Consonance Preferences Within an Unconventional Tuning System

Ronald S. Friedman, & Douglas A. Kowalewski
University at Albany, State University of New York

Dominique T. Vuvan
Skidmore College

W. Trammell Neill
University at Albany, State University of New York

Recently, Bowling, Purves, and Gill (2018a), found that individuals perceive chords with spectra resembling a harmonic series as more consonant. This is consistent with their vocal similarity hypothesis (VSH), the notion that the experience of consonance is based on an evolved preference for sounds that resemble human vocalizations. To rule out confounding between harmonicity and familiarity, we extended Bowling et al.'s (2018a) procedure to chords from the unconventional Bohlen-Pierce chromatic just (BPCJ) scale. We also assessed whether the association between harmonicity and consonance was moderated by timbre by presenting chords generated from either piano or clarinet samples. Results failed to straightforwardly replicate this association; however, evidence of a positive correlation between harmonicity and consonance did emerge across timbres following post hoc exclusion of chords containing intervals that were particularly similar to conventional equal-tempered dyads. Supplementary regression analyses using a more comprehensive measure of harmonicity confirmed its positive association with consonance ratings of BPCJ chords, yet also showed that spectral interference independently contributed to these ratings. In sum, our results are consistent with the VSH; however, they also suggest that a composite model, incorporating both harmonicity as well as spectral interference as predictors, would best account for variance in consonance judgments.

Received: January 31, 2019, accepted October 18, 2020.

Key words: consonance, harmonicity, chords, Bohlen-Pierce, vocal similarity

Among the many types of chords that appear within various musical styles, some tend to be perceived as more "consonant," that is, more pleasant or stable than others (Bowling & Purves, 2015). The consonance of a chord has been found to be related to its familiarity. As noted by Johnson-Laird, Kang, and Leong (2012), experiments dating back to the early twentieth century showed that irregularly tuned music (Meyer, 1903) as well as conventionally dissonant two-note chords (Valentine, 1914) became less disagreeable with repeated hearings. More recently, McLachlan, Marco, Light, and Wilson (2013) trained nonmusicians to perform a pitch-matching task in which they had to adjust the pitch of a probe tone until it matched that of a target component tone in an immediately preceding chord. Following multiple pitch-matching sessions using a randomly selected set of chords, McLachlan et al. (2013) discovered that participants came to rate the chords that appeared in the pitch-matching task, relative to chords to which they were not exposed, as slightly more consonant. These lab-based exposure effects are in accord with historical evidence that chords once considered highly dissonant in Western classical music became palatable among listeners as composers began to use them more frequently (see Rameau, 1722/1971, as cited in Johnson-Laird et al., 2012, p. 22).

Beyond the influence of familiarity, in the late nineteenth century, Helmholtz (1877/1954) proposed that perceptions of consonance in isolated chords are related to the absence of rapid amplitude fluctuations (i.e., "beating") caused by patterns of constructive and destructive interference between the partials of constituent chord tones: The partials of the most consonant chords either tend to align or to be spaced far enough apart to minimize beating and the unpleasant sensation of "roughness" with which it is associated. Building upon Helmholtz' historic contribution, a number of models have been developed to predict variance in consonance ratings of isolated chords by estimating the overall extent of beating between adjacent partials (e.g., Hutchinson & Knopoff, 1978; Kameoka & Kuriyagawa, 1969; Plomp & Levelt, 1965). However, as recently reviewed by Bowling and Purves (2015), whereas

Music Perception, volume 38, issue 3, pp. 313–330, issn 0730-7829, electronic issn 1533-8312. DOI: https://doi.org/10.1525/mp.2021.38.3.313

auditory roughness may contribute to the perception of consonance, the ability of such spectral interference-based models to comprehensively account for consonance ratings of isolated chords may be challenged on a number of grounds: First, although dichotic presentation of chord tones eliminates beating by preventing interactions between tone components, it does not reliably alter patterns of consonance perception (e.g., Guernsey, 1928). Second, when chords are made up of synthesized tones from which interacting partials have been artificially removed, this fails to consistently heighten their perceived consonance (Nordmark & Fahlén, 1988). Third, even when the addition of tones to a chord promotes beating by increasing the overall number of interacting partials, this does not necessarily affect the perceived consonance of the chord (e.g., Nordmark & Fahlén, 1988).

Finally, a series of studies by McDermott and his colleagues (McDermott, Lehr, & Oxenham, 2010; Cousineau, McDermott, & Peretz, 2012; McDermott, Schultz, Undurraga, & Godoy, 2016) has provided particularly compelling evidence for the independence of auditory roughness and perceived consonance. In the first of these studies, McDermott et al. (2010) created a novel measure of individual differences in aversion to beating by presenting participants with pairs of pure tones (i.e., tones of a single frequency) featuring closely spaced fundamental frequencies (F0s). They then compared participants' preferences for these tone pairs when they were presented diotically (i.e., each ear is presented with both tones), which would enable beating, versus dichotically (i.e., each ear is presented with only one tone from the pair), which would impede beating. Here, a stronger aversion to beating was operationalized as diminished preference for the diotically presented pure-tone pairs. McDermott et al. (2010) also gauged individual differences in consonance preference, operationalized as participants' relative preference for conventionally consonant versus dissonant chords. Individual differences in aversion to beating were not reliably associated with measures of consonance preference. This suggests that the experience of consonance (i.e., the extent to which a chord sounds pleasant versus unpleasant) is not primarily based on the amount of spectral interference between the partials making up the chord tones.

In the second study, Cousineau et al. (2012) used another type of individual differences-based method for investigating the relationship between beating and consonance perception. Here, they recruited participants with congenital amusia, a neurological disorder associated with deficits in processing pitch (Ayotte, Peretz, & Hyde, 2002; Vuvan, Nunes-Silva, & Peretz, 2015). Using the same auditory stimuli that were presented to participants in the earlier study by McDermott et al. (2010), Cousineau et al. (2012) found that although amusics were significantly less averse to conventionally dissonant chords than individuals in a control group, they were equally averse to beating, again suggesting that the latter is not at the root of consonance judgments.

Finally, in the third study, McDermott et al. (2016) employed a cross-cultural approach by recruiting participants from an indigenous group in Bolivia, the Tsimane', members of which have had very limited exposure to Western music. Much like the amusics in their earlier study, McDermott and his colleagues found that the Tsimane' did not reliably show a preference for conventionally consonant versus dissonant chords, although they did distinguish between them perceptually. However, they shared Western listeners' aversion to beating, providing converging evidence that the experience of consonance is not reliant upon auditory roughness due to interference between spectral components.

Interestingly, whereas McDermott and his colleagues showed that preference for conventionally consonant chords is unrelated to evaluations of beating, they found that it is strongly associated with a predilection for sounds resembling a harmonic series, in which partials are related by integer multiples of the F0. Specifically, McDermott et al. (2010) compared consonance ratings of complex tones consisting of harmonic partials with those of complex tones, the partials of which had been perturbed in frequency such that they did not conform with a harmonic series. Preference for harmonic over inharmonic complex tones was reliably correlated with preference for consonance. Relatedly, Cousineau et al. (2012) found that although congenital amusics do retain an aversion to beating, they are unable to discriminate between harmonic and inharmonic complex tones, suggesting that this deficit in periodicity perception underlies their failure to show preference between conventionally consonant and dissonant chords.

Consistent with this work, Bowling, Purves, and Gill (2018a), recently conducted a study in which they developed a parsimonious means of estimating the overall *harmonicity* of a given chord, the extent to which the partials present in its component tones collectively produce a harmonic complex sound wave. They then had participants rate the consonance of every possible two-note (dyad), three-note (triad), and four-note chord (tetrad) that could be constructed within the range of an octave using a Western scale. Across all chord types, participants tended to find relatively harmonic chords more consonant, providing converging evidence for the notion that harmonicity fundamentally contributes to the perception of consonance (see also, Parncutt, 1989; Terhardt, 1984).

The results of these behavioral studies are consistent with the findings of physiological studies examining neural responses to chords varying in harmonicity. For instance, Tramo, Cariani, Delgutte, and Braida (2001) found that patterns of action potentials in the auditory nerve (AN) are distinct for musical dyads that are high versus low in harmonicity: For relatively harmonic dyads, including the perfect 4th or 5th, the pattern of neural spikes clearly reflected the periodicity of the constituent tones in the dyad (e.g., the tonic and dominant scale degrees for the perfect 5th) as well as that of the missing F0 (i.e., the subharmonic of which the constituent pitches are themselves integer multiples). In contrast, the pattern of spikes for relatively inharmonic dyads, including the minor 2nd and tritone, demonstrated almost no such representation of constituent or harmonically related tones. Tramo et al. (2001) found a high correlation between the strength of the representation of the missing F0 in the distribution of AN spikes and an independent measure of the self-reported consonance of the dyads, suggesting a preference for dyads that elicit a more harmonic pattern of neural activity (see also, Cariani, 2019).

Similar findings have also been reported with respect to neural activation at higher levels of the auditory system. For instance, Bidelman and Krishnan (2009) recorded the frequency-following response (FFR)—an evoked potential that tracks the cycles of a periodic sound stimulus—as it was elicited by each of a set of conventional musical dyads varying in consonance/dissonance. Again, relatively harmonic (e.g., perfect 5th), relative to inharmonic (e.g., major 7th) dyads elicited more harmonic neural response patterns. Moreover, the harmonicity of the FFR was robustly correlated with listeners' subjective consonance judgments. Mirroring these results, Itoh, Suwazono, and Nakada (2010) subsequently reported that the magnitude of an auditory cortical evoked potential (the N2) to a dyadic interval was associated with the harmonicity of the dyad, which in turn predicted its self-reported consonance. Collectively, these and related neurophysiological findings demonstrate that the "auditory system exploits the harmonicity of sound to code the perceptual pleasantness of music" (Bidelman, 2013, p. 4).

At a functional level of analysis, why might individuals experience harmonic combinations of tones as more consonant? According to Bowling et al. (2018a), harmonicity may be associated with the consonance of a chord because it is a distinguishing characteristic of human vocalizations. Specifically, they posit that given the vital communicative role of harmonic vocal sounds (verbal and non-verbal), our perceptual systems have evolved to reinforce attention to harmonic environmental sounds more generally. Restated, their vocal similarity hypothesis (VSH) proposes that the biological foundation for consonance lies in a canalized attraction to the harmonic stimuli that play a critical role in audio-vocal communication. Although it does not preclude the role of experience in shaping the preference for harmonic chords, the VSH implies that this preference does not require prior exposure to such chords. Indeed, Bowling, Purves, and Gill (2018b) have argued that the correlation between harmonicity and consonance judgments is "unlikely to depend on a particular tuning system" (p. E4958). This suggests that even if individuals were exposed to chords that do not exist within the musical systems to which they have been enculturated, they should be inclined to evaluate unconventional chords more favorably when these are more harmonic.

However, in all of the aforementioned studies, including that by Bowling et al. (2018a), the chord stimuli employed were generated based either on the equal tempered intervals to which most Western listeners are accustomed or, as in Bowling et al.'s (2018a) case, just tempered intervals that are quite similar to their equal tempered counterparts. As discussed earlier, it has been empirically established that tone combinations tend to be perceived as more consonant as they gain familiarity. This raises the possibility that prior findings demonstrating a link between harmonic similarity and consonance—including those of Bowling et al. (2018a)—may have at least partially resulted from confounding between the harmonicity of the chords presented and participants' familiarity with these chords. To address whether harmonicity predicts consonance ratings even within a highly unconventional tuning system, and to help rule out the possibility that this predictive relationship is an artifact of familiarity, we replicated Bowling et al.'s (2018a) study using chords generated from the Bohlen-Pierce chromatic just scale (hereafter, BPCJ scale; Mathews, Pierce, Reeves, & Roberts, 1988; see also, Loy, 2006). Unlike the Western just or equal-tempered chromatic scales, which divide an octave into 12 intervals, the BPCJ scale divides a *tritave* (representing a 3:1 ratio between tones or the span of an octave plus a major fifth) into 13 intervals based on a series of odd integer frequency ratios (see Table 1). The selection of the BPCJ in the present study was motivated by a number of factors: First, none of the non-unison

TABLE 1. *Bohlen-Pierce Dyads: Defining Features and Consonance Ratings*

Dyad	Ratio	CD	DV	Scale Degree		F0 (Hz)		Consonance Ratings		
				Tone 1	Tone 2	Tone 1	Tone 2	Overall	Piano	Clarinet
1	27:25	0.0205	1.08	0	1	251.57	271.69	2.58	2.25	2.84
2	25:21	0.0013	1.19	0	2	238.88	284.38	4.53	4.68	4.41
3	9:7	0.0258	1.29	0	3	228.93	294.33	3.61	3.73	3.52
4	7:5	0.0142	1.40	0	4	218.03	305.24	3.60	3.28	3.87
5	75:49	0.0323	1.53	0	5	206.77	316.49	3.62	3.77	3.50
6	5:3	0.0151	1.67	0	6	196.22	327.04	4.61	4.44	4.75
7	9:5	0.0182	1.80	0	7	186.88	336.38	3.61	3.59	3.63
8	49:25	0.0400	1.96	0	8	176.78	346.48	3.42	3.02	3.74
9	15:7	0.0239	2.14	0	9	166.49	356.77	3.18	3.19	3.18
10	7:3	0.0451	2.33	0	10	156.98	366.28	3.67	3.75	3.61
11	63:25	0.0002	2.52	0	11	148.65	374.61	4.89	5.40	4.46
12	25:9	0.0506	2.78	0	12	138.51	384.75	3.05	3.25	2.87
13	3:1	0.0034	3.00	0	13	130.82	392.45	4.13	4.37	3.93

Note: Scale degrees represent number of BPCJ semitone intervals above the lowest tone (labeled "0"). Overall consonance ratings represent grand means of the ratings tendered by participants for each dyad, collapsed across trials. Ratings are also reported separately for participants in the piano ($n = 51$) and clarinet timbre ($n = 59$) conditions. CD = chromatic distance score. DV = Decimal value of frequency ratio.

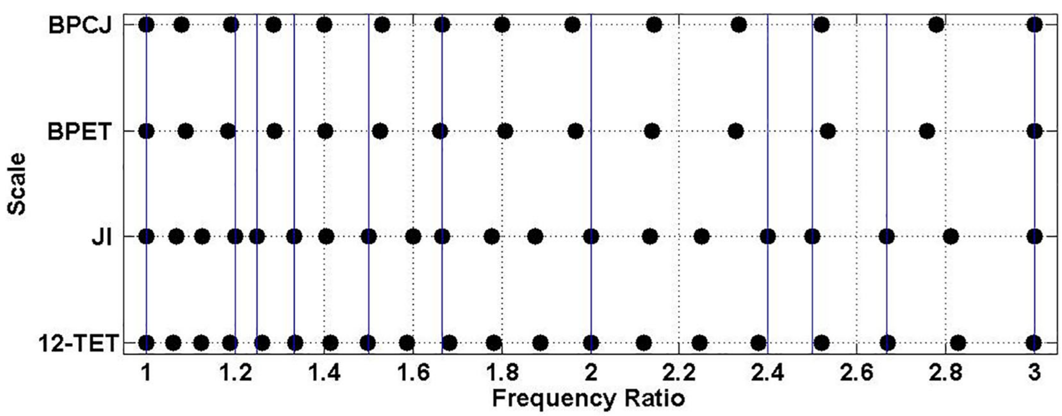

FIGURE 1. Comparison of frequency interval ratios in the BPCJ, Bohlen-Pierce equal tempered (BPET), Western just (JI), and Western 12-TET scales. Circles denote scale tones. Solid vertical lines indicate locations of prominent simple interval ratios (1/1, 6/5, 5/4, 4/3, 3/2, 5/3, 2, 12/5, 5/2, 8/3, and 3/1).

intervals created by chords constructed based on the BPCJ appear in the twelve-tone equal-tempered (12-TET) chromatic scale used in the vast majority of Western music and therefore these exact intervals are virtually never experienced by Western-enculturated listeners. As such, the scale is well-suited for testing the association between harmonicity and consonance in unfamiliar chords. (See Figure 1 for a comparison of the intervals in the BPCJ, Bohlen-Pierce equal tempered, Western just, and 12-TET scales). Second, the BPCJ contains multiple intervals based on small integer ratios (e.g., 7/3, 7/5, 3/1; see Table 1) enabling substantial variance in the harmonicity of chords built from the scale and ensuring that many of these chords are comparable in harmonicity to those in Western scales. Third, earlier work by Matthews et al. (1988; see Table II, p. 1217) established that there is substantial variance in consonance ratings of BPCJ-based chords, with several triads rated as very dissonant (below a "2" on a 7-point scale) and many others rated as highly consonant (exceeding a "5" on the same scale). Sufficient variance in such ratings is necessary to avoid floor effects in any assessment of correlations between harmonicity and consonance. Finally, although there is an equal-tempered version of the Bohlen-Pierce chromatic scale based on division of the tritave into 13 logarithmically

TABLE 2. ***Bohlen-Pierce Triads: Defining Features and Consonance Ratings***

	Scale Degree			F0 (Hz)			Average Consonance Ratings		
Triad	Tone 1	Tone 2	Tone 3	Tone 1	Tone 2	Tone 3	Overall	Piano	Clarinet
1	0	1	2	239.99	259.19	285.71	3.06	2.87	3.22
2	0	1	3	233.20	251.86	299.83	3.10	2.95	3.21
3	0	1	4	225.54	243.59	315.76	3.10	2.76	3.37
4	0	1	5	217.38	234.77	332.73	3.26	3.08	3.41
5	0	1	6	209.49	226.25	349.15	3.10	2.85	3.29
6	0	1	7	202.29	218.47	364.12	3.04	2.90	3.15
7	0	1	8	194.28	209.82	380.79	2.94	2.91	2.96
8	0	1	9	185.87	200.74	398.29	3.14	2.97	3.27
9	0	1	10	177.85	192.07	414.97	2.87	2.68	3.03
10	0	1	11	170.63	184.28	429.98	3.19	3.13	3.23
11	0	1	12	161.57	174.50	448.82	2.77	2.89	2.67
12	0	1	13	154.51	166.87	463.52	3.25	3.41	3.13
13	0	2	3	225.79	268.80	290.30	3.29	3.19	3.38
14	0	2	4	218.60	260.24	306.04	4.12	3.92	4.28
15	0	2	5	210.93	251.11	322.85	3.97	4.17	3.81
16	0	2	6	203.49	242.25	339.15	4.06	4.29	3.87
17	0	2	7	196.69	234.16	354.04	4.29	4.40	4.20
18	0	2	8	189.11	225.13	370.65	3.81	4.14	3.54
19	0	2	9	181.13	215.63	388.13	3.44	3.17	3.67
20	0	2	10	173.50	206.55	404.84	3.40	3.29	3.50
21	0	2	11	166.63	198.36	419.90	3.56	3.43	3.66
22	0	2	12	157.98	188.07	438.84	3.40	3.61	3.22
23	0	2	13	151.22	180.02	453.65	4.48	4.67	4.32
24	0	3	4	212.95	273.80	298.14	3.18	3.13	3.22
25	0	3	5	205.67	264.43	314.80	4.23	4.49	4.02
26	0	3	6	198.59	255.33	330.98	3.85	3.45	4.17
27	0	3	7	192.11	246.99	345.79	3.65	3.41	3.85
28	0	3	8	184.87	237.69	362.34	3.79	3.79	3.79
29	0	3	9	177.23	227.87	379.79	3.36	3.36	3.36
30	0	3	10	169.92	218.47	396.49	3.57	3.58	3.57
31	0	3	11	163.32	209.99	411.58	3.72	3.58	3.84
32	0	3	12	155.01	199.30	430.58	3.31	3.36	3.27
33	0	3	13	148.49	190.92	445.48	3.86	4.11	3.66
34	0	4	5	199.69	279.56	305.64	3.37	3.31	3.42
35	0	4	6	193.01	270.21	321.68	4.00	3.88	4.09
36	0	4	7	186.88	261.63	336.38	3.84	3.81	3.86
37	0	4	8	180.02	252.03	352.84	3.56	3.53	3.59
38	0	4	9	172.77	241.88	370.23	3.31	3.17	3.43
39	0	4	10	165.82	232.15	386.92	3.84	4.00	3.70
40	0	4	11	159.53	223.34	402.02	3.55	3.35	3.72
41	0	4	12	151.59	212.22	421.08	3.30	3.21	3.38
42	0	4	13	145.35	203.49	436.05	3.23	3.29	3.17
43	0	5	6	187.00	286.22	311.67	3.19	3.25	3.15
44	0	5	7	181.24	277.41	326.24	3.76	3.93	3.62
45	0	5	8	174.78	267.53	342.58	3.50	3.68	3.35
46	0	5	9	167.95	257.06	359.88	3.16	2.88	3.39
47	0	5	10	161.37	246.99	376.53	3.55	3.47	3.61
48	0	5	11	155.40	237.86	391.62	4.51	4.86	4.22
49	0	5	12	147.86	226.31	410.72	3.42	3.66	3.23
50	0	5	13	141.92	217.22	425.75	3.78	3.42	4.07
51	0	6	7	175.72	292.87	316.30	2.94	2.90	2.97
52	0	6	8	169.64	282.74	332.50	3.89	3.44	4.26
53	0	6	9	163.19	271.99	349.70	3.44	3.15	3.68
54	0	6	10	156.98	261.63	366.28	3.91	4.09	3.77
55	0	6	11	151.33	252.21	381.35	4.29	4.29	4.28

(continued)

TABLE 2. *(continued)*

	Scale Degree			F0 (Hz)			Average Consonance Ratings		
Triad	Tone 1	Tone 2	Tone 3	Tone 1	Tone 2	Tone 3	Overall	Piano	Clarinet
56	0	6	12	144.16	240.27	400.45	3.63	3.45	3.77
57	0	6	13	138.51	230.85	415.53	4.13	4.49	3.84
58	0	7	8	164.89	296.81	323.19	3.25	2.88	3.56
59	0	7	9	158.79	285.83	340.27	3.53	3.27	3.74
60	0	7	10	152.90	275.22	356.77	3.84	4.09	3.63
61	0	7	11	147.54	265.56	371.79	4.15	4.46	3.90
62	0	7	12	140.72	253.29	390.88	3.37	3.39	3.35
63	0	7	13	135.33	243.59	405.98	3.94	4.03	3.87
64	0	8	9	153.81	301.48	329.60	3.18	3.04	3.29
65	0	8	10	148.28	290.63	345.98	3.80	3.59	3.98
66	0	8	11	143.23	280.73	360.93	4.35	4.30	4.40
67	0	8	12	136.79	268.12	379.98	3.32	3.27	3.35
68	0	8	13	131.69	258.12	395.08	3.70	3.38	3.96
69	0	9	10	143.33	307.13	334.43	3.20	3.06	3.32
70	0	9	11	138.60	297.01	349.28	3.48	3.65	3.34
71	0	9	12	132.57	284.08	368.25	3.53	3.46	3.58
72	0	9	13	127.77	273.80	383.32	3.37	3.30	3.43
73	0	10	11	134.09	312.88	337.91	3.35	3.38	3.33
74	0	10	12	128.44	299.69	356.77	3.55	3.57	3.54
75	0	10	13	123.93	289.17	371.79	4.12	4.15	4.10
76	0	11	12	124.63	314.07	346.19	3.17	3.46	2.93
77	0	11	13	120.38	303.36	361.15	4.74	5.38	4.22
78	0	12	13	115.80	321.68	347.41	3.22	3.50	2.99

Note: Scale degrees represent number of BPCJ semitone intervals above the lowest tone (labeled "0"). Overall consonance ratings represent grand means of the ratings tendered by participants for each triad, collapsed across trials. Ratings are also reported separately for participants in the piano ($n = 51$) and clarinet timbre ($n = 59$) conditions.

equal intervals (e.g., Loy, 2006), the harmonicity algorithm used by Bowling et al. (2018a) is only amenable for use with just scales such as the BPCJ. Therefore, the use of the BPCJ enabled a close replication of a key aspect of their methodology.

The methodological approach taken at present may be seen as the inverse of that adopted by McDermott et al. (2016) in their study of consonance perception among the Tsimane'—instead of testing judgments of Western (12-TET-based) chords by non-Western-enculturated listeners, we tested judgments of non-Western (BPCJ-based) chords by Western-enculturated listeners. Specifically, mirroring the procedure of Bowling et al. (2018a), we presented participants with and asked them to rate the consonance of all possible dyadic and triadic combinations of BPCJ notes within the span of a tritave. (We opted not to include a block of tetrad ratings as this would lead to a substantial increase in the number of chords presented—beyond the already increased number of BPCJ relative to Western dyads and triads—raising concerns regarding participant fatigue). Expanding upon Bowling et al.'s procedure, in which chords were only presented to participants in a single piano timbre, participants in the current study were presented with chords in either a piano or a clarinet timbre. (Timbre was also manipulated between- as opposed to within-participants to avoid fatigue—administering both conditions to each participant would have entailed a doubling of the number of stimuli presented). The motivation for adding a condition with a clarinet timbre was two-fold: First, Bowling et al. (2018b, p. E4958) have argued that "differences in timbre have little effect on consonance, provided that spectra are harmonic." We sought to test this proposition empirically by assessing whether the association between harmonicity and perceived consonance was equivalent when a timbre quite distinct from that used in their original study was employed. Second, we specifically opted to use a clarinet timbre as the odd-numbered partials of tones using this timbre tend to have higher amplitude than the even-numbered partials. Given that the intervals of the BPCJ are based on odd integer frequency ratios, it has been proposed that chords derived from the scale may therefore sound more consonant when played using timbres, such as the clarinet or a synthesized square wave, that give emphasis to odd-numbered harmonics (Loy, 2006; see also, Sethares, 2005). In line with this notion, we speculated

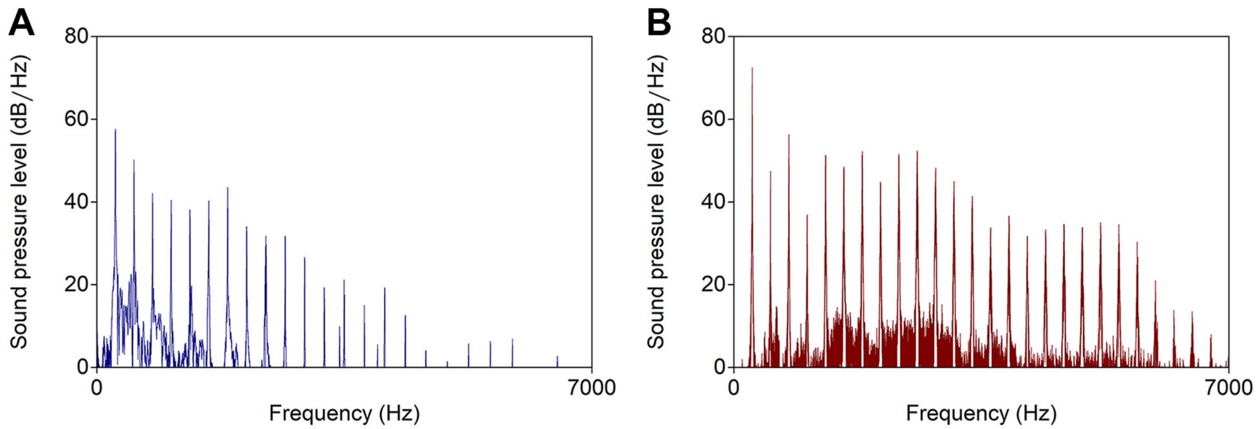

FIGURE 2. Power spectra of examples of the piano timbre (A) and clarinet timbre (B) used in the generation of chord stimuli.

that there might be a more robust association between harmonicity and perceived consonance when BPCJ-based chords were presented using a clarinet as opposed to piano timbre.

As alluded to above, the harmonicities of each BPCJ dyad and triad were computed using the harmonic similarity algorithms devised by Bowling and his colleagues (2018a; Gill & Purves, 2009). We also assessed whether these harmonicity scores were positively associated with consonance ratings of these unfamiliar chords, thereby helping to rule out the possibility that Bowling et al.'s (2018a) effects were artifactually based on the familiarity rather than the harmonicity of the chords they tested and providing more conclusive evidence in favor of their vocal similarity hypothesis. As a secondary means of testing the VSH, we also assessed the correlation between consonance ratings of BPCJ chords and a distinct, frequency interval-based vocal similarity measure developed by Bowling and his colleagues (2018a).

Method

PARTICIPANTS

Participants were 110 undergraduate students at the University at Albany (55 female; Age: M = 18.81, SD = 1.39) who completed the study for course credit in an introductory psychology course. Thirty (27.27%) reported having at least one year of formal training in music theory and sixty-one (55.45%) reported at least one year of formal training on a musical instrument.

MATERIALS

Scale degrees and frequencies of the tones in each dyad and triad are reported in Tables 1 and 2. Following Bowling et al. (2018a), all chords were equated for average pitch height at 261.63 Hz (C4). Test stimuli were created from single-note piano and bass clarinet samples from the University of Iowa's Electronic Music Studios digital archive (Fritts, 2019). (Power spectra of examples of the piano and clarinet tones are presented in Figure 2.) These were pitch shifted and combined into chords using Audacity (v. 2.3.0) software. Specifically, the archived sample with the closest frequency to each BPCJ target pitch was shifted to the frequency of that pitch using Audacity's "Change Pitch" function. This function enables shifting the frequency of an audio sample by manually inputting the desired frequency value into a dialog box. Pitch-shifted samples were then combined into chords using Audacity and then RMS-equated in PRAAT (v. 6.0.43; Boersma & Weenink, 2018). Due to the naturalistic quality of the samples, the length of the finished chords varied for the clarinet condition (from 1.9–2.4 s, depending on how long the original note was played) but remained constant for the piano condition (2.5 s). All audio files, including samples from the digital archive, were mono. Synthesized chords were saved as uncompressed WAV files (sampling rate = 44,100 Hz, bit depth = 32) and were presented via headphones (Koss UR-20) at a fixed volume deemed comfortable in informal testing with undergraduate and graduate research assistants.

HARMONICITY

Following Bowling et al. (2018a), we computed harmonicity scores for each chord by first finding the greatest common divisor (GCD) of the F0s of the tones of a chord made up of intervals with the exact same frequency ratios (e.g., 5/3 and 3/1), yet with tone frequencies set at integer values (e.g., 100 Hz for an exemplar with tones of 300, 500, and 900 Hz) so as to simplify calculations. We then determined the complete harmonic series beginning on this GCD, assigning the highest frequency in this series to the least common multiple (LCM) of the F0s of the chord tones (e.g., 4500 Hz). (As discussed by Bowling et al., 2018a, the LCM is used as a cut-off for purposes of computing harmonicity because it represents the point after which the pattern of harmonics repeats for any given chord. Using a different value risks truncating or extending the harmonic pattern arbitrarily). Finally, we calculated the percentage of harmonics in this complete series that were actually present in the complex spectrum produced by the chord (see Bowling et al., 2018a, and Gill & Purves, 2009, for additional details and see Appendix for a rudimentary worked example). The resulting harmonicity values are presented in the Supplementary Materials (see https://mp.ucpress.edu).

FREQUENCY INTERVALS ANALYSIS

As a secondary means of assessing the VSH, Bowling et al. (2018a) also tested whether chords were perceived as more consonant when the absolute frequency intervals between their tones were more akin to the absolute frequency intervals that tend to occur between harmonics within human vocalizations. According to Bowling et al. (2018a), the minimum absolute frequency interval between adjacent harmonics in human vocalizations is approximately 50 Hz. Consequently, when the F0 s of the tones making up a particular chord form intervals smaller than 50 Hz, this reduces the vocal similarity of this chord, leading it to be perceived as less consonant.

Based on this assumption, Bowling et al. (2018a) made a series of comparisons of consonance ratings among pairs of Western chords in which the minimum interval between the F0 s of at least one of the chords was less than 50 Hz. For to-be-compared chords that met this criterion, they predicted that: (1) when one chord featured a minimum interval between F0 s of less than 50 Hz whereas the second chord featured a minimum interval between F0 s of greater than 50 Hz, the second chord would be perceived as higher in consonance; and, (2) when both chords featured a minimum interval between F0 s of less than 50 Hz, the chord featuring a greater minimum interval would be perceived as higher in consonance. In the present study, we conceptually replicated this frequency intervals analysis using BPCJ dyads and triads.

PROCEDURE

Following Bowling et al. (2018a), consonant chords were defined for participants as "relatively pleasant or attractive" and dissonant chords as "unpleasant or unattractive." Participants then heard a set of conventionally consonant and dissonant chords, after which they rated two 12-TET "practice" chords (a perfect 4th and a minor 7th) on a 7-point Likert scale anchored at 1 (*very dissonant*) and 7 (*very consonant*). Afterward, they were randomly presented with and asked to rate each BPCJ dyad using the same scale. Upon completing 2 blocks of 13 dyad ratings, participants rated all BPCJ triads in 2 consecutive blocks of 78. Participants were randomly assigned to hear all chords in either a piano or a clarinet timbre.

Results

INTERRATER RELIABILITY

Intraclass correlations (ICC; two-way random effects, consistency, average measures model) were first computed to assess interrater reliability in average consonance ratings. These were uniformly high for both dyads (ICC_{piano} = .96; $ICC_{clarinet}$ = .92) and triads (ICC_{piano} = .92; $ICC_{clarinet}$ = .85), suggesting that dataset provides fairly reliable estimates of the consonance values for those chords.

HARMONICITY ANALYSIS

Dyads. Average consonance ratings for dyads appear in Table 1. A scatterplot of the relationship between

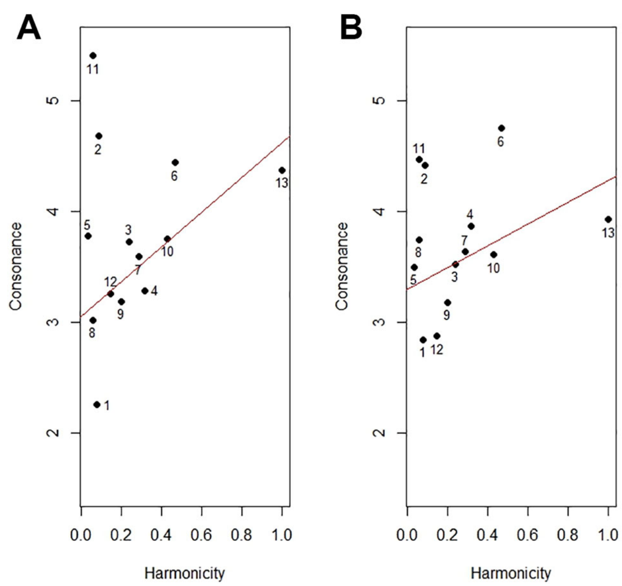

FIGURE 3. Scatterplots of relationship between harmonic similarity scores and consonance ratings of BPCJ dyads. Labeled points refer to dyad number as listed in Table 1. Fitted regression lines exclude quasi-tempered intervals (dyads 2 and 11). (A) Piano timbre. (B) Clarinet timbre.

harmonic similarity scores and consonance ratings of BPCJ dyads appears in Figure 3. Shapiro-Wilk tests revealed that harmonicity scores for dyads were non-normally distributed. Therefore, Spearman rank order correlations were used for our analyses. Inconsistent with the original findings and subsequent predictions of Bowling et al. (2018a, 2018b), initial analyses revealed no statistically reliable association between harmonicity and consonance, irrespective of whether a piano or clarinet timbre was used, $\rho_{\text{overall}} = .19$, $p = .534$.

In trying to account for this failure to replicate, we conducted a number of post hoc exploratory analyses. First, we found that two of the BPCJ dyads that we had presented were particularly similar in their tone ratios to twelve-tone equal tempered (12-TET) dyads. Specifically, the ratio for dyad 2 is 25:21 (1.19) is extremely close to that of the 12-TET minor third (1.189) and the 63:25 (2.52) tone ratio of dyad 11 is within rounding error of that of the 12-TET major tenth (2.5198). Indeed, given their close resemblance to existing 12-TET intervals, these two specific dyads have been previously classified as "quasi-tempered" BPCJ intervals (Haluška, 2004) and both were rated by participants as very high in consonance despite their relatively low harmonicity.

To confirm that the two dyads at issue were indeed more similar to 12-TET intervals than the remaining dyads in the present set, we calculated the absolute difference between each BPCJ dyad's frequency ratio and that of the closest chromatic 12-TET dyad. Scores on this ad hoc "chromatic distance" measure appear in Table 1. They reveal that the ratios for the two so-called quasi-tempered dyads (2 and 11) were indeed the most similar to 12-TET dyads among the intervals tested. A subsequent correlational analysis showed that chromatic distance was reliably predictive of consonance ratings, $\rho = -.58$, $p < .05$; however, when dyads 2 and 11 were dropped from the analysis, this correlation was no longer reliable, $p = .318$. This is consistent with the possibility that consonance ratings for the two quasi-tempered dyads may have been unduly influenced by their familiarity, undermining our ability to detect an effect of harmonicity on consonance.

Based on these findings, we recomputed the raw correlations between harmonicity and consonance, excluding ratings for dyads 2 and 11 from the analysis. Consistent with the correlational analysis reported by Bowling et al. (2018a), the results showed that increased harmonicity was now robustly associated with higher consonance ratings for BPCJ dyads, both overall and within each timbre condition separately, $\rho_{\text{overall}} = .65$, $p < .05$; $\rho_{\text{piano}} = .61$, $p < .05$; $\rho_{\text{clarinet}} = .68$, $p < .05$ (see Figure 3). Following Bowling et al.'s (2018a) analytical strategy, we also computed a series of Bonferroni-adjusted post hoc comparisons between consonance ratings of all dyads included in the aforementioned correlational analysis. In line with results reported for Western dyads by Bowling et al. (2018a), for significant comparisons, the harmonic similarity analysis correctly predicted the chord perceived as more consonant in 89% (24 out of 27) of cases.

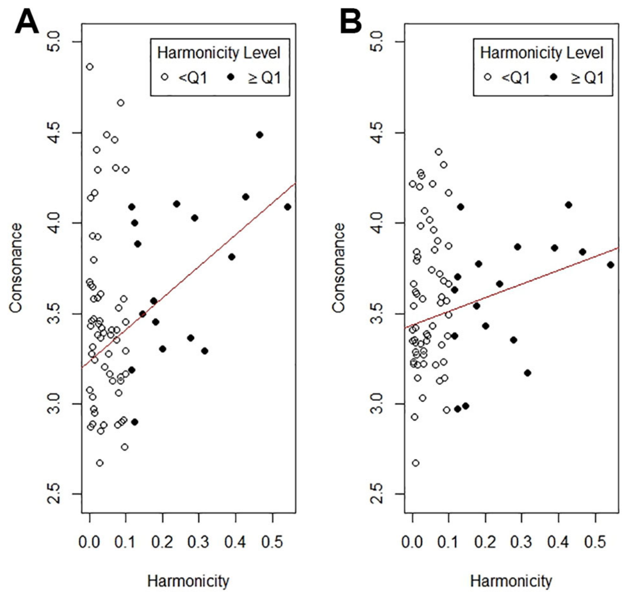

FIGURE 4. Scatterplots of relationship between harmonic similarity scores and consonance ratings of BPCJ triads. Fitted regression lines exclude triads that contain quasi-tempered intervals. (A) Piano timbre. (B) Clarinet timbre.

Triads. Average consonance ratings for dyads appear in Table 1. A scatterplot of the relationship between harmonic similarity scores and consonance ratings of BPCJ triads appears in Figure 4. Shapiro-Wilk tests revealed that consonance ratings as well as harmonicity scores for the triads were non-normally distributed, therefore Spearman rank correlations were used for the analyses reported below. As in the case of BPCJ dyads, these analyses revealed no statistically significant associations between harmonicity and consonance, irrespective of whether a piano or clarinet timbre was used, $\rho_{\text{overall}} = .14$, $p = .217$. These findings fail to replicate Bowling et al.'s (2018a) results, which showed a highly reliable rank order correlation between harmonicity and consonance ratings of conventional triads played in a piano timbre ($\rho = .66$, $p < .001$, according to our calculations from their published data).

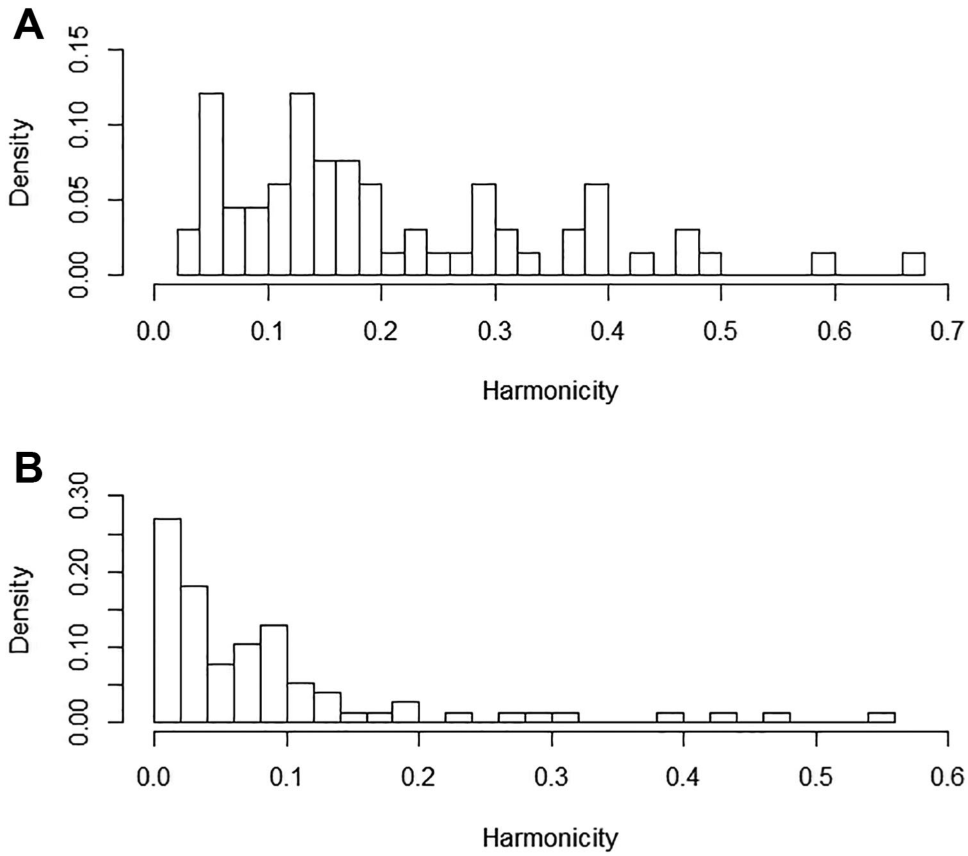

FIGURE 5. (A) Density histogram of harmonic similarity scores for diatonic triads (calculations of Bowling et al., 2018a). (B) Density histogram of harmonic similarity scores for BPCJ triads.

In line with our analysis of BPCJ dyads, we recomputed the harmonicity analyses excluding the 23 of 78 chords that contained dyads 2 and dyad 11 (i.e., the intervals corresponding to the frequency ratios 25/21 and 63/25, respectively; see Table 1). Following these exclusions, correlations between harmonicity and average consonance ratings were now reliable both overall, $\rho = .32$, $p < .05$, and within the piano timbre condition, $\rho = .30$, $p < .05$, but did not attain significance within the clarinet timbre condition, $\rho = .26$, $p = .060$ (see Figure 4). These correlations remained substantially weaker than those reported by Bowling et al. (2018a) for conventional triads.

Another possible contribution to the relative weakness of the foregoing results may be tied to differences between the harmonicity distributions for conventional and BPCJ triads (see Figure 5). Specifically, the mean harmonicity of conventional triads, based on Bowling et al.'s (2018a) data is 0.20 ($SD = 0.16$), whereas that for BPCJ triads is only 0.09 ($SD = 0.11$). Moreover, harmonicity values for BPCJ triads are more positively skewed (skewness = 2.24), that is, more concentrated in the lower end of the distribution, than are conventional triads (skewness = 1.06). Indeed, the first quartile of harmonicity scores for conventional triads (0.103) exceeds the third quartile of harmonicity scores for the BPCJ triads (0.099). This suggests that the association between harmonicity and consonance may emerge more robustly at higher levels of harmonicity. As a post hoc method of evaluating this possibility, we recomputed the regression analysis separately for BPCJ triads with harmonicity values that were at or above the first quartile of harmonicity scores for conventional triads. This value is virtually identical to the mean harmonicity for conventional triads for which common labels (e.g., major, minor, diminished, augmented) are available, which we computed as 0.102 from the data of Bowling et al. (2018a). Within this more conventional range of harmonicity there was a markedly stronger positive correlation between harmonic similarity scores and average consonance ratings, $\rho_{\text{overall}} = .49$, $p < .05$; $\rho_{\text{piano}} = .51$, $p < .05$; $\rho_{\text{clarinet}} = .41$, $p = .103$. The emergence of this more pronounced linear relationship at higher levels of harmonicity is visually apparent in Figure 4. Mirroring the results reported by Bowling et al. (2018a) for Western triads within a similar range of harmonicity, for pairs of BPCJ triads that were significantly different in consonance according to Bonferroni-adjusted post-hoc comparisons, the harmonic similarity analysis correctly predicted the chord perceived as more consonant in 81% (17 out of 21) of cases.

Altogether, the results of our harmonicity analyses do not straightforwardly replicate the findings of Bowling et al. (2018a). However, a similar pattern of results can be detected with unconventional triads, given the important provision that quasi-tempered intervals are excluded or that harmonicity is within a range relatively akin to that of conventional triads. Notably, throughout the series of harmonicity analyses reported above, there were nominal differences between the correlations obtained in the piano and clarinet timbre conditions. However, across analyses, these differences were inconsistent in their direction and magnitude and not a single difference between correlations approached statistical significance (using independent samples tests based on Fisher's r to z transformation). This fails to support our hypothesis that the use of a clarinet timbre should strengthen the correlation between consonance and harmonicity using BPCJ-based chords. However, it does support Bowling et al.'s (2018b) contention that these associations should be relatively unaffected by timbre.

FREQUENCY INTERVAL ANALYSIS

Using the full set of BPCJ stimuli, the frequency interval-based vocal similarity analysis devised by Bowling et al. (2018a) was applicable to 29% of BPCJ dyads and 75% of BPCJ triads. By way of comparison, 44% of Western dyads and 93% of Western triads were eligible to be analyzed using this measure in Bowling et al.'s (2018a) original dataset. (BPCJ-based intervals are generally larger than Western intervals as they are based on the span of a tritave as opposed to an octave. As such, BPCJ chords are less likely to contain intervals smaller than 50 Hz between the F0 s of constituent tones.) The results of the analysis were consistent with the predictions and original findings of Bowling et al. (2018a): For both BPCJ dyads and triads, the chord with higher vocal similarity according the frequency interval measure was rated as more consonant in the majority of cases (61% or 14 out of 23 dyads and 77% or 1735 out of 2262 triads).

SUPPLEMENTARY ANALYSES

Although the primary objective of our study was to conceptually replicate the procedure of Bowling et al. (2018a) using unconventionally tuned chords, we also broadened our analyses to incorporate additional measures that were not originally taken into account in their study. First, although Bowling et al.'s (2018a) pattern matching-based harmonicity measure is elegant in its simplicity, as noted by Harrison and Pearce (2020), its validity may be undercut by a failure to take into account how the auditory system accommodates small deviations in the frequencies of partials when fitting a complex sound wave to a harmonic template during the pitch estimation process. This means that very small changes in the tuning of chord tones may translate into inordinately large changes in harmonicity estimates for that chord. Harrison and Pearce (2020) address this shortfall using a more comprehensive measure that models the information-theoretic uncertainty involved in estimating a fundamental frequency from a set of simultaneously presented complex tones. Essentially, the model builds an idealized representation of the combined spectrum of a chord—one that allows for some modicum of variance in the frequencies of constituent partials—and estimates the extent to which this representation converges upon a candidate for a matching harmonic spectrum among competing harmonic templates (see Harrison & Pearce, 2018, for computational details). In addition, unlike Bowling et al.'s (2018a) harmonicity measure, which provides distinct harmonicity estimates for different voicings of the same chord (e.g., different inversions of a C major triad), Harrison and Pearce's (2020) measure operates on representations of pitch class, rather than on representations of specific pitches, leading it to compute voicing-invariant estimates of harmonicity. This may better reflect the possibility that the auditory system computes similarity between spectra (including that between the complex spectrum of a chord and that of a harmonic template) using pitch class-based spectral representations. In their recent study comparing various harmonicity-based models, Harrison and Pearce (2020) found that their own harmonicity measure was best able to predict consonance ratings. In light of these findings, we reanalyzed consonance ratings for BPCJ chords in our dataset using this alternative measure, assuming that it might offer a more powerful test of the VSH by more realistically capturing how harmonic, and thereby vocal, similarity may be represented psychologically.

Second, theoretically speaking, harmonicity should be negatively correlated with spectral interference, as the frequency distance between harmonic partials renders them less likely to activate overlapping critical bands along the basilar membrane. As discussed earlier, whereas the notion that consonance may result from an absence of auditory roughness due to spectral interference was proposed over a century ago, Bowling et al. (2018a; Bowling & Purves, 2015) have argued against the viability of this hypothesis on both conceptual and empirical grounds. Accordingly, they have excluded measures of spectral interference from their analyses.

However, the results of several recent studies suggest that it may be premature to dismiss the role of spectral interference in judgments of consonance. For instance, in contrast to earlier findings by Cousineau et al. (2012), Marin, Thompson, Gingras, and Stewart (2015) demonstrated that congenital amusics do in fact show systematic preferences between conventionally consonant and dissonant chords and suggested that this effect may be fully attributed to spectral interference. In addition, Parncutt, Reisinger, Fuchs, and Kaiser (2019) proposed that prior studies (e.g., McDermott et al., 2010) may have found relatively weak and inconsistent relationships between spectral interference and consonance because they restricted their investigations to relatively "smooth" chords such as major and minor triads. In line with this possibility, they put forth evidence from a sweeping corpus study of seven centuries of Western polyphonic music, showing that chord prevalence—an indirect measure of consonance—is reliably and uniquely associated with both roughness as well as harmonicity. Similar conclusions have been drawn by Harrison and Pearce (2020), who returned to Bowling et al.'s (2018a) dataset and found that spectral interference was in fact a nominally stronger predictor of consonance than harmonicity, leading them to advocate for a composite model of consonance judgments for isolated chords (see also, Goffinet, 2018).

Most pertinent to the present study, Smit, Milne, Dean, and Weidemann (2019), recently reported the results of a study in which they asked participants to make consonance judgments of both BPCJ as well as Bohlen-Pierce equal tempered (BPET) triads. In contrast with our procedure, Smit et al. (2019) did not attempt to replicate the procedure of Bowling et al. (2018a) and, as such, did not take into account their measure of harmonicity, did not test all possible triads within a tritave, and did not equate the average pitch height of component chord tones (although they did statistically control for absolute pitch height). In addition, they did not manipulate chord timbre and played a series of 13 randomized individual BP chromatic scale tones prior to presentation of each target triad, whereas chords in our study were rated without any added musical context. Despite these differences, Smit et al. (2019) also found evidence that consonance ratings of BP triads were positively correlated with harmonicity, as gauged using a relatively comprehensive measure akin to that of Harrison and Pearce (2020). Moreover, consistent with the work of both Harrison and Pearce (2020) and Parncutt et al. (2019), Smit et al.'s (2019) analyses revealed that that spectral interference was independently (negatively) associated with judgments of consonance, once again supporting a composite model.

Given these recent advances in the field, we conducted a series of supplementary regression analyses including both measures of harmonicity and spectral interference as predictors of consonance ratings. The spectral interference produced by each BPCJ chord in our stimulus set was estimated using Hutchinson and Knopoff's (1978) measure, as implemented by Harrison and Pearce (2020). This models the overall roughness engendered by the interaction of the partials comprising chord tones within critical bands on the basilar membrane (Roederer, 2008). Harrison and Pearce (2020) found that this classic interference measure was in fact the strongest predictor of consonance ratings in Bowling et al.'s (2018a) dataset, contributing to unique variance in these ratings above and beyond harmonicity.

Zero-order correlations between all predictor and outcome variables in our supplementary analyses appear

TABLE 3. *Spearman Correlations between Consonance Ratings for Bohlen-Pierce Dyads and Triads and Harmonicity/Roughness Measures*

Dyads (13)					
Variable	1	2	3	4	5
1.PCR	–				
2.CCR	$.74^{**}$	–			
3.BHM	.13	.29	–		
4.HPHM	.52	$.64^{*}$	.41	–	
5.SIM	$-.80^{**}$	$-.65^{*}$	−.23	−.31	–
Triads (78)					
Variable	1	2	3	4	5
1.PCR	–				
2.CCR	$.68^{**}$	–			
3.BHM	.07	.15	–		
4.HPHM	$.31^{*}$	$.26^{*}$	$.37^{**}$	–	
5.SIM	$-.79^{**}$	$-.77^{**}$	$-.27^{*}$	$-.39^{**}$	–

Note. PCR = average piano consonance ratings; CCR = average bass clarinet consonance ratings; BHM = Bowling et al. (2018a) harmonicity measure; HPHM = Harrison and Pearce (2018) harmonicity measure; SIM = Hutchinson and Knopoff (1978) spectral interference measure. $^{*}p < .05$. $^{**}p < .01$.

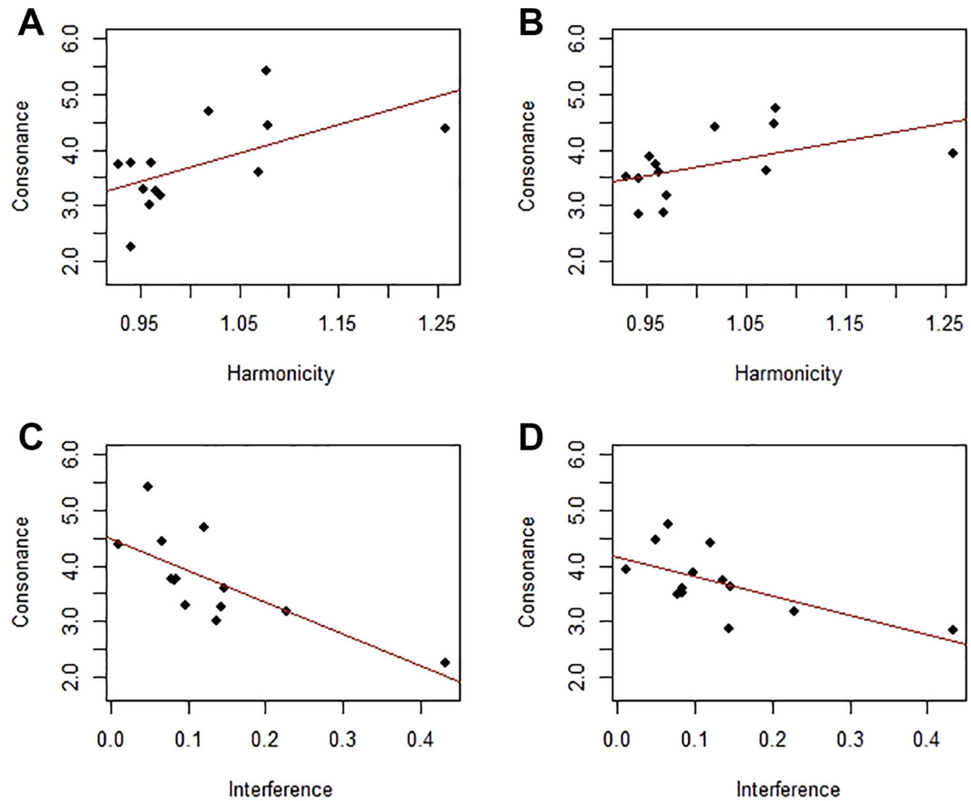

FIGURE 6. Scatterplots of relationship between harmonicity/spectral interference scores and consonance ratings of BPCJ dyads with fitted regression lines. (A) Harmonicity/Piano timbre. (B) Harmonicity/Clarinet timbre. (C) Interference/Piano timbre. (D) Interference/Clarinet timbre.

in Table 3. Harmonicity and spectral interference scores for all BPCJ chords appear in the Supplementary Materials (see https://mp.ucpress.edu). Scatterplots of the relationships between these scores and consonance ratings appear in Figures 6 and 7. In the case of BPCJ dyads, when substituting the harmonicity measure of Harrison and Pearce (2020) for that of Bowling et al. (2018a), the positive, zero-order correlation between harmonicity and consonance was marginally reliable for chords played in a piano timbre, $\rho = .52$, $p = .071$, and significant for those played in a clarinet timbre; $\rho = .64$, $p < .05$, conceptually replicating Bowling et al.'s (2018a) results. We next performed multiple regression analyses on consonance ratings for dyads, entering scores on both Harrison and Pearce's (2020) harmonicity measure as well as Hutchinson and Knopoff's (1978) interference measure as predictors. As shown in Table 4, these analyses revealed negative effects of interference on consonance ratings of chords appearing in both a piano, $\beta = -.60$, $t = -2.76$, $p < .05$, as well as a clarinet timbre, $\beta = -.51$, $t = -1.97$, $p = .076$, suggesting that chords featuring more overall beating between constituent partials were judged as sounding less consonant. However, the effects of harmonicity were no longer statistically significant when shared variance with interference was taken into account.

With respect to BPCJ triads, when substituting the harmonicity measure of Harrison and Pearce (2020), reliable positive correlations between harmonicity and consonance once again emerged both for chords played in a piano timbre, $\rho = .31$, $p < .01$, as well as in a clarinet timbre, $\rho = .26$, $p < .05$. We next conducted multiple regression analyses on consonance ratings for triads, again entering the supplementary measures of harmonicity and interference as predictors. As shown in Table 4, these analyses revealed significant effects of interference on consonance ratings of chords appearing in both a piano, $\beta = -.68$, $t = -8.43$, $p < .001$, as well as a clarinet timbre, $\beta = -.70$, $t = -8.21$, $p < .001$. In addition, there was a significant positive effect of harmonicity on ratings of consonance in triads played in a piano timbre, $\beta = .17$, $t = 2.12$, $p < .05$. The effect of harmonicity on ratings of triads in a clarinet timbre did not approach statistical

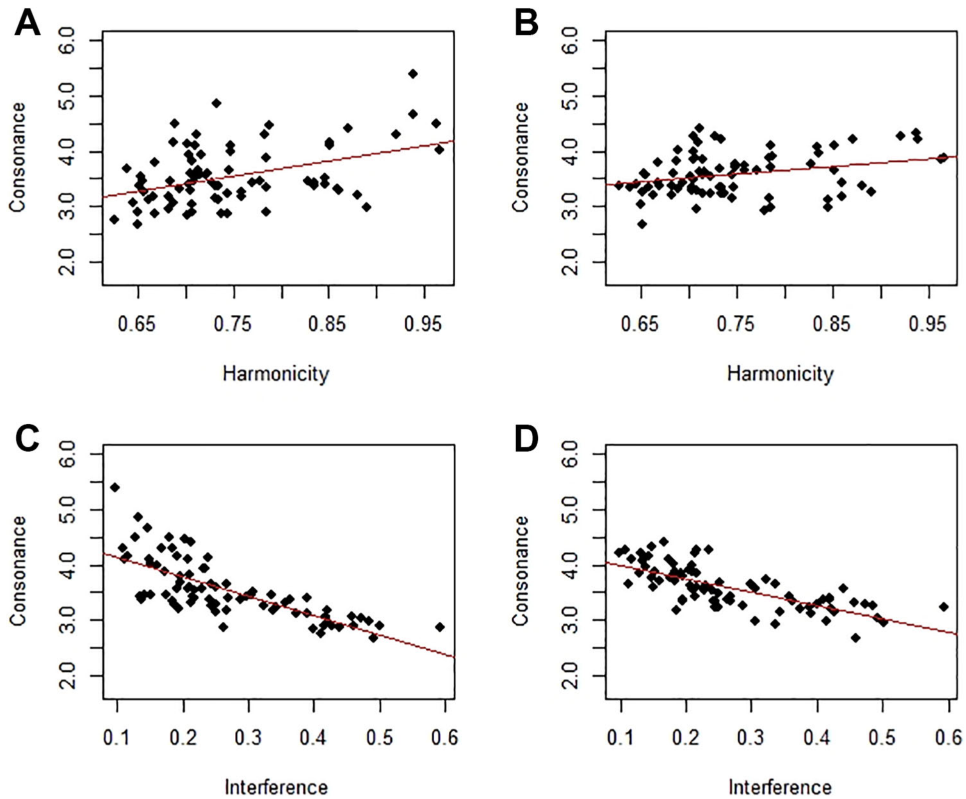

FIGURE 7. Scatterplots of relationship between harmonicity/spectral interference scores and consonance ratings of BPCJ triads with fitted regression lines. (A) Harmonicity/Piano timbre. (B) Harmonicity/Clarinet timbre. (C) Interference/Piano timbre. (D) Interference/Clarinet timbre.

TABLE 4. *Summary of Regression Analyses for Variables Predicting Bohlen-Pierce Chord Consonance Ratings*

	Dyads								Triads							
	Piano				Clarinet				Piano				Clarinet			
Variable	*B*	*SE B*	β	*t*	*B*	*SE B*	β	*t*	*B*	*SE B*	β	*t*	*B*	*SE B*	β	*t*
HPHM	2.73	1.93	0.31	1.41	1.68	1.64	0.27	1.02	1.13	0.53	0.17	2.12*	0.24	0.40	0.05	0.60
SIM	−4.64	1.68	−0.60	−2.76*	−2.82	1.43	−0.51	−1.97	−3.20	0.38	−0.68	−8.43***	−2.37	0.29	−0.70	−8.21***
Adjusted R^2	.55				.35				.57				.51			
F	8.20**				4.24*				51.58***				41.16***			

Note. HPHM = Harrison and Pearce (2018) harmonicity measure; SIM = Hutchinson and Knopoff (1978) spectral interference measure. *$p < .05$. **$p < .01$. ***$p < .001$.

significance when shared variance with interference was taken into account, $p = .554$.[1, 2]

Discussion

In this study, we tested whether the positive correlation between harmonicity and consonance, as recently reported by Bowling et al. (2018a), generalizes to stimuli produced using an unconventional tuning system. The results failed to straightforwardly replicate this

[1] All regression models were checked for conformity with assumptions of linearity, normality (skewness, kurtosis), and heteroscedasticity. For all models, these assumptions were met with a single exception: A partial residual plot of the regression model of ratings of BPCJ triads played in a piano timbre revealed a deviation from linearity. When corrected using a square root transformation of the dependent measure, the results of the analysis were virtually indistinguishable from those reported above.

[2] To confirm our conclusions, we recomputed all multiple regression analyses using a third measure of harmonicity, developed by Stolzenburg (2015; R implemention by Harrison and Pearce, 2020). As shown in Supplementary Table 1, the pattern of findings was identical using this alternative harmonicity measure.

association using the measure of harmonicity originally employed by Bowling et al. (2018a). However, the hypothesized effect did reliably emerge following the post hoc exclusion from the analysis of a set of chords containing intervals that were particularly similar to existing 12-TET dyads. This is consistent with the possibility that consonance ratings for these stimuli were inflated due to familiarity effects, thereby distorting the overall predictive model. Post hoc analyses also indicated that the correlation between harmonicity and consonance for BPCJ triads only began to approach the magnitude of that reported by Bowling et al. (2018a) within a range of harmonicity typical of conventional triads. This may suggest that listeners have difficulty discriminating between chords on the basis of harmonicity when it falls within the comparatively low range characteristic of BPCJ triads. Alternatively, these relatively weak correlations may reveal limitations of the harmonicity measure employed by Bowling et al. (2018a). For instance, the measure does not tolerate near-misses from perfect harmonicity, failing to take into account that the auditory system responds similarly to intervals that deviate slightly from small integer ratios (e.g., Tramo et al., 2001). It also assumes that the constituent tones of a chord contain energy at all partials and in equal measure. These aspects of Bowling et al.'s (2018a) harmonicity measure may diminish its predictive utility when modeling how the auditory system processes naturalistic chord stimuli with a full range of timbral variation.

In line with this proposition, the findings of Bowling et al. (2018a) were conceptually replicated with both BPCJ dyads and triads and newly shown to generalize across both piano and clarinet timbres when harmonicity was measured using the more comprehensive measure recently introduced by Harrison and Pearce (2020). Multiple regression analyses confirmed an independent effect of harmonicity on consonance ratings of BPCJ chords, but only for triads presented in a piano timbre. The majority of variance in consonance ratings across all conditions was not uniquely accounted for by harmonicity, but by spectral interference. Alongside evidence that listeners may be induced to prefer inharmonic spectra when these are generated to minimize beating between partials (Sethares, 2005), these results argue against the contention that "roughness models do not explain consonance" (Bowling et al., 2018b, p. E4958). Rather, in line with the proposals of Harrison and Pearce (2020) and Parncutt et al. (2019), as well as the recent empirical findings of Smit and colleagues (2019), they support a composite model of consonance in which spectral interference plays a prominent role, as do harmonicity and cultural familiarity (McLachlan et al., 2013).[3] From this perspective, the principal question currently facing consonance researchers is not *which* of these three factors explains consonance, but *when* and *how* do each of them contribute to affective responses to isolated chords.

Secondary analyses using Bowling et al.'s (2018a) frequency interval measure also conceptually replicated their findings using BPCJ chords. However, the implications of these results are relatively unclear due to limitations of this secondary vocal similarity measure. First, Harrison and Pearce (2020) noted that the measure is difficult to distinguish from overall spectral interference, which is also associated with the presence of F0 intervals smaller than 50 Hz. Second, as it currently stands, the frequency interval measure categorizes chord stimuli built on either pure or complex tones as equally similar to human vocalizations, although the latter are invariably composed of complex waveforms.

Finally, the present results failed to support our own hypothesis that the use of a clarinet timbre would lead to a stronger correlation between consonance and harmonicity for BPCJ-based chords. Again, we had proposed that such an effect might arise due to the shared emphasis of BPCJ tones and clarinet-like timbres on odd integer frequency ratios/odd-numbered partials. In retrospect, upon examination of the power spectra for our clarinet-based stimuli (see e.g., Figure 2), it is apparent that these naturalistic sounds, which were generated from recorded samples of individual notes on a clarinet, featured considerable energy in even-numbered harmonics. As such, future research examining how the "fit" between timbre and tuning influences consonance judgments may have to rely on computer-generated stimuli (e.g., synthesized square waves), the spectra of which can be more precisely manipulated (see Sethares, 2005, for a discussion of this approach). As noted above, the failure to detect an effect of timbre is consistent with Bowling et al.'s (2018b) proposal that associations between harmonicity and consonance should be relatively consistent across timbres, provided that the latter are harmonic.

Taken together, the present results converge with those of Bowling et al. (2018a) to offer additional support for the VSH, yet suggest that further research is needed to assess whether preferences for isolated chords are indeed

[3] Recent work by Lahdelma and Eerola (2020) suggests that cultural familiarity not only contributes to consonance judgments but may alter how listeners conceptualize terms such as "consonance", "pleasantness", and "tension" when asked to judge the qualities of isolated chords. On the basis of their findings, Lahdelma and Eerola (2020) recommend controlling for cultural familiarity in all research on consonance perception. The use of chords derived from an alternative tuning system such as the BPCJ may be particularly useful in this regard.

at least partly driven by an evolved preference for sounds that resemble human vocalizations. In conjunction with recent studies reporting independent effects of harmonicity on consonance preference based on both behavioral data (Harrison & Pearce, 2020; Smit et al., 2019) and corpus analyses (Parncutt et al., 2019), the present findings using BPCJ chords do lend credence to the notion that the preference for harmonic, and thereby voice-like, musical sounds is rooted in human auditory neurobiology and does not require prior exposure to these sounds (Bowling & Purves, 2015; Purves, 2017).

Nonetheless, proponents of the VSH will also need to contend with ostensibly inconsistent evidence, including McDermott et al.'s (2016) provocative recent findings of weaker or non-existent preferences for harmonic chords in individuals with limited exposure to Western harmony (cf. Bowling, Hoeschele, Gill, & Fitch, 2017) as well as the well-documented prevalence of inharmonic timbres within certain musical systems (e.g., Indonesian gamelan, Thai classical music; Sethares, 2005). These observations would suggest that the preference for harmonicity does require enculturation within a musical system that reinforces attention to and valuation of harmonic sounds (see also, Plack, 2010).

In conclusion, the results of the present study suggest that Bowling et al.'s (2018a) correlational findings in support of the VSH were not merely an artifact of the familiarity or timbre of the chords that they used in their procedure. Our findings also converge with those of Harrison and Pearce (2020), Parncutt et al. (2019), and Smit et al. (2019), by demonstrating that consonance ratings of unconventionally tuned chords are independently associated not only with harmonicity, but with spectral interference as well. We hope that the behavioral data available from this large-scale study, which uniquely includes responses to both dyads and triads in two distinct timbres, will help inspire additional research aimed at assessing the VSH and more generally assist scholars working to develop more comprehensive models of musical consonance.

Author Note

We are grateful to Carly Feldstein, Faith Gaillard, Anthony Leva, and Ryley Scott for their assistance in running these experiments and to Daniel Bowling for his invaluable comments and suggestions.

Correspondence concerning this article should be addressed to Ronald S. Friedman, Department of Psychology, University at Albany, State University of New York, 1400 Washington Avenue, Albany, NY 12222. E-mail: rfriedman@albany.edu

References

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, *125*, 238–251. DOI: 10.1093/brain/awf028

Bidelman, G. (2013). The role of the auditory brainstem in processing musically relevant pitch. *Frontiers in Psychology*, *4*, 1–13. DOI: 10.3389/fpsyg.2013.00264

Bidelman, G., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *The Journal of Neuroscience*, *29*, 13165–13171. DOI: 10.1523/JNEUROSCI.3900-09.2009

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer* [Computer software]. Version 6.0.43. Retrieved from http://www.praat.org

Bowling, D., Hoeschele, H., Gill, K., & Fitch, W. (2017). The nature and nurture of musical consonance. *Music Perception*, *35*, 188–121. DOI: 10.1525/mp.2017.35.1.118

Bowling, D. L., & Purves, D. (2015). A biological rationale for musical consonance. *Proceedings of the National Academy of Sciences*, *112*, 11155–11160. DOI: 10.1073/pnas.1505768112

Bowling, D. L., Purves, D., & Gill, K. Z. (2018a). Vocal similarity predicts the relative attraction of musical chords. *Proceedings of the National Academy of Sciences*, *115*, 216–221. DOI: 10.1073/pnas.1713206115

Bowling, D. L., Purves, D., & Gill, K. Z. (2018b). Reply to Goffinet: In consonance, old ideas die hard. *Proceedings of the National Academy of Sciences*, *115*, E4958–E4959. DOI: 10.1073/pnas.1805570115

Cariani, P. (2019). Musical intervals, scales, and tunings: Auditory representations and neural codes. In P. J. Renfrow & D. J. Levitin (Eds.), *Foundations in music psychology: Theory and research* (pp. 149–218). Cambridge, MA: MIT Press.

Cousineau, M., McDermott, J. H., & Peretz, I. (2012). The basis of musical consonance as revealed by congenital amusia. *Proceedings of the National Academy of Sciences*, *109*, 19858–19863. DOI: 10.1073/pnas.1207989109

Fritts, L. (2019). Piano and clarinet samples. [Audio database]. *Musical instrument samples*. Retrieved from http://theremin.music.uiowa.edu/MIS.html#

Gill, K. Z., & Purves, D. (2009). A biological rationale for musical scales. *PLoS* One, *4*, e8144. DOI: 10.1371/journal.pone.0008144

Goffinet, J. (2018). Little evidence for the vocal similarity hypothesis. *Proceedings of the National Academy of Sciences*, *115*, E4957. DOI: 10.1073/pnas.1804577115

Guernsey, M. (1928). Consonance and dissonance in music. *American Journal of Psychology, 40*, 173–204. DOI: 10.2307/1414484

Harrison, P. M. C., & Pearce, M. T. (2018). *An energy based generative sequence model for testing sensory theories of Western harmony.* In Proceedings of the 19th International Society for Music Information Retrieval Conference. Paris, France.

Harrison, P. M. C., & Pearce, M. T. (2020). Simultaneous consonance in music perception and composition. *Psychological Review, 127*, 216–244. DOI: 10.1037/rev0000169

Haluška, J. (2004). *The mathematical theory of tone systems.* New York: Marcel Dekker.

Helmholtz, H. L. F. von (1954). *On the sensations of tone* (A. J. Ellis, Trans.). New York: Dover. (Original work published 1877)

Hutchinson, W., & Knopoff, L. (1978). The acoustic component of Western consonance. *Interface*, *7*, 1–29. DOI: 10.1080/09298217808570246

Itoh, K., Suwazono, S., & Nakada, T. (2010). Central auditory processing of noncontextual consonance in music: An evoked potential study. *Journal of the Acoustical Society of America*, *128*, 3781–3787. DOI: 10.1121/1.3500685

Johnson-Laird, P. N., Kang, O. E., & Leong, Y. C. (2012). On musical dissonance. *Music Perception*, *30*, 19–35. DOI: 10.1525/mp.2012.30.1.19

Kameoka, A., & Kuriyagawa, M. (1969). Consonance theory. *Journal of the Acoustical Society of America*, *45*, 1451–1469. DOI: 10.1121/1.1911623

Lahdelma, I., & Eerola, T. (2020). Cultural familiarity and musical expertise impact the pleasantness of consonance/dissonance but not its perceived tension. *Scientific Reports*, *10*, 8693. DOI: 10.1038/s41598-020-65615-8

Loy, G. (2006). *Musimathics: The mathematical foundations of music.* (Vol. 1). Cambridge, MA: MIT Press.

Marin, M. M., Thompson, W. F., Gringras, B., & Stewart, L. (2015). Affective evaluation of simultaneous tone combinations in congenital amusia. *Neuropsychologia*, *78*, 207–220. DOI: 10.1016/j.neuropsychologia.2015.10.004

Mathews, M. V., Pierce, J. R., Reeves, A., & Roberts, L. A. (1988). Theoretical and experimental explorations of the Bohlen-Pierce scale. *Journal of the Acoustical Society of America*, *84*, 1214–1222. DOI: 10.1121/1.396622

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology,* 20, 1035–1041. DOI: 10.1016/j.cub.2010.04.019

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535*, 547–550. DOI: 10.1038/nature18635

McLachlan, N., Marco, D., Light, M., & Wilson, S. (2013). Consonance and pitch. *Journal of Experimental Psychology: General*, *142*, 1142–1158. DOI: 10.1037/a0030830

Meyer, M. (1903). Experimental studies in the psychology of music. *American Journal of Psychology, 14*, 456–476.

Nordmark, J., & Fahlén, L. E. (1988). Beat theories of musical consonance. *KTH Royal Institute of Technology STL-QPSR*, *29*, 111–122.

Parncutt, R. (1989). *Harmony: A psychoacoustical approach.* Berlin: Springer.

Parncutt, R., Reisinger, D., Fuchs, A., & Kaiser, F. (2019). Consonance and prevalence of sonorities in Western polyphony: Roughness, harmonicity, familiarity, evenness, diatonicity. *Journal of New Music Research, 48,* 1–20. DOI: 10.1080/09298215.2018.1477804

Plack, C. J. (2010). Musical consonance: The importance of harmonicity. *Current Biology*, *20*, R476–R478. DOI: 0.1016/j.cub.2010.03.044

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, *38*, 548–560. DOI: 10.1121/1.1909741

Purves, D. (2017). *Music as biology: The tones we like and why.* Cambridge, MA: Harvard University Press.

Roederer, J. G. (2008). *The physics and psychophysics of music: An introduction*. New York: Springer.

Sethares, W. (2005). *Tuning, timbre, spectrum, scale*. New York: Springer. DOI: 10.1007/b138848

Smit, E. A., Milne, A. J., Dean, R. T., & Weidemann, G. (2019) Perception of affect in unfamiliar musical chords. *PLoS ONE*, *14*, e0218570. DOI: 10.1371/journal.pone.0218570

Stolzenburg, F. (2015). Harmony perception by periodicity detection. *Journal of Mathematics and Music*, *9*, 215–238. DOI: 10.1080/17459737.2015.1033024

Terhardt, E. (1984). The concept of musical consonance: A link between music and psychoacoustics. *Music Perception*, *1*, 276–295. DOI: 10.2307/40285261

Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological foundations for the theory of harmony in western tonal music. *Annals of the New York Academy of Sciences*, *930*, 92–116. DOI: 10.1111/j.1749-6632.2001.tb05727.x

Valentine, C. W. (1914). The method of comparison in experiments with musical intervals and the effects of practice on the appreciation of discords. *British Journal of Psychology*, *7*, 118–135. DOI: 10.1111/j.2044-8295.1914.tb00247.x

Vuvan, D. T., Nunes-Silva, M., & Peretz, I. (2015). Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 69,* 186–200. DOI: 10.1016/j.cortex.2015.05.002

Appendix: Worked Example of Harmonicity Computation (Bowling et al., 2018a)

Assuming a triad made up of tones of 300, 500, and 900 Hz, the greatest common divisor (GCD) of the tones is 100 Hz and the least common multiple (LCM) is 4500 Hz.

A full harmonic series based on a F0 of 100 Hz with a cutoff of 4500 Hz would contain partials with the following frequencies:

100	200	**300**	400	**500**	**600**	700	800	**900**	**1000**
1100	**1200**	1300	1400	**1500**	1600	1700	**1800**	1900	**2000**
2100	2200	2300	2400	**2500**	2600	**2700**	2800	2900	**3000**
3100	3200	**3300**	3400	**3500**	**3600**	3700	3800	**3900**	**4000**
4100	**4200**	4300	4400	**4500**					

For each of the constituent chord tones, harmonic series with the same LCM-based cutoff would be as follows:

300 Hz

300	600	900	1200	1500	1800	2100	2400	2700	3000
3300	3600	3900	4200	4500					

500 Hz

500	1000	1500	2000	2500	3000	3500	4000	4500

900 Hz

900	1800	2700	3600	4500

Combining the frequencies present within the latter 3 harmonic series results in a chord spectrum including the following 21 partials:

300	500	600	900	1000	1200	1500	1800	2000	2100
2400	2500	2700	3000	3300	3500	3600	3900	4000	4200
4500									

This set of 21 partials overlaps with 46% of the full set of 45 partials present in the full harmonic series based on the GCD of 100 Hz, leading to a harmonic similarity score of 0.46 (21/45). The overlapping partials are highlighted in bold in the full harmonic series above.