An Interval-Size Illusion: The Influence of Timbre on the Perceived Size of Melodic Intervals

Frank A. Russo
University of Toronto

William Forde Thompson
University of Toronto

Please Cite:
Russo, F. A., & Thompson, W. F. (2005). An interval-size illusion: The influence of timbre on the perceived size of melodic intervals. *Perception & Psychophysics* *67*(4), 559-568.
doi:10.3758/BF03193514

Perception & Psychophysics
2005, 67 (4), 559-568

An interval size illusion: The influence of timbre on the perceived size of melodic intervals

FRANK A. RUSSO and WILLIAM FORDE THOMPSON
University of Toronto, Toronto, Ontario, Canada

In four experiments, we investigated the influence of timbre on perceived interval size. In Experiment 1, musically untrained participants heard two successive tones and rated the pitch distance between them. Tones were separated by six or seven semitones and varied in timbre. Pitch changes were accompanied by a congruent timbre change (e.g., ascending interval involving a shift from a dull to a bright timbre), an incongruent timbre change (e.g., ascending interval involving a shift from a bright to a dull timbre), or no timbre change. Ratings of interval size were strongly influenced by timbre. The six-semitone interval with a congruent timbre change was perceived to be larger than the seven-semitone interval with an incongruent timbre change (interval illusion). Experiment 2 revealed similar effects for musically trained participants. In Experiment 3, participants compared the size of two intervals presented one after the other. Effects of timbre were again observed, including evidence of an interval illusion. Experiment 4 confirmed that timbre manipulations did not distort the perceived pitch of tones. Changes in timbre can expand or contract the perceived size of intervals without distorting individual pitches. We discuss processes underlying interval size perception and their relation to pitch perception mechanisms.

One of the most basic findings in music cognition is that listeners are highly sensitive to relative pitch. *Relative pitch* refers to relations between pitches—for example, whether one pitch is higher or lower than another, or whether two pitches are proximate (forming a small pitch interval) or far apart (forming a large pitch interval). From about the age of 6, relative pitch processing is evident for most listeners (Takeuchi & Hulse, 1993). Relative pitch is exemplified by the psychological similarity between two pitch intervals that have the same fundamental frequency ratio but differ in absolute pitch values. For example, the pitch interval formed by tones with fundamental frequencies of 200 and 300 Hz (the pitch distance between the first and second *twinkle*s in "Twinkle, Twinkle Little Star") is psychologically similar to the pitch interval formed by tones with fundamental frequencies of 400 and 600 Hz, and both intervals are defined by the same musical label (perfect fifth).

Musically trained listeners possess explicit knowledge of interval categories that is manifested in the ability to assign labels to intervals (e.g., *perfect fifth* or *octave*). Untrained listeners possess implicit knowledge of interval categories (Smith, Nelson, Grohskopf, & Appleton, 1994). Such knowledge is evident when a familiar melody (e.g., "Happy Birthday") is recognized even though it is sung at a novel pitch level. Untrained listeners can also tell whether a melody is sung poorly or out of tune.

Factors other than fundamental frequency affect relative pitch. For example, intervals with similar harmonic functions in Western music, such as a perfect fifth (seven semitones) and a perfect fourth (five semitones) have greater psychological similarity than intervals with different harmonic functions (Krumhansl, 1979; Schellenberg & Trainor, 1996). Expectations can also affect relative pitch. The same interval will be perceived as slightly larger or smaller depending on whether it is larger or smaller than expected (Russo & Thompson, in press; Shepard & Jordan, 1984) and whether it is moving toward or away from the center of an established pitch range (Russo & Thompson, in press). Moreover, when embedded in a melody, the same interval will create a stronger or weaker melodic accent depending on the rhythmic context (e.g., Boltz & Jones, 1986; Jones, Moynihan, MacKenzie, & Puente, 2002).

The effects of timbre on the perception of pitch are complex. Certain manipulations to the frequency and phase spectrum of a tone can result in changes to the perceived pitch of that tone (e.g., Galembo, Askenfelt, Cuddy, & Russo, 2001; Moore & Glasberg, 1990; Singh & Hirsh, 1992). Other kinds of timbral manipulations do not affect the perceived pitch but influence the speed with which listeners can classify individual tones by pitch (Krumhansl & Iverson, 1992, Experiment 1; Melara & Marks, 1990a; Pitt, 1994, Experiment 2). Timbre can also influence the perception of pitch relations, including pitch discrimination thresholds (Warrier & Zatorre, 2002; but see Semal & Demany, 1991, 1993) and judgments of pitch direction (Singh & Hirsh, 1992).

This research was supported by an operating grant awarded to the second author from the Natural Sciences and Engineering Research Council of Canada. We thank E. Glenn Schellenberg and three anonymous reviewers for comments on an earlier version of this article. Correspondence concerning this article should be addressed to F. A. Russo, Department of Psychology, University of Toronto at Mississauga, 3359 Mississauga Road North, Mississauga, ON, L5L 1C6 Canada (e-mail: frusso@utm.utoronto.ca).

The latter findings, although relevant to our investigation, involved tones with fundamental frequencies that differed by 4% or less, which is less than the frequency difference for the smallest interval used in Western music (the semitone). To date, no research has revealed influences of timbre on the perceived size of intervals used in Western music. In the present investigation, we tested the hypothesis that timbre affects the perceived size of intervals found in music. We also tested the possibility that certain timbral manipulations could result in an interval illusion, so that a six-semitone interval could be perceived as larger than a seven-semitone interval.

EXPERIMENT 1

In Experiment 1, we examined whether the perceived size of pitch intervals can be expanded or contracted depending on the timbres of the component tones. Two intervals from the Western diatonic system, the tritone (six semitones) and the perfect fifth (seven semitones), were presented to participants in ascending and descending directions. Each tone in the interval was presented with one of two timbres: dull or bright. Participants rated the size of melodic intervals. All participants were musically untrained and thus did not possess explicit knowledge of interval categories.

The spectra of the dull and bright timbres (modeled after those used by Warrier & Zatorre, 2002) are represented in Figure 1. The energy in the *dull timbre* was weighted in the lower partials, and the energy in the *bright timbre* was weighted in the higher partials so that the normalized spectral centroid (i.e., f_c/f_0, where f_c is the amplitude-weighted mean of the frequency spectrum and f_0 is the fundamental frequency) of the dull timbre was lower than that of the bright timbre. The tones used to create melodic intervals were assigned dull or bright timbres, yielding four timbral manipulations: dull–dull, dull–bright, bright–dull, and bright–bright. Thus, each melodic interval involved a timbral manipulation that was either static or dynamic. For intervals with a dynamic timbral manipulation, the shift in the normalized spectral centroid was either congruent or incongruent with the shift in pitch. A congruent timbral shift meant that the higher pitched tone in the interval had a high normalized spectral centroid and the lower pitched tone had a low normalized spectral centroid.

We anticipated that interval size ratings would be influenced by timbral manipulations. Specifically, we predicted that intervals with a congruent timbral shift would be perceived to be larger than intervals with an incongruent timbral shift. For example, an ascending interval involving a timbral shift from dull to bright should be perceived as larger than the same interval involving a timbral shift from bright to dull. By making the higher pitched tone in a melodic interval "brighter" in timbre than the lower pitched tone, the distance between the two tones should be emphasized, increasing the psychological size of the interval.

Method

Participants. Thirteen participants with 2 years or less of music instruction ($M = 0.7$ years, $SE = 0.3$) and no continued music activity were recruited from the University of Toronto community. These participants included 10 females and 3 males and ranged in age from 17 to 20 years, with a mean age of 18.9 years. All participants were given partial course credit. No participant reported having abnormal hearing.

Stimuli. The stimuli were digitally synthesized on a PowerMac G4 using SoundEdit 16 software (Macromedia, 2000) and presented over Sennheiser HD280 headphones. Dull and bright timbres differed in the intensity of their 11 harmonically related partials. The sixth partial served as the referent and was measured to be 64 dB SPL at the listener's outer ear. The normalized spectral

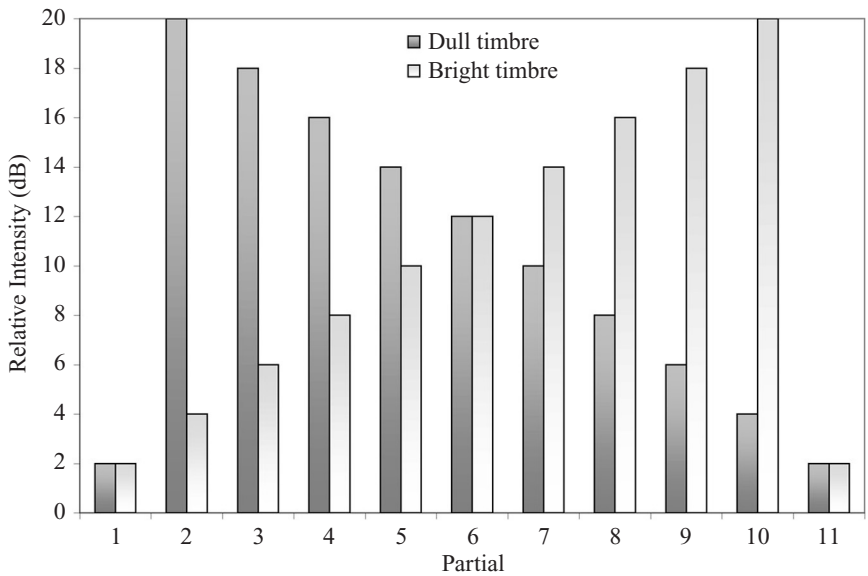

Figure 1. Spectra of dull and bright timbres used in all experiments.

centroid of the dull timbre was approximately a perfect fifth lower than that of the bright timbre (2:2.87).

Individual tones varied in their respective fundamental frequencies from a low of 174.6 Hz (F_3) to a high of 329.6 Hz (E_4). All tones were 1 sec in duration. Tones were always presented in pairs that were separated in pitch by a tritone (six semitones) or a perfect fifth (seven semitones). Each pair was presented in one of four timbral manipulations: (1) incongruent, (2) dull-static, (3) bright-static, or (4) congruent. The names for these timbral manipulations signify the relationship between the shift in the normalized spectral centroid (i.e., timbre) and the shift in fundamental frequency (i.e., pitch). The incongruent condition involved a shift in normalized spectral centroid that was incongruent with the corresponding shift in pitch: from dull to bright for descending intervals and from bright to dull for ascending intervals. The dull-static condition involved a dull timbre for both tones. The bright-static condition involved a bright timbre for both tones. In the congruent condition, the higher pitch in the interval had a bright timbre and the lower pitch had a dull timbre. That is, the timbral shift was from dull to bright for ascending intervals and from bright to dull for descending intervals. The interstimulus interval for all tone pairs was 0 sec.

Procedure. The concepts of interval direction and size were explained to the participants. Their responses involved a judgment about direction (i.e., either up or down) as well as a size rating. The size rating was made on a five-point scale in which a rating of 1 indicated a small interval and a rating of 5 indicated a large interval. The participants were told that they should make a response within 3 sec. If a response was not made within 3 sec, the trial was terminated and the words TIMED OUT were flashed on the screen.

Melodic intervals were presented in ascending and descending directions in each of three possible transpositions (the lower note of an interval was set to one of F_3, G_3, or A_3). A single block of trials encompassed all combinations of two interval sizes (six or seven semitones), two directions (up or down), three transpositions, and four timbral manipulations (incongruent, dull-static, bright-static, congruent). There were three consecutive blocks of trials, yielding 144 trials in total. Trials within each block were independently randomized for each participant. Stimuli were presented and responses collected by a PowerMac G4 running PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993).

Results and Discussion

An alpha level of .05 was used for all statistical tests. Timed-out trials represented 0.8% of all trials in Block 1. In the event that a trial was timed out, the rating for the next equivalent trial was substituted in its place (i.e., from Block 2, or if necessary, Block 3). Pitch direction (ascending or descending) was judged correctly in 97.1% of trials.

Ratings of interval size were subjected to a three-way analysis of variance (ANOVA) with repeated measures on direction (up or down), interval (six or seven semitones), and timbre (incongruent, dull-static, bright-static, or congruent). Figure 2 displays mean size ratings for perfect fifths and tritones in congruent, static, and incongruent timbral manipulations. Means are collapsed across ascending and descending directions because there was no main effect of direction and this factor did not interact significantly with any other factor.

The main effect of interval was not significant [$F(1,12) < 1$]. The lack of differentiation between the tritone and the perfect fifth is surprising, because the ability to judge interval size is the most basic consequence of sensitivity to relative pitch and should be evident even in untrained participants. Two factors may account for the absence of an effect: First, the tritone is a musically unstable interval that is difficult to encode and reproduce, and judgments involving the tritone tend to be associated with a high degree of error (Krumhansl, 1990, 2000). Second, the tritone and perfect fifth intervals are very similar in size, differing by only one semitone. It may have been difficult for listeners with no music training to differentiate intervals so similar in size.

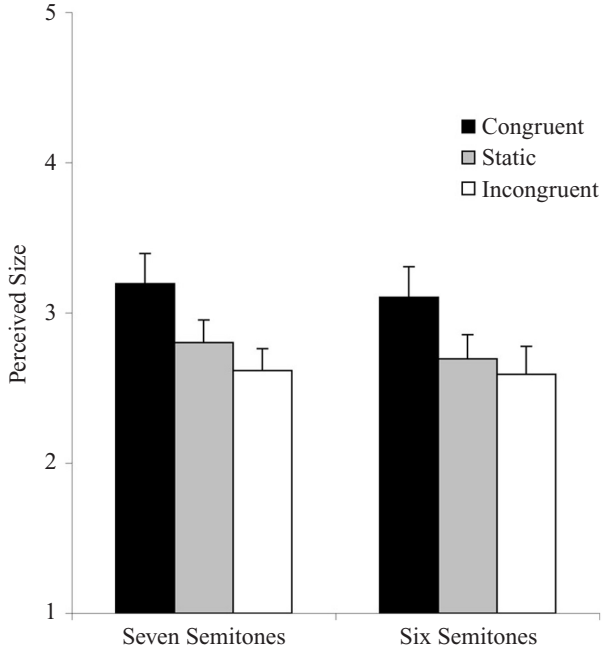

Figure 2. Experiment 1: Mean ratings of interval size by untrained participants for intervals of seven and six semitones, collapsed across ascending and descending intervals.

The main effect of timbre was significant [$F(3,36) = 8.12$, $p < .001$], supporting our primary hypothesis. Intervals presented in the incongruent condition ($M = 2.60$, $SE = 0.16$) were rated as smaller than those in the congruent condition ($M = 3.15$, $SE = 0.13$) [$F(1,12) = 10.43$, $p < .01$]. There was no significant difference in ratings of intervals presented in the dull-static ($M = 2.70$, $SE = 0.16$) and bright-static conditions ($M = 2.79$, $SE = 0.14$) [$F(1,12) = 1.21$, n.s.]. Intervals presented in the congruent condition were rated as larger than those in the static conditions [$F(1,12) = 15.17$, $p < .01$]; however, intervals presented in the static conditions were not rated as significantly larger than those in the incongruent condition [$F(1,12) = 2.23$, $p = .16$].

A planned comparison revealed that the effect of timbre gave rise to an interval illusion. Tritones in the congruent condition ($M = 3.10$, $SE = 0.15$) were rated as significantly larger than perfect fifths in the incongruent condition ($M = 2.62$, $SE = 0.18$) [$F(1,12) = 6.14$, $p < .05$]. This finding indicates that under certain timbral manipulations, an interval of six semitones sounded

significantly larger to untrained listeners than an interval of seven semitones.

Note that the above effects do not reflect absolute differences in timbre between interval tones (which were the same for the congruent and incongruent conditions). If this were the case, both conditions involving a change in timbre (congruent and incongruent) should have been assigned higher ratings than conditions involving no change in timbre. We did not observe this effect. Rather, ratings reflected whether changes in timbre were congruent or incongruent with changes in pitch. The finding suggests that the perception of interval size is not merely determined by the actual pitch difference between the two tones of an interval but includes an evaluation of spectral changes. This evaluation of spectral change appears to be based on change in the normalized spectral centroids of component tones; however, we cannot rule out the possibility that it is based on some other related aspect of the spectra.

Although the finding is striking, it is conceivable that such influences of timbre are limited to musically untrained listeners, who have no knowledge of interval categories and cannot reliably differentiate the two interval sizes presented (six and seven semitones). These factors might have encouraged participants to draw from other differences between tones, such as overall brightness. Experiment 2 was conducted to test this possibility.

EXPERIMENT 2

The results of Experiment 1 indicated that ratings of interval size by untrained participants were strongly influenced by timbre. This effect was so robust that under some conditions an interval of six semitones was rated as larger than an interval of seven semitones. Although the finding is a remarkable instance of an auditory illusion, participants did not have explicit knowledge of interval categories and may have been especially susceptible to influences from acoustic variables other than differences between fundamental frequencies. Musically trained participants might not be susceptible to such effects because they are able to classify intervals of six and seven semitones and can use their knowledge of music as a basis for judging interval size. Experiment 2 was conducted to test this possibility. On the basis of our experience of the effect, however, we predicted that trained participants would experience strong effects of timbre on perceived interval size in spite of their explicit knowledge of interval categories.

Method

Participants. Thirteen participants with a minimum of 10 years of music instruction ($M = 14.3$ years, $SE = 0.9$) and continued activity in music were recruited from the University of Toronto community. These participants included 12 females and 1 male and ranged in age from 18 to 21 years, with a mean age of 19.1 years. All participants were given partial course credit. No participant reported having abnormal hearing.

Stimuli and Procedure. The stimuli and procedure were identical to those for Experiment 1.

Results and Discussion

Timed-out trials represented 1.8% of all trials in the first block. In the event that a trial was timed out, the rating for the next equivalent trial was substituted in its place (i.e., from Block 2 or, if necessary, Block 3). Pitch direction was judged correctly in 99.6% of trials. The data were analyzed as in Experiment 1.

Figure 3 displays mean ratings in the congruent, static, and incongruent conditions for ascending (upper panel) and descending (lower panel) intervals. The main effect of interval was highly significant [$F(1,12) = 87.75$, $p < .0001$], confirming that trained participants could reliably differentiate the size of the two intervals. The strong effect of interval may reflect the explicit knowledge of interval categories possessed by trained participants. There was also a significant interaction between interval and pitch direction [$F(1,12) = 14.13, p < .01$], suggesting that the ability to differentiate the size of the two intervals depended on whether that interval was ascending or descending. As seen in Figure 3, the effect of interval was more pronounced for ascending intervals than for descending intervals; that is, participants were better able to discriminate melodic intervals of six and seven semitones when those intervals involved upward pitch motion.

The main effect of timbre was also significant [$F(3,36) = 16.69$, $p < .0001$], confirming that judgments of interval size were influenced by changes in timbre even for musically trained participants. Intervals presented in the incongruent condition ($M = 2.86$, $SE = 0.16$) were rated as smaller than those in the congruent condition ($M = 3.51$, $SE = 0.13$) [$F(1,12) = 71.83, p < .0001$]. Intervals in the dull-static condition ($M = 3.17$, $SE = 0.16$) were rated similarly to intervals in the bright-static condition ($M = 2.97$, $SE = 0.14$) [$F(1,12) = 4.52$, n.s.]. Intervals presented in the congruent condition were rated as larger than those in the static conditions [$F(1,12) = 24.86$, $p < .0001$], and intervals presented in the static conditions were rated as marginally larger than those in the incongruent condition [$F(1,12) = 4.67$, $p = .05$]. Thus, as with untrained participants, ratings of interval size by trained participants were influenced by timbre. The strong effects of timbre on ratings of interval size are surprising because trained participants are highly attuned to melodic intervals and have explicit knowledge of interval categories.

There was a significant interaction between timbre and pitch direction [$F(3,36) = 3.85, p < .05$], indicating that the effect of timbre on judgments depended on whether the interval was ascending or descending. Specifically, timbral manipulations had a greater influence on judgments of descending intervals than on judgments of ascending intervals. In short, ascending intervals were more accurately differentiated than descending intervals and were less susceptible to the effects of timbre. One

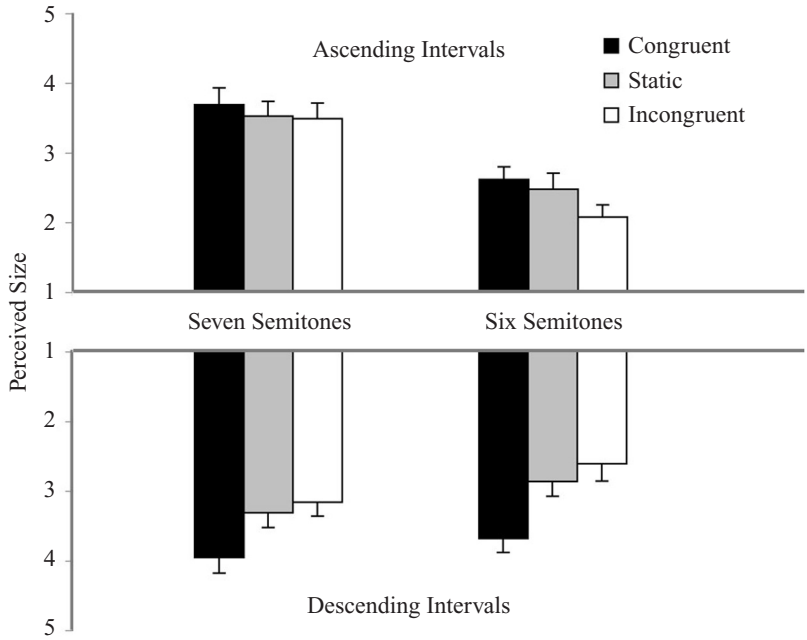

Figure 3. Experiment 2: Mean ratings of interval size by trained participants for intervals of seven and six semitones. Upward bars (top panel) represent ascending intervals, and downward bars (bottom panel) represent descending intervals.

explanation for this effect relates to familiarity. Large intervals such as the tritone and perfect fifth occur more often in the ascending direction than in the descending direction (Meyer, 1973, p. 145; Vos & Troost, 1989). Trained participants may have particularly stable mental representations of these ascending intervals, leading to better discrimination and reduced susceptibility to the effects of timbre.

Indeed, although trained participants were strongly affected by timbral manipulations for both ascending and descending directions, there was no interval illusion observed for ascending intervals. For those intervals, the congruent tritone ($M = 2.67$, $SE = 0.22$) was rated as smaller than the incongruent perfect fifth ($M = 3.45$, $SE = 0.21$) [$F(1,12) = 24.09, p < .001$]. For descending intervals, however, the congruent tritone ($M = 3.69$, $SE = 0.16$) was rated as marginally larger in size than the incongruent perfect fifth ($M = 3.21$, $SE = 0.20$) [$F(1,12) = 4.43$, $p = .06$]. Thus, for descending intervals, there was evidence for an interval illusion even among trained participants.

EXPERIMENT 3

The results of Experiments 1 and 2 revealed strong effects of timbre on ratings of interval size. A comparison of ratings for different conditions revealed a striking illusion in which listeners perceived an interval of six semitones to be larger than an interval of seven semitones. Because our task did not involve a direct comparison of these two intervals, evidence for this illusion was indirect. Experiment 3 was conducted to verify the presence of the interval size illusion by asking participants to make direct comparisons of the size of two intervals presented one after the other.

In each trial, participants judged which of two intervals was larger. As in Experiments 1 and 2, the two intervals were a tritone and a perfect fifth, and each interval involved a shift in timbre from the first to the second tone. Shifts in timbre either emphasized or de-emphasized the difference in size of the two intervals. In the emphasis condition, the tritone had an incongruent timbral shift (making it seem smaller) and the perfect fifth had a congruent timbral shift (making it seem larger). In the de-emphasis condition, the tritone had a congruent timbral shift (making it seem larger) and the perfect fifth had an incongruent timbral shift (making it seem smaller).

Method

Participants. Twenty-eight undergraduate students, 14 musically trained and 14 musically untrained, were recruited from the University of Toronto community. The musically untrained participants had 2 years or less of music instruction ($M = 1.1$, $SE = 0.3$) and no continued music activity. These participants included 11 females and 3 males and ranged in age from 18 to 19 years, with a mean age of 18.4 years. The musically trained participants had 10 years or more of music instruction ($M = 12.5$, $SE = 0.7$) and some continued music activity. These participants included 10 females and 4 males and ranged in age from 18 to 20 years, with a mean age of 18.3 years. All participants were given partial course credit. No participant reported having abnormal hearing.

Stimuli. Presentations were identical to the congruent and incongruent intervals used in Experiments 1 and 2. All trials consisted of two intervals presented sequentially and in the same pitch

direction (i.e., both ascending or both descending). In the emphasis condition, the tritone had an incongruent timbral shift and the perfect fifth had a congruent timbral shift. In the de-emphasis condition, the tritone had a congruent timbral shift and the perfect fifth had an incongruent timbral shift. The interstimulus interval between the last tone of the first interval and the initial tone of the second interval was 1 sec.

Procedure. Each participant completed four blocks of trials consisting of two blocks each of ascending and descending trials. Block presentation adhered to one of two possible orders: (1) ascending, descending, descending, ascending, or (2) descending, ascending, ascending, descending. The two block orders were counterbalanced across participants.

The concepts of interval direction and size were explained to the participants. They were told that their task was to judge whether the first or second interval presented in each trial was larger in size. Response latencies were limited to a maximum of 5 sec. If a response was not made within 3 sec, the phrase RUNNING OUT OF TIME was flashed on the screen. If a response was not made within 5 sec, the phrase OUT OF TIME was flashed on the screen and the trial was terminated.

To discourage strategies based on a comparison of the absolute pitch height of tones, intervals were presented at each of the three transpositions used in Experiment 1 (i.e., the lower pitch of each interval was set to one of F_3, G_3, or A_3). For each transposition condition, the two intervals within a given trial never had the same higher pitch and were presented in both orders (either the fifth or the tritone was presented first). Thus, two timbral conditions (i.e., emphasis or de-emphasis) were crossed with three transposition conditions and two orders, resulting in 12 trials per block. The order of trials within blocks was randomized independently for each participant.

Trials were controlled and responses collected by a PowerMac G4 running PsyScope software (Cohen et al., 1993). The participants heard stimuli over Sennheiser HD280 headphones.

Results and Discussion

There were no timed-out trials. The number of correct responses out of 12 (3 transpositions $\times$ 2 orders $\times$ 2 blocks) was obtained for each combination of pitch direction (ascending and descending) and timbral condition (emphasis and de-emphasis). Figure 4 displays the mean percentages correct in each condition for trained and untrained participants.

Scores were subjected to a 2 $\times$ 2 $\times$ 2 mixed-design ANOVA with timbral condition and pitch direction as the within-subjects variables and training as the between-subjects variable. As predicted, there was a main effect of timbre with higher scores in the emphasis condition than in the de-emphasis condition [$F(1,26) = 33.46, p < .0001$].

There was a main effect of training, with trained participants receiving higher scores than untrained participants [$F(1,26) = 43.16, p < .0001$]. The two-way interaction between training and pitch direction and the three-way interaction among training, timbre, and pitch direction were significant [$F(1,26) = 4.40$ and 5.00, respectively; $p < .05$ for both]. To explore these interactions, orthogonal contrasts were conducted for trained and untrained participants to ascertain differences in discrimination accuracy as a function of timbral condition and pitch direction.

For untrained participants, discrimination accuracy was higher for the timbral emphasis than for the timbral de-emphasis condition in both ascending and descending pitch directions [$F(1,26) = 33.34$ and 13.18, respectively; $p < .01$ for both]. As seen in Figure 4, mean accuracy among untrained participants in the de-emphasis condition was only 41.7%. This finding is consistent with results obtained in Experiment 1 and indicates that untrained listeners experienced an interval illusion in close to 60% of trials in the de-emphasis condition.

For trained participants, discrimination in the ascending pitch direction was extremely accurate and did not

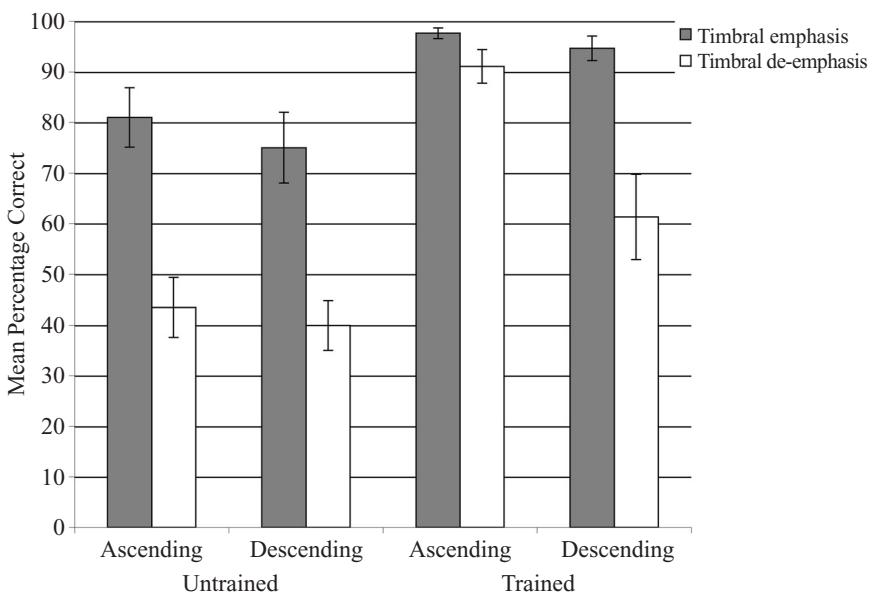

Figure 4. Experiment 3: Mean percentage correct by untrained and trained participants for ascending and descending intervals presented with timbral emphasis and de-emphasis.

differ between the emphasis and the de-emphasis conditions [$F(1,26) < 1$]; however, it is possible that the lack of a difference between these conditions was due to a ceiling effect. Discrimination accuracy in the descending pitch direction was significantly higher in the emphasis condition than in the de-emphasis condition [$F(1,26) = 11.88$, $p < .01$]. Indeed, for descending intervals, mean accuracy among trained participants for the de-emphasis condition was only 61.3%. That is, for descending intervals in the de-emphasis condition, trained listeners experienced an interval illusion in close to 40% of trials. This susceptibility by trained participants to the interval illusion for descending intervals is consistent with the results of Experiment 2. The rate at which trained listeners experienced this illusion is remarkable given that they had explicit knowledge of the intervals being tested and that the intervals were presented one after the other.

To summarize, timbre strongly influenced judgments of interval size for both musically untrained and trained participants. These findings, together with the results obtained in Experiments 1 and 2, suggest that the experience of pitch distance involves not only an evaluation of the difference in pitch of interval tones but also an evaluation of differences in the spectra of component tones.

Still another explanation, however, is that the effect of timbre observed in Experiments 1, 2, and 3 arose merely because timbre manipulations interfered with the ability of participants to extract the fundamental frequencies of the individual tones involved, leading to misperceptions of the pitch of individual tones. Although timbral manipulations were carefully designed to influence timbral brightness without distorting overall pitch, it is conceivable that for some listeners the manipulations resulted in either misperceptions of a few hertz or octave confusions. Such misperceptions of pitch would explain the effects of timbre on perceived interval size without implicating an additional process of evaluating differences in the spectra of individual tones. Experiment 4 was conducted to verify that our timbral manipulations did not distort the pitches of individual tones.

EXPERIMENT 4

Musically trained and untrained participants were asked to perform a pitch-matching task (after Platt, Racine, Stark, & Weiser, 1990). Standard tones possessed either dull or bright timbre and were presented at each of two pitches (F_3, C_4). On each trial, participants adjusted the pitch of a pure tone until it seemed to match the pitch of the standard tone. We expected that pure-tone matches would differ significantly depending on the pitch of the standard tones. Because our timbral manipulations were designed to influence brightness without distorting pitch, however, we did not expect pure-tone matches to differ significantly depending on the timbre of the standard tone.

Method

Participants. The participants were identical to those tested in Experiment 3.

Stimuli and Apparatus. Standard tones were drawn from those tested in Experiments 1 and 2. Specifically, two had a fundamental frequency of 174.61 Hz (F_3) and two had a fundamental frequency of 261.63 Hz (C_4). At each pitch level, one of the two tones had a bright timbre and one had a dull timbre. Pure-tone comparisons were generated using SoundEdit 16 software (Macromedia, 2000).

Procedure. Each participant completed 12 trials (i.e., three repetitions of the four standards). On each trial, the standard tone was presented followed by a pure-tone comparison. Both the standard and comparison tones were 1 sec in duration, and the interstimulus interval was 0.5 sec. The presentation of standard and comparison tones was repeated in a continuous loop until the end of the trial, with 1 sec separating each presentation. The starting frequency of the comparison tone was randomly selected from the frequency range spanning ±6% to ±33% of the standard tone's fundamental frequency. By using a computer mouse, the participants were able to adjust the frequency of the comparison tone up or down in 1-Hz steps. Adjustments were limited to the frequency range spanning one octave above and below the comparison tone's starting frequency. Thus, it was possible for the participants to make octave confusions on every trial. For roughly half of the trials, it was possible to make the kind of octave confusions that would account for the timbre effect observed in Experiments 1 and 2—that is, an octave above the pitch of the standard tone for a bright timbre or an octave below the pitch of the standard tone for a dull timbre. The participants were instructed to use the mouse to click a button displayed on the computer screen called SUBMIT RESPONSE once they were confident that the comparison tone matched the standard tone. Trials were controlled and responses collected with software programmed in REAL Basic (REAL Software, 2003).

Results and Discussion

Matched frequencies for standard tones with dull and bright timbres are plotted in Figure 5. Matched frequencies were subjected to a $2 \times 2 \times 2$ mixed-design ANOVA, with timbre and pitch as the within-subjects variables and training as the between-subjects variable. There was no main effect of timbre, and timbre did not interact with any other factor. An analysis of matching errors revealed that all matches were within one-half octave of the pitch of the standard tone and that over 90% of matches were within three semitones (92.8% and 96.4% for untrained and trained participants, respectively). Notably, matches resembling octave errors never occurred. As predicted, there was a significant main effect of pitch [$F(1,26) = 218.92$, $p < .0001$]. Matched frequencies for C_4 were higher than those for F_3. These findings verify that timbre manipulations did not give rise to minor pitch distortion or to octave confusion. Thus, the effects of timbre on judgments of interval size observed in Experiments 1–3 cannot be explained by misperceptions of pitch but, rather, implicate a process of evaluating spectral differences between component tones.[1]

GENERAL DISCUSSION

Experiment 1 indicated that timbre influenced perceived interval size for musically untrained participants. Under some conditions, intervals of six semitones were perceived to be larger than intervals of seven semitones, giving rise to an interval illusion. Experiment 2 indicated that musically trained participants are also susceptible to

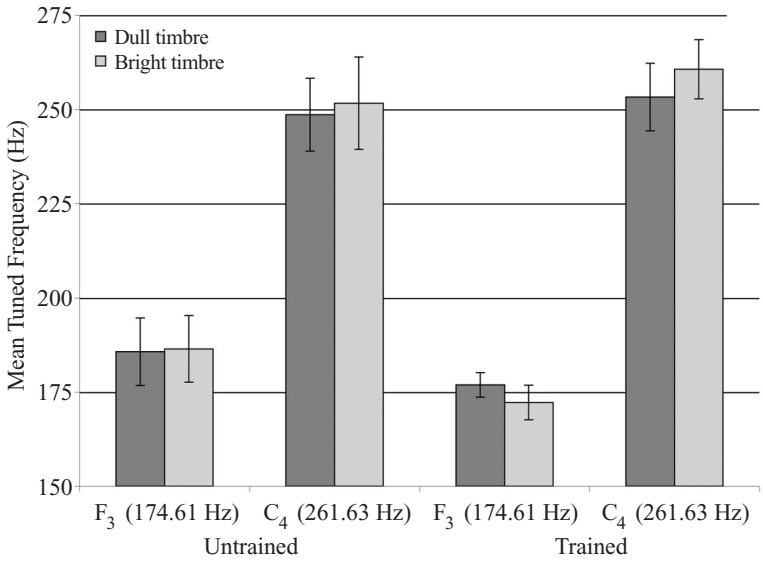

Figure 5. Experiment 4: Mean frequency matches tuned by untrained and trained participants for standard tones of F_3 and C_4 presented with dull and bright timbres.

the effect of timbre on perceived interval size and show evidence of an interval illusion for descending pitch intervals. Experiment 3 verified the findings of Experiments 1 and 2 by asking participants to compare intervals directly and confirmed that the timbral manipulations led to a genuine perceptual illusion.

Importantly, ratings of interval size did not merely reflect global differences in the timbre of interval tones, which were identical for congruent and incongruent conditions. Indeed, ratings of interval size in Experiment 2 (trained participants) were marginally lower for incongruent conditions involving a change in timbre than for conditions involving no change in timbre. Experiment 4 also confirmed that ratings were not the result of listeners' misperceiving the pitch of individual tones. Timbral manipulations had no effect on pitch matching. Instead, our results suggest that the perception of interval size is determined by whether a timbral shift is congruent or incongruent with a pitch shift. When a timbral shift is congruent with a pitch shift (i.e., in the same direction), the perceived size of the interval expands; when it is incongruent with a pitch shift, the perceived size of the interval contracts.

Music theorists have noted that the size of a melodic interval has special aesthetic significance (see, e.g., Meyer, 1956, 1973; Narmour, 1990). To the extent that musicians manipulate timbre in performance for aesthetic purposes, these findings are relevant to performance practice. Although highly mechanized instruments do not allow manipulations of timbre (e.g., the piano), it is possible to vary timbre on many instruments. For example, a singer may produce two notes with the same pitch but with dramatically different timbres. Similarly, a cellist or violinist is capable of altering the brightness of tones. Thus, one implication of the current findings is that performers may be able to manipulate timbre as a way of expanding or contracting the perceived size of melodic intervals, depending on the aesthetic intent.

Another implication of the current findings is that pitch and timbre are not perceived independently. There is a large body of research concerned with interactions among perceptual dimensions. Garner (1974) reviewed a series of findings illustrating that when a stimulus is classified on one dimension, interference may arise from orthogonal variation of an irrelevant dimension. Such dimensions interact and are said to be *integral dimensions*. Conversely, *separable dimensions* show no such interference. Melara and Marks (1990a) observed interactions between pitch and timbre and argued that when an attribute of one dimension is extracted (e.g., a timbre), that attribute creates a "context" in which other dimensions are perceived. For example, the attribute of a high pitch has one perceptual meaning when combined with the attribute of a bright timbre but a different meaning when combined with the attribute of a dull timbre. The context established by extracting a particular nonpitch attribute "acts to weight perceptually the extraction of pitch information" (Melara & Marks, 1990b, p. 399).

The influence of timbre on perceived interval size was best illustrated by the interval illusion. Specifically, intervals of a tritone were rated as larger than intervals of a perfect fifth under certain conditions. Illusions have received considerable attention because perceptual mechanisms are often revealed most clearly under conditions in which veridical perception breaks down (Deutsch, 1975; Thompson, Hall, & Pressing, 2001). The Müller-Lyer illusion—in some respects analogous to the illusion reported here—illustrates that two lines of equal length

can be made to appear different depending on characteristics of their end points. Similarly, Experiments 1 and 2 revealed that two melodic intervals of equal pitch distance can be made to appear smaller or larger depending on the timbres of the tones that define the end points of those intervals.

To our knowledge, this study is the first to report an extra-pitch influence on the perceived size of intervals used in the Western diatonic system. Previous research has indicated that timbre can affect judgments of whether one tone is higher or lower than another (Singh & Hirsh, 1992) and whether two tones differ in pitch (Warrier & Zatorre, 2002), but these effects were observed only for very small intervals involving tones that differed in fundamental frequency by 4% or less. Indeed, Warrier and Zatorre speculated that spectral shape is unlikely to affect pitch judgments when pitches differ in fundamental frequency by more than 4%. Although their prediction may be valid for some tasks (e.g., pitch matching in Experiment 4), we observed large and reliable effects of timbre on the perceived size of intervals used in Western tonal music.

Although timbral manipulations had no effect on the pitch-matching task, they influenced perceived relations between pitches. This pattern of results may relate to the mode of listening emphasized by the task. It has often been suggested that at least two mechanisms are involved in pitch perception (e.g., Grimault, Micheyl, Carlyon, & Collet, 2002; Terhardt, Stoll, & Seewann, 1982). The first is referred to as *place* or *spectral pitch* and the second as *residue*, *temporal*, or *periodicity pitch*. Spectral pitch derives the fundamental frequency from analysis of the resolved partials, whereas periodicity pitch derives the fundamental frequency from the repetition rate of the waveform produced by the interaction of unresolved partials on the basilar membrane. Singh and Hirsh (1992) have suggested that both pitch mechanisms are always involved in pitch perception, but that "analytic" modes of listening favor spectral pitch and "synthetic" modes of listening favor periodicity pitch (see also Houtsma, 1979).

The emphasis on spectral pitch in the analytic mode provides a strong connection between pitch and timbre. Both spectral pitch and timbre perception are thought to receive input from a preliminary analysis of the spectrum. Although some listeners may be predisposed toward synthetic listening (e.g., musically trained listeners), the interval size task employed in Experiments 1–3 may have encouraged an analytic mode of listening. Conversely, the pitch-matching task employed in Experiment 4 may have emphasized the synthetic mode.

The above explanation implies that the perception of interval size occurs subsequent to pitch perception mechanisms but favors an analytic mode of listening. Another possibility is that interval size is assessed in the early stages of auditory processing when partials are resolved by auditory filters. Several contemporary models of pitch perception propose that individual partials are resolved by auditory filters in the early stages of pitch processing followed by temporal analysis of neural firing patterns (e.g., Meddis & Hewitt, 1991; Moore, 1997; Patterson et al., 1992; see also Terhardt, 1974). It is possible that processes involved in determining the phenomenal size of a melodic interval may not only operate on the output from such pitch mechanisms but also be engaged at the early stages of pitch processing, reflecting a comparison of energy across all of the resolved partials. The feasibility of such comparison is supported by research showing that listeners are highly adept at making comparisons of the output from different auditory filters, particularly when sounds are separated by interstimulus intervals of 0.5 sec or less (Green, 1988). Thus, interval size perception may include computations of spectral analyses arising from early stages of processing as well as pitch analyses arising from later stages. Presumably, training in music encourages greater attention and weighting to computations arising at later stages of processing, thereby reducing potential influences of timbre on pitch-related tasks.

REFERENCES

Boltz, M., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, **18**, 389-431.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, **25**, 257-271.

Deutsch, D. (1975). Musical illusions. *Scientific American*, **233**, 92-104.

Galembo, A., Askenfelt, A., Cuddy, L. L., & Russo, F. A. (2001). Effects of relative phases on pitch and timbre in the piano bass range. *Journal of the Acoustical Society of America*, **110**, 1649-1666.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Green, D. M. (1988). *Profile analysis*. Oxford: Oxford University Press.

Grimault, N., Micheyl, C., Carlyon, R. P., & Collet, L. (2002). Evidence for two pitch encoding mechanisms using a selective auditory training paradigm. *Perception & Psychophysics*, **64**, 189-197.

Houtsma, A. J. M. (1979). Musical pitch of two-tone complexes and predictions by modern pitch theories. *Journal of the Acoustical Society of America*, **66**, 87-99.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, **13**, 313-319.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, **11**, 346-374.

Krumhansl, C. L. (1990). Tonal hierarchies and rare intervals in music cognition. *Music Perception*, **7**, 309-324.

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, **126**, 159-179.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 739-751.

Macromedia (2000). *SoundEdit 16* [Software]. San Francisco: Author.

Meddis, R., & Hewitt, M. J. (1991). Virtual pitch and phase sensitivity of a computer model of the auditory periphery: I. Pitch identification. *Journal of the Acoustical Society of America*, **89**, 2866-2882.

Melara, R. D., & Marks, L. E. (1990a). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics*, **48**, 169-178.

Melara, R. D., & Marks, L. E. (1990b). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 398-414.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Meyer, L. B. (1973). *Explaining music: Essays and explorations*. Berkeley: University of California Press.

Moore, B. C. J. (1997). *An introduction to the psychology of hearing* (4th ed.). New York: Academic Press.

Moore, B. C. J., & Glasberg, B. R. (1990). Frequency discrimination of complex tones with overlapping and non-overlapping harmonics. *Journal of the Acoustical Society of America*, **87**, 2163-2177.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.

Patterson, R. D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., & Allerhand, M. (1992). Complex sounds and auditory images. In Y. Cazals, K. Horner, & L. Demany (Eds.), *Auditory physiology and perception: Proceedings of the 9th International Symposium on Hearing* (pp. 429-446). Oxford: Pergamon.

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 976-986.

Platt, J. R., Racine, R. J., Stark, M., & Weiser, M. (1990). Pitch interactions in the perception of isolated musical triads. *Perception & Psychophysics*, **48**, 59-67.

REAL Software (2003). *REAL Basic 5* [Programming language]. Austin, TX: Author.

Russo, F. A., & Thompson, W. F. (in press). The subjective size of melodic intervals over a two-octave range. *Psychonomic Bulletin & Review*.

Schellenberg, E. G., & Trainor, L. J. (1996). Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: Tests of adult and infant listeners. *Journal of the Acoustical Society of America*, **100**, 3321-3328.

Semal, C., & Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. *Journal of the Acoustical Society of America*, **89**, 2404-2410.

Semal, C., & Demany, L. (1993). Further evidence for an autonomous processing of pitch in auditory short-term memory. *Journal of the Acoustical Society of America*, **94**, 1315-1322.

Shepard, R. N., & Jordan, D. S. (1984). Auditory illusions demonstrating that tones are assimilated to an internalized musical scale. *Science*, **226**, 1333-1334.

Singh, P. G., & Hirsh, I. J. (1992). Influence of spectral locus and F_0 changes on the pitch and timbre of complex tones. *Journal of the Acoustical Society of America*, **92**, 2650-2661.

Smith, J. D., Nelson, D. G. K., Grohskopf, L. A., & Appleton, T. (1994). What child is this? What interval was that? Familiar tunes and music perception in novice listeners. *Cognition*, **52**, 23-54.

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin*, **113**, 345-361.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, **55**, 1061-1069.

Terhardt, E., Stoll, G., & Seewann, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, **71**, 679-688.

Thompson, W. F., Hall, M. D., & Pressing, J. (2001). Illusory conjunctions of pitch and duration in unfamiliar tone sequences. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 128-140.

Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception*, **6**, 383-396.

Warrier, C. M., & Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception & Psychophysics*, **64**, 198-207.

NOTE

1. Two additional control experiments were performed to investigate questions relating to our findings.

In the first control experiment, we used a forced choice matching task to confirm that timbral manipulations did not result in octave confusions. Trials consisted of a standard tone followed by two comparison tones. One comparison tone matched the fundamental frequency of the standard tone, and the other was an octave above or below the fundamental frequency of the standard tone. Nine musically trained participants indicated which comparison tone matched the pitch of the standard tone. All participants were highly accurate, with a mean of 92.01% correct ($SE = 2.86$). Error rates in the predicted direction (i.e., octave-shift up for a bright timbre or octave-shift down for a dull timbre) were not significantly different from error rates that were opposite to predictions [$t(8) = 2.14$, n.s.]. Thus, octave confusions were extremely rare and, when they occurred, were unrelated to timbre.

In an additional control experiment, evidence was obtained that labeling relies on the output of pitch perception mechanisms and does not consider all partials. For this experiment, 9 musically trained participants were asked to label intervals that were identical to those described in Experiment 1 (i.e., perfect fifths and tritones presented with incongruent or congruent timbral shifts). Participants were highly accurate in the labeling task, with a mean of 97.11% ($SE = 1.27\%$). This finding is consistent with the notion that interval labeling depends more on output from later stages of pitch processing than does the phenomenal experience of interval size.

(Manuscript received July 7, 2003;
revision accepted for publication August 2, 2004.)