups in use of the scale, aptitude for the task, and general affect. Both age groups made similar use of the scale and did not demonstrate differences in aptitude for the task.

Comparison between the younger and the older group revealed a difference in the perceived pleasantness of musical dyads. The effect of age group interacted with the effect that different combinations of notes had on the perceived pleasantness of musical dyads; in general, the older group distinguished between categories of dyad less than did the younger group and rated dissonant combinations as more pleasant. This difference in relative rating of different categories of musical dyad can be expressed as a difference in consonance preference. On average, the younger group had greater consonance preference than the older group and age in years predicted a decline in consonance preference. It is also noticeable, however, that the range of consonance preference scores is greater in the older group ($SD = 0.39$) than in the younger group ($SD = 0.22$), with a number of older listeners having consonance preference scores comparable to the younger group. This will be discussed further below in relation to NCI scores.

The older age group also rated the positive affective voice (“happiness”) as less pleasant than did the younger group relative to the other categories of voice. However, the effect of age group on preference interacted with stimulus type; the effect of age group was greater for consonance preference than happiness preference. Moreover, the correlation between age and consonance preference remained significant when the effect of happiness preference was controlled—indeed, the effect size increased. The results therefore suggest that different mechanisms were responsible for the perceived pleasantness of musical dyads and affective voices. In particular, the results provide evidence that the perception of consonance is dependent upon integrity of neural temporal coding in a way that the cognitive labeling of affective voices is not. However, another interpretation is that the greater exposure to (and importance of) affective voices during the lifetime may make the neural representations of these stimuli more robust to the effects of aging. Whatever the reason for this discrepancy, the affective voices ratings suggest that the conso-

nance preference results are not due to a general flattening of affect with age.

The distinction between perfect consonance, imperfect consonance, and dissonance is central to Western music (Rameau, 1971) and the hierarchical arrangement of notes by their ratio to a “tonic” determines musical “key” and listeners’ expectations to hear a particular note in a melodic context (Krumhansl and Kessler, 1982). The resolution or nonresolution of tension evoked by sense of key contributes to the emotional response evoked by music. Engagement with music in old age provides significant benefit to cognitive, social, and emotional well-being (for review, see Creech et al., 2013) and different aspects of music use are associated with different dimensions of well-being (Laukka, 2007). Further work is needed to determine whether a reduction in consonance preference with age is associated with a decline in listening to and engaging with music and what the repercussions of this are for quality of life.

A further question left unanswered by the present study is whether the age-related decline in consonance preference in non-musicians seen here also occurs for musicians. Previous work has found a correlation between musical experience and preference for consonance in young, normal hearing participants (McDermott et al., 2010; Bones et al., 2014) and between musical experience and NCI scores (Bones et al., 2014). Previous work also suggests that older musicians perform better in behavioral tests of temporal processing than older nonmusicians (Parbery-Clark et al., 2011) and that auditory training may even reverse age-related deficits in temporal coding (Anderson et al., 2013). A possible direction for future work therefore is to test the hypothesis that the auditory training involved in musicianship protects against an age-related decline in consonance perception.

Age is associated with weaker representation of harmonic structure by neural temporal coding

Tufts et al. (2005) reported that four listeners with mild to moderate hearing loss (mean age = 69) rated perfect fourth and perfect fifth dyads as being less consonant relative to other dyads compared with four normal hearing listeners (mean age = 50). The loss of outer hair cell function responsible for elevated auditory thresholds also results in a decline in the tuning of the cochlear response (Ruggero and Rich, 1991; Ruggero et al., 1997) necessary for resolving individual harmonics; Tufts et al. (2005) therefore interpreted the results of the study as being a consequence of impaired cochlear frequency selectivity. Bidelman and Heinz (2011) modeled the response of the auditory nerve (AN; Zilany et al., 2009) to musical dyads in normal hearing and mild to moderately hearing impaired listeners using the audiometric data reported by Tufts et al. (2005). Bidelman and Heinz (2011) used a pooled autocorrelation function calculated from post-stimulus time histograms of AN spikes to determine the salience of temporal information relevant to each dyad. The contrast between consonant and dissonant dyads in the output of the normal hearing model was greater than in that of the hearing impaired model, suggesting that the consequences of hearing impairment for the perception of musical consonance can be accounted for by temporal coding at the level of the AN.

Similar behavioral results to those reported by Tufts et al. (2005) were found in the present study, in which hearing thresholds were controlled: all participants had hearing thresholds ≤ 20 dB HL at 250, 500, 1000, and 2000 Hz and therefore had clinically normal hearing at the frequency range of the dyads used, but not above this frequency. At high stimulus levels, the FFR to musical dyads is likely to be generated by regions of the cochlea tuned to

frequencies above the stimulus (Gardi and Merzenich, 1979; Dau, 2003; O. Bones and C.J. Plack, unpublished observations). However, because the response of the cochlea at places tuned above the frequency of the stimulus is linear and unaffected by outer hair cell gain (Ruggero et al., 1997), it is unlikely that the outer hair cell loss likely responsible for elevated hearing thresholds at these frequencies (Moore, 2007) in older participants would significantly affect generation of the FFR. Nonetheless, since hearing loss at high frequencies is highly correlated with age, partial correlation of consonance preference and age controlling for differences in hearing threshold at high frequencies was not meaningful and, therefore, an effect of hearing loss cannot be fully ruled out.

We argue, however, that the results of the present study are evidence for an age-related decline in neural temporal coding. There is a growing body of evidence demonstrating age-related deficits in temporal processing independently of hearing loss (Strouse et al., 1998; Clinard et al., 2010; Hopkins and Moore, 2011; Anderson et al., 2012; Neher et al., 2012; Ruggles et al., 2012; Marmel et al., 2013). Clinard et al. (2010) found that age predicted pitch discrimination and magnitude of the FFR at the stimulus frequency, but did not find a significant correlation between the two measures. Marmel et al. (2013) also found that age predicted both pitch discrimination and the representation of the stimulus in the FFR and that the FFR predicted behavioral performance. Importantly, the two measures remained significantly correlated when controlling for the effect of hearing loss.

It has been demonstrated previously that the representation of the harmonic structure of musical dyads in the neural temporal coding represented by the FFR is predictive of their pleasantness and that individual differences in the distinction between categories of dyad correspond to individual differences in the representation of the harmonic structure of dyads by neural coding (Bones et al., 2014). The present study demonstrates that age is associated with a decline in the representation of the harmonic structure of dyads in neural temporal coding and that this may correspond to a decline in the perceptual distinction of categories of musical dyads. It is also notable that within the older group there is a trend for greater NCI scores being associated with greater consonance preference (Fig. 6D). That older listeners with greater NCI scores had consonance preference scores comparable to listeners in the younger group and that (consistent with previous work; Bones et al., 2014) NCI is associated with consonance preference across the entire sample suggests that it is a decline in the neural representation of harmonicity rather than age per se that drives a decline in consonance perception.

References

- Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N (2012) Aging affects neural precision of speech encoding. *J Neurosci* 32:14156–14164. [CrossRef Medline](#)
- Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N (2013) Reversal of age-related neural timing delays with training. *Proc Natl Acad Sci U S A* 110:4357–4362. [CrossRef Medline](#)
- Belin P, Fillion-Bilodeau S, Gosselin F (2008) The montreal affective voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behav Res Methods* 40:531–539. [CrossRef Medline](#)
- Bidelman GM, Heinz MG (2011) Auditory-nerve responses predict pitch attributes related to musical consonance-dissonance for normal and impaired hearing. *Journal of the Acoustical Society of America* 130:1488–1502. [CrossRef Medline](#)
- Bidelman GM, Krishnan A (2009) Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *J Neurosci* 29:13165–13171. [CrossRef Medline](#)
- Bones O, Hopkins K, Krishnan A, Plack CJ (2014) Phase locked neural activity in the human brainstem predicts preference for musical consonance. *Neuropsychologia* 58C:23–32.
- Brugge JF, Anderson DJ, Hind JE, Rose JE (1969) Time structure of discharges in single auditory nerve fibers of the squirrel monkey in response to complex periodic sounds. *J Neurophysiol* 32:386–401. [Medline](#)
- Caspary DM, Palombi PS, Hughes LF (2002) Gabaergic inputs shape responses to amplitude modulated stimuli in the inferior colliculus. *Hear Res* 168:163–173. [CrossRef Medline](#)
- Caspary DM, Schatteman TA, Hughes LF (2005) Age-related changes in the inhibitory response properties of dorsal cochlear nucleus output neurons: Role of inhibitory inputs. *J Neurosci* 25:10952–10959. [CrossRef Medline](#)
- Clinard CG, Tremblay KL, Krishnan AR (2010) Aging alters the perception and physiological representation of frequency: evidence from human frequency-following response recordings. *Hear Res* 264:48–55. [CrossRef Medline](#)
- Cousineau M, McDermott JH, Peretz I (2012) The basis of musical consonance as revealed by congenital amusia. *Proc Natl Acad Sci U S A* 109:19858–19863. [CrossRef Medline](#)
- Creech A, Hallam S, McQueen H, Varvarigou M (2013) The power of music in the lives of older adults. *Research Studies in Music Education* 35:87–102. [CrossRef](#)
- Dau T (2003) The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. *Journal of the Acoustical Society of America* 113:936–950. [CrossRef Medline](#)
- Frisina RD, Walton JP (2006) Age-related structural and functional changes in the cochlear nucleus. *Hear Res* 216–217:216–223. [Medline](#)
- Furman AC, Kujawa SG, Liberman MC (2013) Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *J Neurophysiol* 110:577–586. [CrossRef Medline](#)
- Gardi J, Merzenich M (1979) Effect of high-pass noise on the scalp-recorded frequency following response (ffr) in humans and cats. *Journal of the Acoustical Society of America* 65:1491–1500. [CrossRef](#)
- Goblick TJ Jr, Pfeiffer RR (1969) Time-domain measurements of cochlear nonlinearities using combination click stimuli. *Journal of the Acoustical Society of America* 46:924–938. [CrossRef Medline](#)
- Gockel HE, Carlyon RP, Mehta A, Plack CJ (2011) The frequency following response (ffr) may reflect pitch-bearing information but is not a direct representation of pitch. *Journal of the Association for Research in Otolaryngology* 12:767–782. [CrossRef Medline](#)
- Gockel HE, Farooq R, Muhammed L, Plack CJ, Carlyon RP (2012) Differences between psychoacoustic and frequency following response measures of distortion tone level and masking. *Journal of the Acoustical Society of America* 132:2524–2535. [CrossRef Medline](#)
- Hopkins K, Moore BC (2011) The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *Journal of the Acoustical Society of America* 130:334–349. [CrossRef Medline](#)
- Klug A, Bauer EE, Hanson JT, Hurley L, Meitzen J, Pollak GD (2002) Response selectivity for species-specific calls in the inferior colliculus of Mexican free-tailed bats is generated by inhibition. *J Neurophysiol* 88:1941–1954. [Medline](#)
- Koch U, Grothe B (1998) Gabaergic and glycinergic inhibition sharpens tuning for frequency modulations in the inferior colliculus of the big brown bat. *J Neurophysiol* 80:71–82. [Medline](#)
- Krishnan A (1999) Human frequency-following responses to two-tone approximations of steady-state vowels. *Audiology and Neurotology* 4:95–103. [CrossRef Medline](#)
- Krishnan A (2007) Frequency-following response. In: *Auditory evoked potentials: basic principles and clinical application* (Burkard RF, Don M, Eggermont JJ, eds), pp 313–333. New York: Lippincott Williams and Wilkins.
- Krishnan A, Plack CJ (2011) Neural encoding in the human brainstem relevant to the pitch of complex tones. *Hear Res* 275:110–119. [CrossRef Medline](#)
- Krishnan A, Xu Y, Gandour J, Cariani P (2005) Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res Cogn Brain Res* 25:161–168. [CrossRef Medline](#)
- Krumhansl CL, Kessler EJ (1982) Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review* 89:334–368. [CrossRef Medline](#)
- Kujawa SG, Liberman MC (2009) Adding insult to injury: Cochlear nerve

- degeneration after “temporary” noise-induced hearing loss. *J Neurosci* 29:14077–14085. [CrossRef Medline](#)
- Laukka P (2007) Uses of music and psychological well-being among the elderly. *Journal of Happiness Studies* 8:215–241. [CrossRef](#)
- Lee KM, Skoe E, Kraus N, Ashley R (2009) Selective subcortical enhancement of musical intervals in musicians. *J Neurosci* 29:5832–5840. [CrossRef Medline](#)
- Lins OG, Picton PE, Picton TW, Champagne SC, Durieux-Smith A (1995) Auditory steady-state responses to tones amplitude-modulated at 80–110 Hz. *Journal of the Acoustical Society of America* 97:3051–3063. [CrossRef Medline](#)
- Makary CA, Shin J, Kujawa SG, Liberman MC, Merchant SN (2011) Age-related primary cochlear neuronal degeneration in human temporal bones. *Journal of the Association for Research in Otolaryngology* 12:711–717. [CrossRef Medline](#)
- Marmel F, Linley D, Carlyon RP, Gockel HE, Hopkins K, Plack CJ (2013) Subcortical neural synchrony and absolute thresholds predict frequency discrimination independently. *J Assoc Res Otolaryngol* 14:757–766. [CrossRef Medline](#)
- McDermott JH, Lehr AJ, Oxenham AJ (2010) Individual differences reveal the basis of consonance. *Curr Biol* 20:1035–1041. [CrossRef Medline](#)
- Mickelson WC, Riemann H (1977) Hugo Riemann’s theory of harmony: a study and history of music theory. Lincoln, NE: University of Nebraska.
- Milbrandt JC, Albin RL, Caspary DM (1994) Age-related decrease in gabab receptor binding in the fischer 344 rat inferior colliculus. *Neurobiology of Aging* 15:699–703. [CrossRef Medline](#)
- Moore, BC (2007) Cochlear hearing loss: physiological, psychological, and technical issues, Ed 2. Chichester, UK: Wiley.
- Moushegian G, Rupert AL, Stillman RD (1973) Laboratory note. Scalp-recorded early responses in man to frequencies in the speech range. *Electroencephalogr Clin Neurophysiol* 35:665–667. [CrossRef Medline](#)
- Neher T, Lunner T, Hopkins K, Moore BC (2012) Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners (I). *Journal of the Acoustical Society of America* 131:2561–2564. [CrossRef Medline](#)
- Parbery-Clark A, Strait DL, Anderson S, Hittner E, Kraus N (2011) Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise. *PLoS One* 6:e18082. [CrossRef Medline](#)
- Plomp R, Levelt WJ (1965) Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America* 38:548–560. [CrossRef Medline](#)
- Puel JL, Pujol R, Tribillac F, Ladrech S, Eybalin M (1994) Excitatory amino acid antagonists protect cochlear auditory neurons from excitotoxicity. *J Comp Neurol* 341:241–256. [CrossRef Medline](#)
- Puel JL, Ruel J, Gervais d’Aldin C, Pujol R (1998) Excitotoxicity and repair of cochlear synapses after noise-trauma induced hearing loss. *Neuroreport* 9:2109–2114. [CrossRef Medline](#)
- Rameau J-P (1971) Treatise on harmony, Ed1. New York: Dover Publications.
- Ruggero MA, Rich NC (1991) Furosemide alters organ of corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane. *J Neurosci* 11:1057–1067. [Medline](#)
- Ruggero MA, Rich NC, Recio A, Narayan SS, Robles L (1997) Basilar-membrane responses to tones at the base of the chinchilla cochlea. *Journal of the Acoustical Society of America* 101:2151–2163. [CrossRef Medline](#)
- Ruggles D, Bharadwaj H, Shinn-Cunningham BG (2012) Why middle-aged listeners have trouble hearing in everyday settings. *Curr Biol* 22:1417–1422. [CrossRef Medline](#)
- Sergeyenko Y, Lall K, Liberman MC, Kujawa SG (2013) Age-related cochlear synaptopathy: An early-onset contributor to auditory functional decline. *J Neurosci* 33:13686–13694. [CrossRef Medline](#)
- Smalt CJ, Krishnan A, Bidelman GM, Ananthakrishnan S, Gandour JT (2012) Distortion products and their influence on representation of pitch-relevant information in the human brainstem for unresolved harmonic complex tones. *Hear Res* 292:26–34. [CrossRef Medline](#)
- Strouse A, Ashmead DH, Ohde RN, Grantham DW (1998) Temporal processing in the aging auditory system. *Journal of the Acoustical Society of America* 104:2385–2399. [CrossRef Medline](#)
- Tufts JB, Molis MR, Leek MR (2005) Perception of dissonance by people with normal hearing and sensorineural hearing loss. *Journal of the Acoustical Society of America* 118:955–967. [CrossRef Medline](#)
- Walton JP, Frisina RD, Ison JR, O’Neill WE (1997) Neural correlates of behavioral gap detection in the inferior colliculus of the young cba mouse. *J Comp Physiol A* 181:161–176. [CrossRef Medline](#)
- Walton JP, Frisina RD, O’Neill WE (1998) Age-related alteration in processing of temporal sound features in the auditory midbrain of the cba mouse. *J Neurosci* 18:2764–2776. [Medline](#)
- Walton JP, Simon H, Frisina RD (2002) Age-related alterations in the neural coding of envelope periodicities. *J Neurophysiol* 88:565–578. [Medline](#)
- Wright D, Field A (2009) Methods: giving your data the bootstrap. *Psychologist* 22:412.
- Zilany MS, Bruce IC, Nelson PC, Carney LH (2009) A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics. *Journal of the Acoustical Society of America* 126:2390–2412. [CrossRef Medline](#)