Effects of Tone-Quality Changes on Intonation and Tone-Quality Ratings of High School and College Instrumentalists

Author(s): John M. Geringer and Michael D. Worthy

Source: Journal of Research in Music Education, Vol. 47, No. 2 (Summer, 1999), pp. 135-149

Published by: Sage Publications, Inc. on behalf of MENC: The National Association for Music Education

Stable URL: http://www.jstor.org/stable/3345719

Accessed: 21/09/2013 14:13

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.
We investigated the effects of variations in tone quality on listeners' perception of both tone quality and intonation. University music and nonmusic major instrumentalists and high school students participating in instrumental ensembles served as listeners (N = 116). High-quality digital samples of clarinet, trumpet, and trombone tones were used. The original tone quality of each instrument was manipulated to produce experimental stimuli of "bright" and "dark" relative to the unaltered tone quality. Results indicated that the more inexperienced instrumentalists rated stimuli that were relatively "brighter" in tone quality as sharper in intonation, and conversely, stimuli of relatively "darker" tone quality were judged to be flatter in intonation. For the brass instruments, listeners judged the unaltered and bright tones as better in tone quality than tones that were relatively dark. However, for the clarinet tones, the bright tone quality was judged to be worse than unaltered or dark-quality stimuli.

John M. Geringer, University of Texas at Austin
Michael D. Worthy, Castleton State College

Effects of Tone-Quality Changes on Intonation and Tone-Quality Ratings of High School and College Instrumentalists

Learning to play in tune is of paramount importance in instrumental music education. Conductors of student ensembles devote a great deal of instructional time to tuning the ensemble and presenting drills and exercises with the intent of improving intonation. Many educators make use of ensemble tuning procedures that require the active participation of student musicians. The students as
well as the teacher must assess the intonation of tones from multiple instrument timbres as well as a variety of tone qualities within the same timbre. Some pedagogues (see, for example, Kohut, 1973, and Thomson, 1995) offer techniques for improving ensemble intonation by addressing aspects of tone quality. Teachers and students may benefit by greater awareness of factors that affect the perception and performance of pitch.

Helmholtz (1877/1954) suggested that the pitch of a complex tone (like those typically produced by musical instruments) was characterized by the fundamental, due to its strength relative to other partials. Upper partials were thought to contribute to quality, but not to the pitch, of the complex tone. Research in the last 30 years, however, has demonstrated that the strength of the fundamental has little to do with pitch perception, and that middle-range harmonics are very influential in the perception of pitch (Plomp, 1966, 1976; Rasch & Plomp, 1982). Plomp (1966) concluded that for fundamentals below approximately 300 Hz (the D one whole step above middle C), pitch is determined primarily by the fourth and higher harmonics; for fundamentals between 300 Hz and 700 Hz, pitch is a function of the third and higher harmonics; for tones between 700 Hz (approximately the F in the second octave above middle C) and 1400 Hz (F in the third octave above middle C), pitch is determined by the second and higher harmonics; and for tones above 1400 Hz, pitch is determined by the fundamental frequency. Plomp (1976) later added that pitch of a complex tone is derived from difference tones of the lower harmonics, and that higher harmonics have little to do with the pitch of complex tones. Singh and Hirsh (1992) asked listeners to discriminate changes in pitch in tones that differed in fundamental frequency, spectral locus (increased intensity of selected harmonics), or both. Changes in frequency affected the perception of pitch; changes in spectral locus affected the perception of timbre and changes in locus influenced the perception of pitch. Higher spectral loci ("brighter" relative tone quality) were associated with higher pitches.

The interaction of pitch and timbre in perception also has been studied via "Garner classification tasks." In this model (Garner, 1974), correct discriminations or reaction times are compared between single variable changes and two variable changes to determine if there is interaction between the variables. Although Garner originally studied the interaction of attributes in visual classification tasks, this model was applied to auditory stimuli by Crowder (1989), Krumhansl and Iverson (1992), Melara and Marks (1990), Moore, Glasberg, and Proctor (1992), and others. Participants in the 1990 Melara and Marks study, for example, were slower to discriminate changes in pitch and timbre when both changed than when only one variable changed. In their "triplex" model of auditory processing, Melara and Marks suggested that timbre, pitch, and loudness are processed through separate channels that are interconnected at some levels. Listeners are able to select any particular channel, but
the information accessed is affected by information leaked from the other two channels. Krumhansl and Iverson (1992) reported that listeners could not attend to pitch without being influenced by timbre, nor could they attend to timbre without being influenced by pitch. Findings in the second part of their study indicated that timbre may be recalled in absolute terms while pitch is recalled in relative terms. This series of studies demonstrated consistently that pitch judgments are impaired when accompanied by changes in timbre.

Music researchers have investigated relationships between tone quality and intonation that have implications for music teaching. Many ensemble performance situations require performers to evaluate intonation across diverse timbres. In an experiment that involved both perception and performance tasks, Ely (1992) found that university woodwind players were more accurate at detecting pitch deviation in duets of unlike woodwind timbres than in duets of like timbres. Results suggested also that the harmonic structure of the stimulus may be consequential when students and teachers are assessing intonation (see also Greer, 1970). Other studies have dealt with the effects of timbral complexity on the performance and perception of intonation (Cassidy, 1989; Swaffield, 1974), discrimination of tone quality and intonation (Madsen & Flowers, 1981; Madsen & Geringer, 1981), and preferences for tone quality and intonation (Geringer & Madsen, 1981, 1989; Madsen & Geringer, 1976). Pitt (1994) used different timbres (trumpet and piano) and diverse pitches to investigate differences between musically trained and untrained listeners. Musicians had little difficulty identifying separate or concomitant changes in timbre or pitch. However, correct discriminations of pitch by nontrained listeners dropped by 30% in trials in which both variables changed.

Musicians in instrumental sections must also make comparisons of pitch with like timbres. Although timbre may be the same, performers may differ in tone quality, particularly in beginning-level ensembles. In a study closely related to the present investigation, Wapnick and Freeman (1980) had 50 undergraduate music majors listen to 48 pairs of clarinet tones and indicate whether the second tone sounded sharp, flat, or the same as the first tone. The tones had been altered to be either bright or dark relative to the original level, and the second tone was either adjusted to 12 cents sharp or 12 cents flat or was unaltered compared to the first. Analysis of error patterns showed that student musicians associated dark variations with flatness and brightness was associated with sharpness. Students were more accurate in making intonation judgments when there was no change in tone quality.

More research is needed in music contexts regarding the interaction of pitch and tone quality, particularly for different instrumental timbres. The present study was designed to investigate the effects of variations in tone quality on listeners' judgments of both tone quality and intonation. We chose the clarinet and two brass instruments (trumpet and trombone) in order to partially replicate and extend
the investigation of Wapnick and Freeman (1980), who used clarinet examples only. The addition of the brass instruments allowed comparisons of judgments between instrument timbres and octave placement of tones, both relevant to ensemble tuning contexts. Three tone-quality conditions (relatively bright, dark, and unaltered tones) were presented in paired comparisons.

**METHOD**

**Participants**

A total of 116 instrumentalists participated in this study. Thirty-six of the participants were music majors enrolled in music degree programs at a large state university in the southern United States. All music majors were third-, fourth-, or fifth-year undergraduates selected randomly from instrumental wind ensembles. Another 36 college students were nonmusic majors selected randomly from those enrolled in concert bands at the same university. Forty-four high school students (tenth and eleventh graders) were selected from those attending a summer high school band camp.

**Preparation of Stimuli**

Stimulus examples were pairs of tones performed on clarinet, trumpet, and trombone. Stimulus tones were selected from the McGill University Master Samples (Opolko & Wapnick, 1987). These digital samples are high-quality recordings of concert performers on compact disc. The stimulus tones we chose for this study were samples of F₄ (349.2 Hz), a perfect fourth above middle C, for clarinet and trumpet and F₃ (174.61 Hz), one octave lower, for trombone. Digital audiotape was used to record both master and experimental tapes and to present stimuli to listeners.

In addition to the extant tone qualities of the selected samples, the tones were altered electronically to produce relatively bright and dark tone qualities. Altered tones were manipulated with a studio-quality (Urei Model 533) 10-octave band active equalizer to produce changes in the tone quality of stimuli. For the “standard” tone-quality condition, the samples were presented unaltered. For the “bright” quality conditions, the clarinet and trumpet samples were altered by increasing the amplitude by 12 dB of the octave bands centered on 1 kHz, 2 kHz, and 4 kHz. All other octave band settings were maintained at normal settings (0 dB). The “bright” trombone sample was presented with the octave bands centered on 500 Hz, 1 kHz, and 2 kHz, increased by 12 dB. For the “dark” tone-quality conditions, the clarinet and trumpet samples were altered by reducing the amplitude of the same octave bands (centered on 1, 2, and 4 kHz) by 12 dB. Similarly, the trombone sample was changed by reducing the energy in the octaves centered on 500 Hz, 1 kHz, and 2 kHz by 12 dB. These alterations to the stimulus tones correspond approximately
with the third through eleventh harmonics for each of the stimulus tones and thus were designed to influence the perception of pitch (Plomp, 1966, 1976). Figure 1 shows power spectra of the three tone-quality conditions for the trumpet. Recording levels were monitored and input levels adjusted using judges during the recording process so that loudness levels remained approximately constant across the examples. The frequency of all tones was monitored after recording to the experimental tape, and it was verified that the original fundamental frequency of the tones had not been altered by the harmonic intensity manipulations. Duration of the tones was that of the original McGill samples and ranged from 3 to 4 seconds. A series of pilot studies was conducted with music major students, and results established that music students were able to discriminate three distinct tone qualities for each of the instruments (90% correct), that altered and unaltered samples were recognized as those of the instruments they represented, and that loudness levels remained essentially the same across the examples.

When the experimental tape was made, the tones were grouped by instrument to approximate a tuning task in a band rehearsal. Three tape orders were prepared. Four trials of tone pairs from one instrument were presented, followed by four trials of a second instrument, and then four trials from the third. This pattern was repeated in counterbalanced order for the remaining 12 trials. Tone-quality conditions were arranged within instruments in randomized blocks in the six sets of pairs, where members of the pair were different dark/standard, standard/dark, dark/bright, bright/dark, standard/bright, and bright/standard. For control purposes, two sets of same-quality stimulus pairs were also presented for each instrument, that is, two of the following sets: dark/dark, bright/bright, and standard/standard. In order to reduce the number of trials, we did not include all three sets of identical pairs for each instrument. Pilot testing had confirmed the ability of listeners to identify these stimuli as the same. Thus, eight sets of paired stimuli were heard for each of the three instruments. Individual tones of a pair were presented consecutively with a 0.5 second interval between tones. An interval of 5 seconds between presentation of the tone-pair trials allowed listeners adequate time to respond on the prepared answer sheets.

**Procedure**

Participants heard the stimulus tape presented via loudspeakers in small groups of 5 to 10 listeners. Students were asked to listen carefully to the tone quality and intonation of the 24 tone pairs and to compare the second tone to the first member of the pair. Participants responded following each presentation of a pair, and used separate 5-point Likert-type scales for rating tone quality and intonation. The tone-quality scale included the following values: 5 (indicating that the second tone of a pair had a lot better tone quality than the first), 4 (a little better tone quality), 3 (no change in tone quality), 2 (a
Figure 1. Power spectra of trumpet tone stimuli: bright tone quality (top), standard tone quality (middle), and dark tone quality (bottom).

little worse tone quality), and 1 (a lot worse tone quality). The scale for intonation included 5 (indicating that intonation of the second tone was a lot sharper than the first), 4 (a little sharper), 3 (no change in intonation), 2 (a little flatter), and 1 (a lot flatter). Listeners were presented with two practice tones (using oboe tones as examples) immediately prior to beginning the experimental trials.

RESULTS

Six of the 24 trials were control trials in that they presented the identical tone-quality condition as a stimulus pair (dark/dark, standard/standard, and bright/bright pairs were each presented twice). Mean ratings for these trials ranged from 2.93 to 3.09 for tone quality, and from 2.91 to 3.07 for intonation. Ratings of three on the
response sheet indicated no change in tone quality or intonation; thus, these means reveal that little differentiation was made by listeners between the first and second examples of these pairs. Standard deviations of the control trials showed relatively little variation in listeners' judgments; all were less than 0.45. Three-way analyses of variance (ANOVAs) showed that there were no significant differences ($p > .01$) in these trials between subject groups, instruments, or tone-quality condition. Thus, for trials where stimuli were identical, listeners appropriately judged these tone pairs as essentially the same in both tone quality and intonation.

Data for the trials that compared the effects of tone-quality alterations were analyzed with two three-way ANOVAs, one for the judgments of tone quality and one for the intonation responses. A multivariate analysis was not used because analysis of intercorrelations showed a lack of homogeneity across cells between the dependent variables. Both analyses included one between-subjects factor (the three groups of subjects) and two within-subjects factors (the three instruments used as stimuli and the six trials that comprised the tone-quality comparisons). Main effect means for both analyses are presented in Table 1. Means represent the rating of the second tone in comparison to the first tone of a trial.

The analysis of the tone-quality ratings revealed no significant ($p > .01$) main effects of group or instrument. It can be seen from Table 1 that the overall group and instrument tone-quality mean ratings are all similar and slightly less than 3.0. There was a significant effect of the six tone-quality comparisons on tone-quality ratings, $F (5, 565) = 36.32$, $p < .0001$. Subsequent Scheffé post hoc tests showed that listeners rated the comparisons of dark to standard tone quality ($M = 3.57$), bright to standard tone quality ($M = 3.33$), and dark to bright ($M = 3.16$) as significantly different than the comparisons in opposite order: standard to dark ($M = 2.21$), standard to bright ($M = 2.55$), and bright to dark ($M = 2.49$).

However, the effects of tone-quality conditions were not independent of group or instrument. There were significant interactions between group and tone-quality comparisons, $F (10, 565) = 3.03$, $p < .001$, and between instrument and tone quality, $F (10, 1130) = 25.97$, $p < .0001$. Figure 2 illustrates the interaction of participant group with the tone-quality comparisons. Ratings of the nonmusic majors were more extreme than music majors and high school students for four of the trials. Nonmajors rated the standard tone compared to dark and bright qualities higher than did the other two groups and rated the dark tones compared to standard and bright tone qualities lower than the other groups. All three groups gave similar ratings to the comparison of dark to bright qualities. The music majors rated the comparison of standard to bright tone quality relatively higher ($M = 2.94$) than did the nonmajors ($M = 2.37$) and high school ($M = 2.34$) listeners.

Figure 3 shows the interaction of instruments with the quality comparisons. Tone-quality ratings for the trumpet and trombone
### Table 1
Mean Tone Quality and Intonation Ratings

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Tone-quality ratings</th>
<th>Intonation ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school</td>
<td>2.87</td>
<td>2.93</td>
</tr>
<tr>
<td>Nonmusic majors</td>
<td>2.83</td>
<td>3.00</td>
</tr>
<tr>
<td>Music majors</td>
<td>2.96</td>
<td>2.99</td>
</tr>
<tr>
<td><strong>Instrument</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trumpet</td>
<td>2.92</td>
<td>2.94</td>
</tr>
<tr>
<td>Trombone</td>
<td>2.92</td>
<td>2.89</td>
</tr>
<tr>
<td>Clarinet</td>
<td>2.82</td>
<td>3.09</td>
</tr>
<tr>
<td><strong>Trials</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bright to standard</td>
<td>3.33</td>
<td>2.98</td>
</tr>
<tr>
<td>Standard to dark</td>
<td>2.21</td>
<td>2.46</td>
</tr>
<tr>
<td>Bright to dark</td>
<td>2.49</td>
<td>2.55</td>
</tr>
<tr>
<td>Standard to bright</td>
<td>2.55</td>
<td>3.19</td>
</tr>
<tr>
<td>Dark to standard</td>
<td>3.57</td>
<td>3.23</td>
</tr>
<tr>
<td>Dark to bright</td>
<td>3.16</td>
<td>3.43</td>
</tr>
</tbody>
</table>

*Note. For each independent variable, the underlined means are significantly different (*p* < .01) from means not underlined, but are not different from each other. Means not underlined are not different from each other.*

Examples were similar to each other throughout the six trials. Ratings for the clarinet examples were different on all trials except the change from bright to standard tone quality. Clarinet ratings were higher than the brass instruments for two of the comparisons: the change from standard to dark and the change from bright to dark. Lower ratings were given to the three clarinet examples where changes were in the direction of brighter tone quality, that is, changes from standard tone quality to bright, from dark to standard, and from dark to bright. Listeners rated the latter two changes for the brass instruments as better in tone quality than any of the other stimulus pairs. The other interactions were not significant.

Analysis of intonation ratings found no significant differences between participant groups. There was a difference in intonation ratings between the instruments, *F*(2, 226) = 9.99, *p* < .0001. Scheffé tests revealed that the clarinet (*M* = 3.09) was rated significantly different than the trumpet (*M* = 2.94) and trombone (*M* = 2.89). There was also a significant effect of the tone-quality conditions on ratings.
of intonation, $F(5, 565) = 44.83, p < .0001$. As can be seen in Table 1, changes in tone quality in the bright direction—from standard to bright tone quality ($M = 3.19$), from dark to standard ($M = 3.23$), and from dark to bright ($M = 3.43$)—were rated higher (indicating a judgment of sharper intonation) than were changes in the dark direction. Similar to the results with the tone-quality ratings, intonation ratings for the tone quality stimuli were not independent of groups or instruments.

There was a significant interaction of groups and quality changes, $F(10, 565) = 7.75, p < .01$. Figure 4 shows that the music majors changed their intonation ratings minimally across the changes in tone quality. Means were between 2.8 and 3.2, although the direction of changes generally was consistent with the other two groups. Both nonmajors and high school students were more extreme than the

Figure 2. Tone-quality ratings of the participant groups for the six tone-quality comparisons.
majors in their intonation ratings except in the case of the bright to standard tone-quality change. Ratings were notably higher (indicating sharper judged intonation) for these two subject groups for quality changes in the bright direction: from dark to bright, from dark to standard, and from standard to bright. The changes of standard to dark and bright to dark were rated lower (indicating flatter judged intonation) by nonmajor and high school groups than by music majors.

There was also a significant interaction between instruments and quality changes, $F(10, 1130) = 4.14, p < .0001$. Figure 5 illustrates that for most quality changes, the intonation of the two brass instruments was judged similarly. For the three trials that changed to a relatively dark tone quality (from bright to standard, from standard to dark, and from bright to dark), the intonation of the clarinet resulted in judgments similar to that of the brass instruments. However, for the
Figure 4. Intonation ratings of the participant groups for the six tone-quality comparisons.

trials wherein stimuli became more bright (changes from standard to bright, dark to standard, and dark to bright), intonation of the clarinet was judged as more sharp compared to the same quality changes in trumpet and trombone. The other interactions were not significantly different.

DISCUSSION

Results of the study indicate that university and high school instrumentalists generally associated stimuli that were relatively brighter in tone quality with a judgment of sharper intonation, and, conversely, darker tone quality stimuli were judged as flatter in intonation. This tendency is in accordance with the results of Singh and Hirsh (1992) and replicates and extends the findings of Wapnick and Freeman (1980), who studied clarinet tones with undergraduate music majors as subjects. In the present study, the association of bright quality with
perceived sharpness and dark quality with flatness was particularly pronounced for the high school and nonmusic major instrumentalists. The changes in quality did not influence intonation judgments of the music majors to the same magnitude as judgments of those with less experience and training. Comparison pairs of clarinet tones that became relatively bright were judged as particularly sharper compared to the identical changes to trumpet and trombone tones. Tones that became relatively dark in quality were judged to be similar in intonation for the three instruments.

It should be remembered that stimuli were altered only regarding relative magnitudes of harmonic energy and that the frequency of stimuli was not changed. It is possible that the procedure used in the study, that is, asking listeners to judge the intonation of these pairs of tones, may have created an expectation of pitch change. However, the more experienced musicians indicated only minimal perceived

---

*Figure 5.* Intonation ratings of the stimulus instruments for the six tone-quality comparisons.
pitch change, providing evidence that expectation alone did not necessarily alter judgment of intonation. Responses of the high school and college nonmusic major participants showed consistent patterns of perceived pitch change, indicating that judgments were not random. Furthermore, these perceptions should not be interpreted as necessarily "wrong" judgments. As documented in the review of literature, pitch perception of tones is affected by more than fundamental frequency. The consistency across listeners, particularly less experienced participants, makes the case that in an ensemble tuning situation in which tones of like timbre are being compared, perceived differences in intonation may, in fact, result from differences other than frequency, such as tone quality as in the present study, and perhaps additional factors.

Listeners' judgments of tone quality also reveal consistent patterns. Overall, these participants rated the unaltered tone quality higher (indicating relatively better tone quality) than alterations in the bright and dark directions. There was a tendency to rate bright tones higher than dark tones particularly for the brass instruments. The nonmusic major group of listeners rated the unaltered stimuli higher than did high school or music major participants. Tone-quality ratings for the trumpet and trombone stimuli were similar to each other, whereas a different pattern of responses was found for clarinet tones. For the brass instruments, listeners consistently judged the standard and bright tone qualities as better in tone quality than tones that changed to a darker tone quality. However, for the clarinet tone, the bright tone quality was judged as worse than standard or dark-quality stimuli. Standard clarinet quality was preferred to bright quality, and dark was preferred to bright quality.

The present study was structured to approximate a tuning task in an instrumental rehearsal. Listeners heard tones presented in pairs and were asked to judge how the intonation of the second tone differed from the first. Judgments of the high school and college nonmusic majors across clarinet and the two brass instruments heard in this study provide additional evidence of the important effects of different tone qualities on perception of intonation. Particularly at beginning levels of ensemble playing, attempts to tune a tone of "poor" quality may be made by adjusting the length of the instrument, when in fact, the source of the perceived intonation mismatch may be a result of other factors. Tones of beginning performers in ensembles often vary widely in quality, and these differences may lead students to conclude that perceived pitch differences may be addressed immediately and adequately by adjusting tube length. Music teachers, especially those who teach novice musicians, should be aware of these effects and should also make their students aware of them. The use of electronic tuners and other visual feedback devices may aid students in discriminating whether differences are related to mismatches of frequency, tone quality, or both. Music educators should encourage students to be aware of and discriminate characteristic "good" tone quality in the context of performance and
to improve tone qualities that are perceived to be undesirable. However, it should be noted as well that emphasizing tone quality exclusively may not be an optimum teaching strategy. The difficulty in separating tone quality and intonation perceptually has been noted (Bergee, 1995). There is some evidence that pitch changes may be more apparent to listeners than are changes in tone quality. Results of studies by Madsen and Geringer (1976) and Geringer and Madsen (1989) suggest that listeners may exercise a minimum degree of tone-quality discrimination in an accompanied context.

In future studies, researchers might investigate more specifically the effects of different tone qualities on performance of both intonation and tone quality. The paired-comparison format of the present study could be extended to tasks in which two or more tone qualities are presented harmonically rather than sequentially. Furthermore, possible effects of multiple timbres on perception and performance of both intonation and tone quality should be considered (see, for example, Ely, 1992). Other aspects of performance, such as loudness levels, harmonic function, and accompaniment conditions, may also exert influence on perception and performance of intonation and tone quality.

REFERENCES

Madsen, C. K., & Flowers, P. J. (1981). Effect of tuning in an attempt to


Submitted September 15, 1998; accepted February 23, 1999.