In Tune or Out of Tune: Are Different Instruments and Voice Heard Differently?

By: John M. Geringer, Rebecca B. MacLeod, Justine K. Sasanfar

Geringer, J. M., MacLeod, R. B., & Sansafar, J. (2015). In tune or out of tune: Are different instruments and voice heard differently? *Journal of Research in Music Education,* 63(1), 89-101. doi: 10.1177/0022429415572025

Abstract:

We studied music majors' perception of intonation in accompanied solo performances of trumpet, violin, and voice. We were interested in whether listeners would judge pitch deviations of equal magnitude in the three solo performances as equivalent in intonation. Participants were 150 graduate and undergraduate music majors drawn from two large music schools and included 50 students representing each of the following areas of applied music study: voice, wind instruments, and string instruments. Listeners heard solo trumpet, violin, and soprano performances of *Ave Maria* (Bach/Gounod) accompanied by piano. Pitch performances of the soloists were altered in four sections of the excerpt and deviated in either the sharp or flat direction within a section by 0, 10, 20, and 30 cents relative to the accompaniment. Listeners judged pitch deviations in the flat direction as more out of tune for the trumpet than equivalent alterations of violin and voice, especially for magnitudes of 20 and 30 cents. In sharp direction changes, violin and trumpet were heard similarly and as more out of tune than the vocal soloist. Additionally, differences in pitch judgments were found between listener groups representing different areas of applied study.

Keywords: instruments | intonation | perception | timbre | voice

Article:

Evidence of the value given to tuning and pitch accuracy by teachers and students can be found throughout music behavior. Soloists and chamber and large ensembles at all levels of experience tune before performing. Music contests and festivals include intonation ratings on adjudication forms. Performers evaluate new instruments on the basis of tone and intonation tendencies. Teachers instruct students in alternate key or finger combinations to produce sometimes subtle pitch variations, and so on. Musicians obviously place great emphasis on the ability to "play in tune." Additionally, research suggests that in some contexts musicians may focus attention on intonation more than other elements of music (Geringer & Madsen, 1981, 1989, 1998; Johnson & Geringer, 2007). All the aforementioned beg the question, "What is in tune?" Are there differences between instruments and voice in how intonation is perceived by listeners? Acousticians, psychologists, and music educators have conducted a large number of studies in pitch perception and performance. Reports have addressed abilities from perspectives of both listener and performer and have been summarized elsewhere (Ballard, 2011; Geringer, 2010). Evidence exists that tone quality and intonation appear to interact in listeners' perception, and problems in one area may be confused with errors in the other (Geringer, Madsen, & Dunnigan, 2001; Madsen & Geringer, 1981). For example, listener responses have indicated a propensity to associate sharper intonation with "brighter" tone qualities and flatter intonation with "darker" tone qualities (Geringer & Worthy, 1999; Wapnick & Freeman, 1980; Worthy, 2000). In studies of tuning performance, timbre affected the ability to match pitch (Ely, 1992; Greer, 1970). More recently, Byo, Schlegel, and Clark (2011) investigated timbre and octave of tuning stimuli on tuning accuracy of high school wind players. Tuning responses were least in tune to the tuba tones compared to clarinet, flute, and oboe tones, even though most participants reported that tuning to the tuba was the most common method used when tuning their ensembles.

A few researchers have noted differences in perception of intonation between different instruments and voice. Vurma and Ross (2006) found that trumpet and voice tones were judged as sharper than viola tones with the same fundamental frequency and noted that the salience of pitch may be different for different instruments. Subsequent study showed that such timbre-induced changes in perceived pitch may achieve magnitudes that are likely to lead to conflicts between subjective and fundamental frequency-based pitch assessments (Vurma, Raju, & Kuuda, 2012). Loosen (1995) investigated the effects of participants' experience performing a specific instrument on perception of accurate tuning. Violinists tended to prefer sharper tuning of scales (closer to Pythagorean tuning) compared to pianists who preferred equal-tempered scales. Kopiez (2003) found evidence of a "burn in" effect demonstrated by two professional trumpet players. Their performances were closer to equal temperament than just intonation, which Kopiez attributed to long-term intonation practice with equal temperament.

Researchers have documented that vibrato influences pitch perception, and some musicians have suggested that vibrato masks intonation errors (Metfessel, 1932). Yoo, Sullivan, Moore, and Fujinaga (1998) reported that listeners required more time to determine the pitch of violin vibrato tones compared to tones without vibrato. Recently, Geringer, MacLeod, Madsen, and Napoles (in press) showed that melodic intervals were perceived as more out of tune when there was no vibrato compared to vibrato performances.

Very few investigators have explicitly studied the range of tuning that is acceptable to listeners. For example, Lindgren and Sundberg (cited in Sundberg, 1979) used different tunings of a vocal performance, and musically experienced listeners were asked to identify tuning errors. Their experienced listeners accepted errors of 50 to 70 cents as in tune; listeners were tolerant especially when mistunings were sharp, in metrically unstressed locations, and when found in emotionally prominent points in the song. Fyk (1982) studied how listeners perceived mistuned intervals of a piano within melodic context. She found that the ability to recognize mistuned intervals varied considerably depending on the particular interval, the direction of mistuning, and music training. Perhaps most importantly, the melodic context itself contributed to different widths of intonation tolerance. A more recent study investigated the range of acceptable tuning more directly (van Besouw, Brereton, & Howard, 2008). Using synthesized tones, musicians judged stimuli as "acceptably well-tuned" when between +10 and −15 cents from perfect tuning. Adding vibrato to tones increased the lower limit of the acceptable range by an additional 10 cents.

Hutchins and Peretz (2012) compared tolerance for mistuned synthesized and actual singing tones. Musicians and nonmusicians were less likely to hear tuning differences if they were actual vocal tones rather than synthesized voices. Nonmusicians needed about 50 cents of separation between two sung notes to hear the tuning difference compared to 30 cents of separation in synthesized tones. Prame (1997) established that professional singers in recordings were sometimes ±40 cents or more from intended pitch. Given that these recordings were commercially released, these deviations were either not noticed or judged acceptable. Seashore (1938) and Sundberg (1979) both noted that listeners seem generous and operate in an "interpretive mode" when listening to singing; this tendency was recently called "vocal generosity" (Hutchins, Roquet, & Peretz, 2012).

Hutchins et al. (2012) studied musician and nonmusician responses to pairs of single tones and scale-based melodies performed with the voice or the violin. The final note varied in tuning to the prior context, and listeners judged whether the final note was in tune or not. They found evidence for a vocal generosity effect in both melodic and single tone conditions; greater mistuning was necessary for listeners to perceive that sung tones were out of tune compared to violin tones.

Geringer et al. (in press) compared musicians' discrimination of intonation in *unaccompanied* melodies performed by trumpet, violin, and voice and whether there were differences in melodies performed with and without vibrato. Across the no vibrato stimuli, violin was judged as more out of tune than voice and trumpet whether melodic intervals were in tune, flat, or sharp. However, melodies performed with vibrato were judged differently. Violin was perceived as least in tune for intervals mistuned in the flat direction and trumpet was heard as least in tune for intervals mistuned sharp. Notably, mistuned melodic intervals were judged as less out of tune for vocal performances compared to violin and trumpet performances.

We designed the present study to investigate listener perception of intonation in *accompanied* solo performances of trumpet, voice, and violin. We were interested in whether music students with primary performance experience in voice, wind instruments, or string instruments would judge intonation errors of equal magnitude in the three types of stimuli as equivalent or different in degree of mistuning. Given that investigators have found effects of specific instrumental experience on both performance and perception of intonation, we thought, for example, that perhaps string players' experience in listening to violin intonation would facilitate a heightened acuity to intonation errors in violin performances compared to voice and trumpet performances. On the other hand, since vibrato in trumpet is minimal, especially compared to the magnitude of frequency modulation in voice and violin, perhaps most listeners would hear intonation errors in trumpet more readily. Specifically, we asked whether university music students would judge the intonation in melodies differently between trumpet, violin, and voice soloists when performances of soloists were in tune, sharp, or flat relative to the piano accompaniment.

Method

Participants

Participants in the study were 150 undergraduate and graduate university music students. All were recruited from music classes and ensembles at two large schools of music in the southeastern United States. Female students numbered 86 (57%), and there were 64 (43%) males, percentages that approximate the proportion of females and males enrolled in the two music schools. All students had completed a minimum of two years of college-level music study. We obtained responses from 50 students in each of the following applied areas: voice, wind instruments, and string instruments.

Preparation of Stimuli

We recorded solo trumpet, voice, and violin performances of the first 23 measures of *Ave Maria* (Bach/Gounod) accompanied by piano. We chose this piece because it has sustained notes, legato articulations, and clearly defined phrases. The presence of sustained notes helped to facilitate listeners' perception of intonation. Recordings of the piano accompaniment were made in a large piano teaching studio with a Steinway B (7-foot) grand. The accompaniment was performed by a professional pianist and was recorded in three different keys (D, $E^{\flat}$, and F major) as appropriate for the soloist recordings and to prevent listeners from accommodating to a single tonic when making intonation judgments. We analyzed the piano notes that were used in accompanying the excerpt in all three keys and found that frequencies conformed to usual tuning practices in equal temperament relative to A_4 = 440 Hz (individual notes in the middle two octaves varied by ±5 cents). Recording equipment for both the accompaniment and solo performances included two AKG C1000S condenser microphones and a Tascam HD-P2 digital audio recorder. All performances were recorded at a sampling frequency of 48 kHz with 24-bit resolution.

Soloist performances were recorded in a studio designed for recording small ensembles and solo performers. Performers were brought to the recording room individually and were given time to warmup, accommodate to the room acoustics, and become familiar with the procedures. The prerecorded piano accompaniment was presented to performers by means of headphones so that the solo performances could be recorded in isolation from the accompaniment. One ear was left uncovered by the headphones so that the soloists could hear themselves acoustically. The three soloists were professional performers and were chosen after consultation with applied faculty and based on their ability to perform with excellent intonation. Recordings were made in comfortable tessitura and keys for each soloist: The violinist performed the excerpt in the key of D major, the trumpeter in (concert) $E^{\flat}$ major, and the soprano vocalist in F major. All three performers made multiple recordings until both they and we were satisfied with their performance, particularly regarding intonation.

All recordings were transferred digitally to computer, and the intonation of the performances was analyzed using Praat (Boersma & Weenink, 2012). Performances that conformed closest to the tuning of the piano were identified. We then edited the selected sound files with Adobe Audition (version 4.0) and Auto-Tune (version 7.09) software in order to produce versions that contained pitches deviating no more than 5 cents from the accompaniment. We used these versions as the master files for subsequent sound editing.

The original four-measure piano introduction (without soloist) was truncated to two measures. We then electronically manipulated the overall level of intonation of each soloist in four sections of the excerpt, each being four to six measures in length. The initial four measures of each solo performance (originally measures 5–8 of the piece) and each subsequent section (measures 9–12, 13–17, and 18–23) were altered in the overall pitch levels of the soloists so that the sections as a whole were in tune; 10, 20, or 30 cents sharp; or 10, 20, or 30 cents flat relative to the accompaniment. We used Adobe Audition (version 4.0) software to accomplish the manipulations. Mistuning of sections within a given example were either in the sharp or flat direction, not both. Stimulus examples were constructed in 12 orders with the restriction that not more than two adjacent changes occurred by a progressive magnitude in either direction. For example, the sequence −20, –10, –30, and 0 was an acceptable presentation order; a sequence of 0, –10, –20, and −30 was not. A total of 12 experimental examples were created, 4 for each instrument, 2 of which contained alterations in the flat direction and 2 in the sharp direction. Four orders of presentation were produced, each of which counterbalanced the order of soloists and direction of deviation. Thus, listeners heard 12 presentations of *Ave Maria* (through measure 23), each example of which contained four sections varying in magnitude of pitch alterations presented continuously.

Procedures

Experimental examples were transferred to compact disc and presented to listeners in groups of 5 to 20 using loudspeakers (M-Audio Studiophile AV 40). All participants were asked to complete an informed consent form prior to commencing the listening task. A prepared response sheet asked participants to indicate their instrument/voice and their gender and provided instructions for the listening task:

You will hear performances of a soloist and pianist. The melody line of the soloist is notated for each performance. Excerpts are divided into four sections, indicated by the vertical lines. Please use the following rating scale for all examples, and CIRCLE ONE number that corresponds to your <u>perception of intonation</u> between the soloist and the accompanist FOR THAT SECTION OF THE EXCERPT. You will rate 4 sections for each excerpt. Your rating indicates your opinion of the intonation of that <u>section as a whole.</u> Each of the 4 sections could be in tune or any degree of out-of-tune from mostly in tune to extremely out-of-tune. Wait until the end of the section to make your decision. First, we will do a practice example to make sure that you understand the directions.

We used an 11-point rating scale, anchored with the words *very in tune* at the low point of the rating scale (0), *mostly in tune* (below numbers 2 and 3), *out of tune* at the midpoint (5), *very out* *of tune* (below numbers 7 and 8), and *extremely out of tune* at the high point (10). Listeners were reminded to rate the intonation across the section and not to rate individual notes. A practice example (trumpet soloist playing in a different key) was provided at the beginning of the listening task to allow participants to hear the excerpt and ask questions prior to the experimental examples. Each example on the response sheet consisted of the notated solo melody line for measures 5 through 23 (all were notated in C major), with clear markers between the four sections with the rating scales provided: one rating scale each for measures 5 through 8, 9 through 12, 13 through 17, and 18 through 23. The words "Decide Now" appeared in the last measure of each section.

Results

Raw data consisted of listener ratings of intonation between the soloist and the accompanist. Table 1 displays overall mean intonation ratings for the three instruments across the pitch alterations in sharp and flat directions. It can be seen that the trumpet soloist was judged as most out of tune overall in both the flat ($M = 4.06$) and sharp ($M = 4.27$) directions, and the voice was perceived as least out of tune in both directions. Standard deviations were consistent and ranged from 0.98 to 1.24. Table 1 also exhibits intonation ratings of the three soloists for each magnitude of pitch deviation from the accompaniment. Each progressive increase in magnitude of deviation resulted in higher intonation ratings (indicating increased perception of out of tune). In the flat direction, listeners perceived trumpet as more out of tune than the violin and voice performances. Trumpet deviations of 20 and 30 cents, in particular, were rated as less in tune. Listeners rated the soprano soloist as more in tune than trumpet and violin for the two highest magnitudes of deviation. Deviations of 20 and 30 cents in the sharp direction produced ratings for the trumpet and violin as similar and much more out of tune than the voice for the same magnitudes.

Table 1. Judged Intonation Means for Magnitude of Deviation and Direction of Change.

	Magnitude of Deviation				
Direction of Change	0 Cents	10 Cents	20 Cents	30 Cents	Overall Mean
Flat direction					
Trumpet	2.35	2.79	4.40	6.70	4.06
Violin	2.41	2.41	3.47	5.26	3.39
Voice	2.00	2.46	2.88	4.44	2.94
Sharp direction					
Trumpet	2.58	3.64	4.40	6.46	4.27
Violin	2.46	3.11	4.74	6.27	4.15
Voice	1.91	2.41	2.75	3.89	2.74

We assessed the internal consistency of the rating scale and obtained a Cronbach's coefficient alpha of .94 overall, with measures for violin of .84, .90 for voice, and .89 for trumpet scale items. We used an alpha level of .01 for all statistical comparisons. Preliminary analysis showed that there was no statistical difference between female and male listeners ($F < 1$) or between the presentation orders ($F < 1$), nor did these factors evidence interaction with other variables in the study. There was a violation of the sphericity assumption for the direction of change (flat vs. sharp) variable ($p < .01$); therefore, we employed a multivariate analysis of variance with the two directions as the variates (using the repeated measures routine of the general linear model in SPSS version 20). There was one between-subjects factor (voice, wind, and string participant groups) and two within-subjects variables (the three stimulus instruments and four levels of deviation magnitude). We found no significant difference in intonation ratings between the three participant groups, $F(4, 294) = 2.68$, $p > .03$, but significant multivariate main effects for the stimulus instruments and the magnitudes of deviation ($p < .001$). Although the three-way interaction was not significant, $p > .02$, significant multivariate interactions were found for all the two-way combinations of variables: participant group and stimulus instrument, $F(8, 588) = 6.10$, $p < .001$, $\eta^2_P = .08$; participant group and magnitude of change, $F(12, 882) = 6.50$, $p < .001$, $\eta^2_P = .08$; and stimulus instrument and magnitude of change, $F(12, 1764) = 48.34$, $p < .001$, $\eta^2_P = .25$. Subsequent univariate analyses revealed significant differences for both sharp and flat variates in all three two-way interactions (following the Bonferroni correction for multiple comparisons). For changes in the flat direction, voice students rated the vocal and violin soloists as more out of tune than did the wind and string students, $F(4, 294) = 5.15$, $p= .002$, $\eta^2_P = .07$. All three participant groups rated the trumpet as most out of tune for the flat stimuli. For sharp direction changes, all participant groups rated the trumpet and violin as more out of tune than the vocal soloist, and again, voice students indicated that the vocal soloist was less in tune than did the wind and string students, $F(4, 294) = 8.20$, $p < .001$, $\eta^2_P = .10$. Regarding the interaction of participant group and magnitude of deviation, string students rated in tune stimuli and 10 cent flat deviations as more in tune than wind and voice students. Wind players gave lower ratings (less out of tune) than did string or voice students at 30 cents flat, $F(6, 441) = 9.45$, $p <.001$, $\eta^2_P = .11$. In the sharp direction, string players again rated 0 and 10 cents deviation as more in tune and gave higher ratings (more out of tune) at 20 and 30 cents sharp compared to wind and voice students, $F(6, 441) = 9.28$, $p < .001$, $\eta^2_P = .11$.

Most relevant to the purpose of this study concerns the univariate interactions between stimulus instrument and magnitude of deviation. As represented in Figure 1 (top figure), the three stimuli were rated similarly at both 0 and 10 cents flat. However, at 20 and 30 cents flat, listeners rated the trumpet as more out of tune than violin, which was judged more out of tune than the vocal soloist, $F(6, 882) = 56.15$, $p < .001$, $\eta^2_P = .28$. Figure 1 (bottom figure) illustrates that for deviations in the sharp direction, listeners rated the trumpet and violin similarly at all magnitudes of deviation. The vocal soloist was rated more in tune than the two instruments at all magnitudes, particularly at 20 and 30 cents sharp, $F(6, 882) = 54.76$, $p < .001$, $\eta^2_P = .27$.

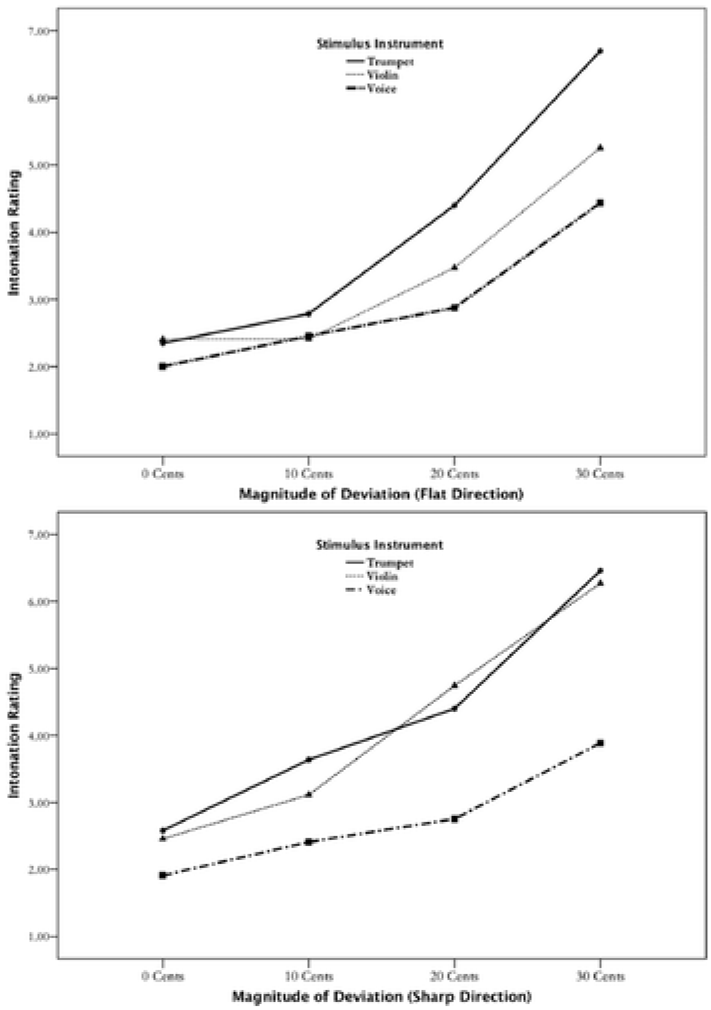

Figure 1. Interactions between stimulus instrument and magnitude of deviation in the flat direction (top figure) and sharp direction (bottom figure).

Note. Higher values indicate perception of greater deviation.

Discussion

We designed the present study to address whether listeners would judge pitch deviations of equal magnitude in three solo performances as equivalent in intonation. Listeners judged phrases that contained very little pitch deviation from the accompaniment similarly between the soloists, although the soprano vocalist was perceived as more in tune than the violinist and trumpeter even with small pitch deviations. Pitches mistuned in the flat direction were perceived as most out of tune when performed by the trumpet at 10, 20, and 30 cents flat. The violin was heard as somewhat more in tune than the trumpet when presented 20 and 30 cents flat; however, the vocalist was judged as most in tune in phrases with these same levels of deviation. Pitch deviations in the sharp direction show little differentiation between the trumpet and violin performances at all magnitudes; however, both were perceived as less in tune than the voice. Voice (soprano) performances were heard as most in tune for phrases that were 10, 20, and 30 cents sharp. Listeners were clearly more tolerant of pitch deviations in the vocalist's performance, particularly in the sharp direction.

Interestingly, mean ratings for even the largest alterations (±30 cents) were just below 7 on the (0–11) rating scale, falling between the *out of tune* and *very out of tune* labels. Two possible explanations occur to us. Although it is clear that music majors discriminated the intonation deviations of 30 cents, apparently they did not regard them as sufficiently mistuned to use the ratings aligned with the label *extremely out of tune*. Perhaps they experience more extreme intonation differences in other contexts. Second, listeners were rating the 4 through 6 measure sections as a whole (not individual tones), and while some notes seem extremely out of tune, other notes are apparently not heard as obviously disparate with the piano accompaniment, corroborating earlier study (Fyk, 1982).

We found that one of the most instructive aspects of this study occurred during the preparation of stimuli. We had spoken with several applied teachers to help identify performers who could provide an initial recording that was as close as possible to being "in tune" with the accompaniment. In the case of the trumpet and violin performers, teachers' choices were quickly confirmed, as both these instrumentalists needed only a few trials to produce a recording that was within 10 cents of the accompaniment. Most notes deviated less than 5 cents from the recorded accompaniment. It was not difficult to subsequently alter (using Auto-Tune or Audition) those few slightly out of tune notes to reach our criterion of ±5 cents relative to the accompaniment. Our experience with obtaining a satisfactory voice recording contrasted markedly. During recording of the first soprano recommended to us, it was our real-time perception that most of the notes were close to the accompaniment, with only a few that apparently would require more than a little pitch correction. However, after analysis of the performed pitches, we were surprised to learn that most of the notes were not within 5 or 10 cents of the target accompaniment but were sometimes 20 or 30 cents away. We then recorded four other singers and obtained similar outcomes and in several cases found more obviously out of tune singing. Eventually we identified a soprano who was recommended by several choir directors. She was able to sing with less deviation from the accompaniment than the others; however, there remained several notes that required pitch correction of 15 to 25 cents.

Both the results of the data analysis and the difficulty we had in finding a singer whose intonation closely approximated the accompaniment appear to fit well with what was recently labeled the "vocal generosity effect" (Hutchins et al., 2012). As noted in the review of literature, Sundberg (1979) agreed with Seashore (1938) that the musical ear is extremely generous and apparently operates in the "interpretive mood" when it listens to singing. Hutchins et al. (2012) found a vocal generosity effect in both melodic and single tone conditions: Their musician and nonmusician participants were 40% more likely to judge a sung note as in tune than its violin equivalent.

Explanations that have been proposed for such differences in listener pitch acuity include those that center on either acoustical factors or top-down factors, such as possible cognitive influences specific to voices. Acoustic features such as vibrato (Geringer et al., in press), timbre, pitch, and spectral variability within tones, long-term average spectra, and attack/decay characteristics either alone or in any combination may serve to affect the range of acceptable tuning tolerated by listeners. For example, differences in long-term average spectra between voice, violin, and trumpet performances in the present study (trumpet had the least number and strength of upper harmonics compared to violin and voice) would seem to predict that listeners might hear the trumpet as lower in pitch than violin or voice for equivalent fundamental frequencies (e.g., Worthy, 2000). This provides one possible explanation for listeners' judging the trumpet as most out of tune for the flat deviations. Proposed cognitive factors appear specific to human voices: Listeners may be more forgiving of mistuning because we recognize that it is a voice (Hutchins et al., 2012). Additionally, we may have more experience with mistuned vocal tones, and perhaps that diversity allows acceptance of a wider range of tuning. As noted previously, variation in vocal timbres among singers might also influence differences in acceptable tuning. Clearly more research is necessary to develop a more detailed explanation for the apparently wider tolerance for vocal intonation.

Regardless of the proposed explanations, it seems consequential for music educators, applied teachers, and performing musicians to realize that listeners apparently apply different standards to different timbres. For example, it would seem useful for a trumpet player to know that it may be more acceptable to deviate on the sharp side of the pitch rather than on the flat side. Further research is necessary to see if that tendency would hold for other brass instruments. We did not include a woodwind instrument in this study, and there may be a number of differences in perception for tuning deviations of equivalent magnitude between the woodwinds given the diversity of note onsets and spectral qualities within this instrument grouping. There was some evidence (in the two-way interaction of participant group and magnitude of deviation) that string players discriminated intonation more accurately than the wind or voice groups of listeners. Perhaps the long-term practice and attention that is given to listening, tuning, and discerning pitches when students learn to play a string instrument may contribute to development of increased pitch acuity. If making more informed perceptual judgments regarding intonation is a goal for other applied areas, and if subsequent research corroborates this conclusion, then devoting greater attention to improvement in discerning pitch accuracy in perception and performance would seem advisable.

We found that differences in perception of intonation accuracy between instruments and voice do appear to exist. Based on this and other recent studies, such judgments apparently are a result of factors beyond inconsistencies in fundamental frequency. Perceptual differences apparently exist between different instruments and voice as a result of a number of additional variables, including characteristics of vibrato (type, rate, and width), timbre, and other factors that warrant further research.

Declaration of Conflicting Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

References

Ballard, D. L. (2011). Relationships between college-level wind instrumentalists' achievement in intonation perception and performance. Bulletin of the Council for Research in Music Education, 187, 19–32. Retrieved from http://www.jstor.org/stable/41162321

Boersma, P., Weenink, D. (2012). Praat: Doing phonetics by computer (Version 5.3.23) [Computer program]. Retrieved from http://www.praat.org/.

Byo, J. L., Schlegel, A. L., Clark, N. A. (2011). Effects of stimulus octave and timbre on the tuning accuracy of secondary school instrumentalists. Journal of Research in Music Education, 58, 316–328. doi:10.1177/0022429410386230

Ely, M. C. (1992). Effects of timbre on college woodwind players' intonation performance and perception. Journal of Research in Music Education, 40, 158–167. doi:10.2307/3345565

Fyk, J. (1982). Perception of mistuned intervals in melodic context. Psychology of Music, Special Issue, 36–41.

Geringer, J. M. (2010). Musicians' preferences for tempo and pitch levels in recorded orchestral music. Journal of Research in Music Education, 58, 294–308. doi:10.1177/0022429410380464

Geringer, J. M., MacLeod, R. B., Madsen, C. K., Napoles, J. (in press). Musicians' perception of melodic intonation in performances with and without vibrato. Psychology of Music. doi:10.1177/0305735614534004

Geringer, J. M., Madsen, C. K. (1981). Verbal and operant discrimination/preference for tone quality and intonation. Psychology of Music, 9, 26–30. doi:10.1177/03057356810090010501

Geringer, J. M., Madsen, C. K. (1989). Pitch and tone quality discrimination and preference: Evidence for a hierarchical model of musical elements. Canadian Music Educator, 30, 29–38.

Geringer, J. M., Madsen, C. K. (1998). Musicians’ ratings of good versus bad vocal and string performances. Journal of Research in Music Education, 46, 522–534. doi:10.2307/3345348

Geringer, J. M., Madsen, C. K., Dunnigan, P. (2001). Trumpet tone quality versus intonation revisited: Two extensions. Bulletin of the Council for Research in Music Education, 148, 65–76. Retrieved from http://www.jstor.org/stable/40319079

Geringer, J. M., Worthy, M. D. (1999). Effects of tone quality changes on intonation and tone quality ratings of high school and college instrumentalists. Journal of Research in Music Education, 47, 135–149. doi:10.2307/3345719

Greer, R. D. (1970). The effect of timbre on brass-wind intonation. Experimental Research in Music: Studies in the Psychology of Music, 6, 65–94.

Hutchins, S., Peretz, I. (2012). A frog in your throat or in your ear? Searching for the causes of poor singing. Journal of Experimental Psychology: General, 141, 76–97. doi:10.1037/a0025064

Hutchins, S., Roquet, C., Peretz, I. (2012). The vocal generosity effect: How bad can your singing be?Music Perception, 30, 147–159. doi:10.1525/mp.2012.30.2.147

Johnson, C. M., Geringer, J. M. (2007). Predicting music majors’ overall ratings of wind band performances: Elements of music. Bulletin of the Council for Research in Music Education, 173, 25–38. Retrieved from http://www.jstor.org/stable/40319468

Kopiez, R. (2003). Intonation of harmonic intervals: Adaptability of expert musicians to equal temperament and just intonation. Music Perception, 20, 383–410. doi:10.1525/mp.2003.20.4.383

Loosen, F. (1995). The effect of musical experience on the conception of accurate tuning. Music Perception, 12, 291–306. doi:10.2307/40286185

Madsen, C. K., Geringer, J. M. (1981). Discrimination between tone quality and intonation in unaccompanied flute/oboe duets. Journal of Research in Music Education, 29, 305–313. doi:10.2307/3345006

Metfessel, M. (1932). The vibrato in artistic voices. In Seashore, C. E. (Ed.), The vibrato: Studies in the psychology of music (pp. 14–117). Iowa City: University of Iowa.

Prame, E. (1997). Vibrato extent and intonation in professional lyric singing. Journal of the Acoustical Society of America, 102, 616–621. doi:10.1121/1.419735

Seashore, C. E. (1938). Psychology of music. New York, NY: McGraw-Hill.

Sundberg, J. (1979). Perception of singing. Speech Transmission Laboratory—Quarterly Progress and Status Report, 20, 1–48.

van Besouw, R. M., Brereton, J. S., Howard, D. M. (2008). Range of tuning tones with and without vibrato. Music Perception, 26, 145–156. doi:10.1525/mp.2008.26.2.145

Vurma, A., Raju, M., Kuuda, A. (2012). Does timbre affect pitch?: Estimations by musicians and non-musicians. Psychology of Music, 39, 291–306. doi:10.1177/0305735610373602

Vurma, A., Ross, J. (2006). Pitch perception of sounds with different timbre. In Baroni, M., Addessi, A. R., Caterina, R., Costa, M. (Eds.), Proceedings of the 9th International Conference on Music Perception and Cognition (pp. 1838–1842). Bologna, Italy: University of Bologna.

Wapnick, J., Freeman, P. (1980). Effects of dark-bright timbral variation on the perception of flatness and sharpness. Journal of Research in Music Education, 28, 176–184. doi:10.2307/3345235

Worthy, M. D. (2000). Effects of tone-quality conditions on perception and performance of pitch among selected wind instrumentalists. Journal of Research in Music Education, 48, 222–236. doi:10.2307/3345395

Yoo, L., Sullivan, D. S., Moore, S., Fujinaga, I. (1998). The effect of vibrato on the response time in determining the pitch relationship of violin tones. In Yi, S. W. (Ed.), Proceedings of the 5th International Conference on Music Perception and Cognition (pp. 209–211). Seoul, Korea: Seoul National University.