

# Journal of Experimental Psychology: Human Perception and Performance

## **Vibrotactile Discrimination of Musical Timbre**

Frank A. Russo, Paolo Ammirante, and Deborah I. Fels

Online First Publication, June 18, 2012. doi: 10.1037/a0029046

### CITATION

Russo, F. A., Ammirante, P., & Fels, D. I. (2012, June 18). Vibrotactile Discrimination of Musical Timbre. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. doi: 10.1037/a0029046

## OBSERVATION

# Vibrotactile Discrimination of Musical Timbre

Frank A. Russo, Paolo Ammirante, and Deborah I. Fels  
Ryerson University

Five experiments investigated the ability to discriminate between musical timbres based on vibrotactile stimulation alone. Participants made same/different judgments on pairs of complex waveforms presented sequentially to the back through voice coils embedded in a conforming chair. Discrimination between cello, piano, and trombone tones matched for F0, duration, and magnitude was above chance with white noise masking the sound output of the voice coils (Experiment 1), with additional masking to control for bone-conducted sound (Experiment 2), and among a group of deaf individuals (Experiment 4a). Hearing (Experiment 3) and deaf individuals (Experiment 4b) also successfully discriminated between dull and bright timbres varying only with regard to spectral centroid. We propose that, as with auditory discrimination of musical timbre, vibrotactile discrimination may involve the cortical integration of filtered output from frequency-tuned mechanoreceptors functioning as critical bands.

*Keywords:* vibrotactile, music, timbre, deaf

A topic of recent investigation in music perception research has been the contribution of extraauditory cues. This interest has arisen in part from a growing interest in the evolutionary origins of music, and a concomitant recognition of the auditory-dominated musical experience afforded by the advent of recorded music as a historical anomaly (Thompson, Graham, & Russo, 2005). Research on the integration of auditory and visual cues has shown that watching a musician perform can influence perception of structural aspects of music (Schutz & Lipscomb, 2007; Thompson & Russo, 2007), as well as emotional aspects (Thompson et al., 2005; Thompson, Russo, & Quinto, 2008; Vines, Krumhansl, Wanderley, & Levitin, 2006). For example, a singer's facial expressions can strongly influence judgments of emotion conveyed by song. These findings suggest that visual cues may indirectly influence auditory perception by conveying information about those intentional acts that give rise to musical sound (Russo, Sandstrom, & Maksimowski, 2011).

While there has been considerable research on the contribution of visual information to the understanding and appreciation of

music, investigation of the contribution from the tactile system has remained relatively obscure. Vibration in the form of air pressure against the skin and within the internal cavities such as the lungs, and physical contact with speakers or instruments is also experienced in live music settings and could contribute to an overall gestalt of live music. Unlike visual cues, tactile information obtained from touching a stereo speaker or the soundboard of a cello consists of the same physical energy used by the auditory system. Furthermore, tactile receptors in the skin and auditory receptors in the cochlea are structurally similar; both are mechanoreceptors whose bending in response to pressure changes triggers neural impulses sent to the brain (Orr, Helmke, Blackman, & Schwartz, 2006). Thus, low-level processes involved in the sensation of vibration may also be similar between modalities. Indeed, it has been shown that tactile and auditory stimuli can be confused when presented simultaneously or in alternation (Bekésy, 1959; Gescheider & Niblette, 1967) and that vibrotactile input can stimulate the auditory cortex (Caetano & Jousmaki, 2006).

The sensation of musical tone by the skin or the ear is generated by periodic vibration. Psychophysical studies have investigated vibrotactile (VT) discrimination of tones in isolation using auditory masking. These studies have focused largely on VT discrimination of the fundamental frequency (F0) and intensity of tones. On the one hand, these findings show clear differences between VT and auditory (A) discrimination, for example, upper F0 discrimination threshold is ~1,000 Hz (Verrillo, 1992) for VT and 20,000 Hz for A, and just-noticeable difference between the F0 of tones has been reported to be between 10 to 30 times higher for VT than A (Franzén and Nordmark, 1975; Goff, 1967).

On the other hand, some striking similarities emerge between VT and A related to interactions between frequency and intensity (Verrillo, 1992). For example, for both VT and A, perceived "loudness" of pairs of pure tones presented either successively (Verrillo & Gescheider, 1975) or simultaneously (Marks, 1979) is

---

Frank A. Russo and Paolo Ammirante, Department of Psychology, Ryerson University, Toronto, ON, Canada; Deborah I. Fels, Centre for Learning Technologies, Ryerson University.

This research was supported by a Discovery grant awarded to the first author from the Natural Sciences and Engineering Research Council of Canada, and Graphics, Animation and New Media (GRAND) Canada, a federally funded Network of Centres of Excellence. We acknowledge Julia Kim, Christopher Lachine, Michael Maksimowski, Lisa Liskovoi, and Gabriel Nespoli for research assistance. We are indebted to the Canadian Cultural Society of the Deaf for assistance with participant recruitment, and to Kelly Ferguson for American Sign Language interpretation.

Correspondence concerning this article should be addressed to Frank A. Russo, 350 Victoria Street, Toronto, Canada, M5B 2K3. E-mail: russo@ryerson.ca

summed only when the frequencies of the tones are widely spaced. In the auditory domain, loudness summation is thought to be related to frequency-tuned hair cell populations arranged tonotopically along the length of the cochlea (Zwislocki & Sokolich, 1974). Widely spaced tone pairs displace nonadjacent populations, effectively filtering neural impulses sent to the brain into separate channels, and leading to a cortical summation of their intensities. Tone pairs that fall within the same *critical band* ( $\sim 1/3$  of an octave) will displace adjacent populations of hair cells and are filtered into the same channel, partially masking summation.

Verrillo and Gescheider (1975) suggested that the mechanism underlying VT loudness summation may be analogous to critical bands in the auditory system. They pointed to differences in frequency tunings between at least two different types of skin receptors: Pacinian corpuscles lying deep within the dermal layer with peak sensitivity to vibration between 225 and 275 Hz, and non-Pacinian corpuscles closer to the epidermis with peak sensitivity between 25 and 40 Hz. Subsequent studies have reported findings consistent with a multiple channel model (Makous, Friedman, & Vierck, 1995; Marks, 1979; Gescheider, Bolanowski, & Hardick, 2001), and four types of skin receptors are generally recognized today (Birbaum & Wanderley, 2007; Bolanowski, Gescheider, Verrillo, & Checkosky, 1988).

One question that arises from these findings is whether frequency-tuned skin receptors can filter the component frequencies or *partials* that fall above the F0 of spectrally rich musical tones. In the auditory system, one means by which the timbres of different musical instruments of identical F0 are distinguished is through the relative intensities of a complex tone's partials. For example, greater spectral energy in the upper partials contributes to the characteristically bright timbre of a clarinet. Here we report 5 experiments that investigated the ability to discriminate musical timbre on the basis of vibrotactile information alone.

## Experiment 1

### Method

**Participants.** Ten undergraduates (7 women) aged 19–31 ( $M = 24.6$ ;  $SD = 3.6$ ) from Ryerson University participated for course credit.

**Apparatus.** Complex vibrotactile waveforms driven by an acoustic signal were presented to the human back via a pair of voice coils embedded in a conforming chair (Karam, Russo, & Fels, 2009). The voice coils were 1-in in diameter and made contact with the left and right sides of the lumbar region of the back. To mask any incidental sounds produced by the voice coils, continuous white noise was played over headphones (75 dBA).

**Stimuli and procedure.** Stimuli were piano, cello, and trombone samples taken from the Iowa Musical Instrument Samples (Fritz, 1997). Tones were approximately 1.5 s in length and presented sequentially with an interstimulus interval (ISI) of 1 s at each of 3 levels of F0: 120, 220, and 440 Hz. Figure 1 shows spectrograms for each instrument at 220 Hz. All stimuli were equated for perceived magnitude of vibration on the basis of iterative amplitude adjustments made by one of the authors (F.A.R.). Two judges who did not serve as participants validated that the amplitude-adjusted tones were equal in magnitude.

Participants made same/different judgments for stimulus pairs. A practice block of 5 trials with feedback was followed by 54 trials without feedback presented in 6 blocks. Each block presented pairs at a single F0, and each of these blocks was repeated. The order of presentation of the blocks was randomized as was the order of trials within each block. Within blocks, all possible timbre pairs were presented once (3 timbres squared = 9 trials). Thus, one-third of the pairs were same and two-thirds were different.

### Results

Median percentage of correct responses was 85.08%.<sup>1</sup> A Friedman's analysis of variance (ANOVA) on all available trials showed no effect of F0,  $\chi^2(2) = 4.431$ ,  $p = .109$ . Each participant's percent correct scores for complementary different pairs at each F0 (e.g., cello/trombone vs. trombone/cello at 110 Hz) were entered into Wilcoxon's paired sample tests, which showed no effect of order of presentation. Table 1 shows percent correct for each stimulus pair after collapsing across participants, F0, and order of presentation for different pairs. Binomial tests on these collapsed data revealed that discrimination was significantly above chance for all stimulus pairs ( $p < .05$ ).

### Discussion

These findings supported our prediction that participants should be able to discriminate the timbre of musical instruments based on vibrotactile information alone. Experiment 2 controlled for the possibility that some residual auditory information remained available through bone conduction.

## Experiment 2

### Method

**Participants.** Five undergraduates (3 women) aged 18–24 ( $M = 21.8$ ;  $SD = 1.9$ ) from Ryerson University participated for course credit.

**Apparatus.** The apparatus was the same as Experiment 1 except that potential bone conduction of the acoustic signal was masked using white noise presented by vibrotactile stimulation over Tactaid VBW32 transducers placed on the left and right mastoids and by auditory stimulation over headphones.

**Stimuli and procedure.** Stimuli and procedure were identical to Experiment 1.

### Results

Median percent correct was 90.74%. As in Experiment 1, F0 did not affect the percentage of correct responses,  $\chi^2(2) = 1.226$ ,  $p = .542$ , and there were no effects of the order of presentation of different pairs at any of the 3 F0s. As shown in Table 2, percent correct was significantly above chance for all stimulus pairs col-

<sup>1</sup> The median  $d'$  values suggest unbiased responding: Experiment 1 (Mdn = 2.55); Experiment 2 (Mdn = 2.60); Experiment 3 (Mdn = 3.48); Experiment 4a (Mdn = 4.06) and Experiment 4b (Mdn = 3.09).

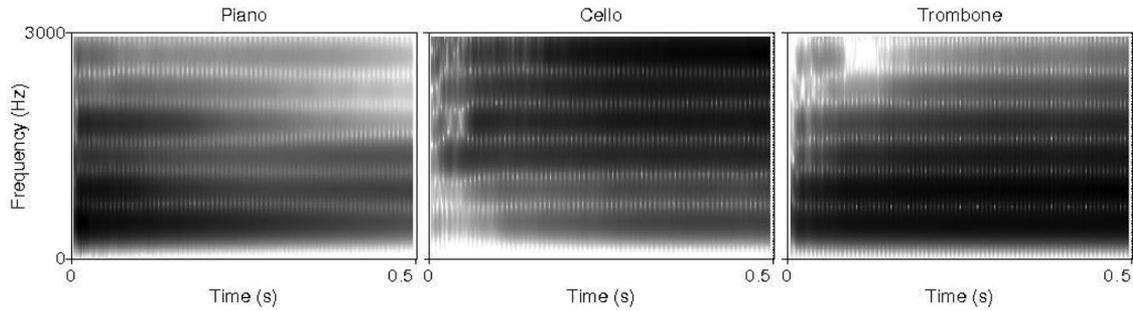


Figure 1. Spectrograms of piano, cello, and trombone tones with F0 of 220 Hz. Shown are the first 500 ms after tone onset and for frequencies below 3,000 Hz.

lapsed across participants, F0, and order of presentation for different pairs ( $p < .05$ ).

### Discussion

Experiment 2 revealed that the findings for Experiment 1 could not be attributed to bone-conducted auditory information instead of vibrotactile discrimination. These data support the hypothesis that different populations of frequency-tuned skin receptors function as critical bands, filtering complex tones into their component frequencies, with relative intensities of these component frequencies contributing to a tactile perception of timbre.

However, spectral content is not the only means by which musical timbre is distinguished. An alternative possibility is that vibrotactile discrimination in Experiments 1 and 2 was based on transient differences in amplitude envelope between instruments. For example, a piano tone has a relatively “sharp” attack (i.e., intensity rises steeply at onset) followed by a rapid decay, whereas a trombone tone has a more gradual attack and decay. Thus, instead of instrument samples, Experiment 3 used synthesized tones identical in amplitude envelope but varied in normalized *spectral centroid*. This is defined as the amplitude-weighted mean of the tone’s frequency spectrum divided by its F0, and its psychological correlate is brightness (Warrier & Zatorre, 2002). Thus, with F0 held constant, a tone containing more spectral energy (i.e., intensity) in the lower partials has a lower spectral centroid and sounds dull, whereas a tone containing more spectral energy in the upper partials has a higher spectral centroid and sounds bright.

## Experiment 3

### Method

**Participants.** Five undergraduates (2 women) aged 21–28 ( $M = 23.8$ ;  $SD = 2.77$ ) from Ryerson University participated for course credit.

Table 1  
Percent Correct for Experiment 1

	Cello	Piano	Trombone
Cello	95		
Piano	92	90	
Trombone	61.5	93	93

**Apparatus.** The apparatus was the same as Experiment 2 (i.e., both auditory and vibrotactile masking were used).

**Stimuli and procedure.** Dull and bright timbres contained 11 harmonically related partials that differed in intensity. The normalized spectral centroid of the dull tone was approximately a perfect fifth lower than the bright tone (Russo & Thompson, 2005). Both tones had an F0 of 220 Hz, were 1 s in duration, and contained identical attack and decay patterns. As in Experiment 1, all stimuli were equated for perceived magnitude of vibration on the basis of iterative amplitude adjustments made by one of the authors (F.A.R.). Two judges who did not serve as participants validated that the amplitude-adjusted tones were equal in magnitude.

Participants again made same/different judgments on stimulus pairs. Pairs were presented sequentially with a 1 s ISI. Following a practice block of 5 trials with feedback, 16 trials were completed without feedback in two blocks. Within each block, all possible timbre pairs were randomly presented twice (8 trials), and the probability of a “same” trial was .5.

### Results

Median percent correct was 93.75%, and did not differ between orders of presentation for different pairs. As shown in Table 3, percent correct collapsed across participants and order of presentation for different pairs was again significantly above chance for all stimulus pairs ( $p < .05$ ).

### Discussion

Experiment 3 showed that vibrotactile discrimination of timbre persists with tone F0, duration, and amplitude envelope held constant. This finding lends stronger support to the critical band hypothesis.

In Experiments 4a and 4b, we tested vibrotactile discrimination of frequency-matched tones in a sample of deaf individuals. Be-

Table 2  
Percent Correct for Experiment 2

	Cello	Piano	Trombone
Cello	73		
Piano	81	73	
Trombone	90	90	87

Table 3  
Percent Correct for Experiment 3

	Dull	Bright
Dull	90	
Bright	95	90

sides providing a rigorous means of controlling for the residual availability of auditory cues, these replications were motivated by an interest in the application of VT stimuli to assistive music technology (Karam et al., 2009).

### Experiments 4a and 4b

#### Method

**Participants.** 19 individuals (9 women) aged 23–64 ( $M = 42.1$ ;  $SD = 12.9$ ) were recruited from Toronto’s deaf community. Participants were compensated \$20.

**Apparatus.** The apparatus was the same as Experiment 1. Eleven of our deaf participants reported some hearing at high intensity and five of these participants wore hearing aids. To eliminate any possibility of auditory stimulation, all participants were asked to wear sound attenuating earmuffs with a noise reduction rating of 26 dB, and those with hearing aids were asked to turn their devices off for the duration of the experiment.

**Stimuli and procedure.** The stimuli and procedure in Experiment 4a were identical to Experiment 1 (musical instruments) and the stimuli and procedure in Experiment 4b were identical to Experiment 3 (dull/bright tones). All instructions to participants were delivered using American Sign Language. Participants completed both experiments in a single session.

#### Results

In Experiment 4a, median percent correct was 92.59%. Friedman’s ANOVA again revealed no effect of F0,  $\chi^2(2) = 4.711$ ,  $p < .10$ . Wilcoxon’s paired sample tests at each F0 showed no significant differences in the order of presentation for different pairs. As shown in Table 4, percent correct collapsed across F0, participants, and order of presentation for different pairs was significantly above chance for all tone pairs ( $p < .05$ ). Figure 2 shows boxplots of percent correct for the musical instrument stimuli among deaf participants in Experiment 4a and hearing participants in Experiments 1 and 2.

In Experiment 4b, median percent correct was 90.63%. There were no order effects for different pairs. As shown in Table 5, percent correct collapsed across participants and order of presentation for different trials was significantly above chance ( $p < .05$ )

Table 4  
Percent Correct for Experiment 4a

	Cello	Piano	Trombone
Cello	90		
Piano	96.5	97	
Trombone	71.5	96	96.5

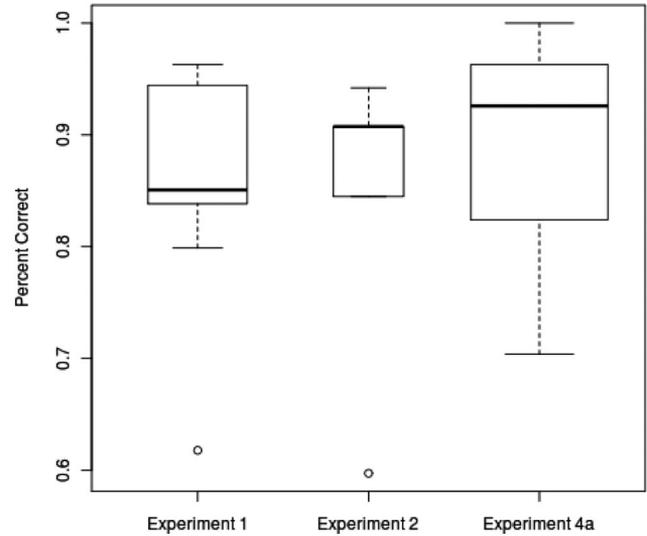


Figure 2. Boxplots of percent correct for the musical instrument stimuli in Experiment 1 (white noise masking), Experiment 2 (additional masking), and Experiment 4a (deaf participants). Outliers are shown as small circles. Box widths are proportional to the square root of the sample size in each experiment.

for all tone pairs. Figure 3 shows boxplots of percent correct for the dull/bright stimuli among deaf participants in Experiment 4b and hearing participants in Experiment 3.

### General Discussion

Five experiments investigated the ability to discriminate between musical timbres based on vibrotactile stimulation alone. Participants made same/different judgments on pairs of complex waveforms presented to the lumbar region of the back through voice coils embedded in a conforming chair. Discrimination between cello, piano, and trombone tones matched for F0, duration, and intensity was well above chance with white noise masking any of the incidental sound emanating from the voice coils (Experiment 1), with additional masking to control for bone-conducted sound (Experiment 2), and among a group of deaf individuals (Experiment 4a). Hearing (Experiment 3) and deaf individuals (Experiment 4b) also successfully discriminated between dull and bright timbres varying only with regard to spectral centroid.

Given that VT and A sensations of tone are derived from the same physical energy, as well as mechanical similarities between VT and A receptor systems (Orr et al., 2006), the current findings may not be surprising. Indeed, they are consistent with previous findings using pairs of pure tones presented successively (Verrillo & Gescheider, 1975) or simultaneously (Marks, 1979) in suggest-

Table 5  
Percent Correct for Experiment 4b

	Dull	Bright
Dull	92	
Bright	83	75

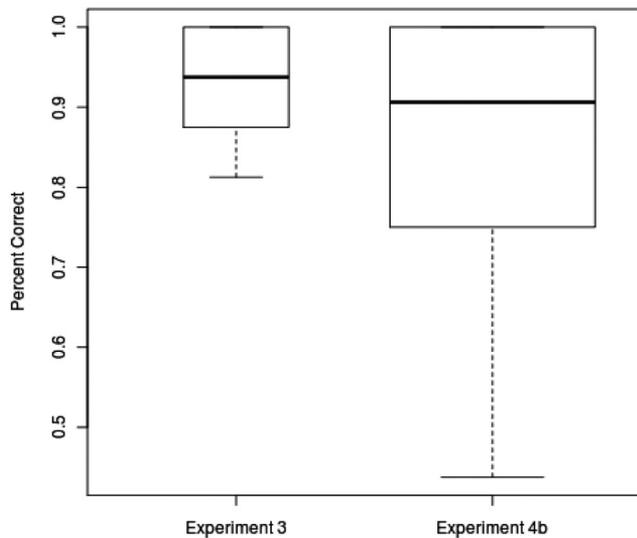


Figure 3. Boxplots of percent correct for the dull/bright stimuli in Experiment 3 (additional masking) and Experiment 4b (deaf participants).

ing that sensation of tone initiates similar perceptual filtering processes in both modalities (Gescheider et al., 2001; Makous et al., 1995; Marks, 1979; Verrillo & Gescheider, 1975). In particular, these findings imply that populations of frequency-tuned skin cells, like cochlear hair cells, act as critical bands, filtering a complex tone into its component frequencies, with vibrotactile discrimination of timbre following from sensitivity to the relative intensities across the bands. Previous investigations of VT discrimination of musical tone have focused on F0, intensity, and timing (see Verrillo, 1992, for a review); here we extend these findings to include timbre. The current findings may be relevant to research on the role of extraauditory feedback in musical performance (Marshall & Wanderley, 2006; Verrillo, 1992). They may also be applied creatively to multimodal composition and expression for artistic or inclusive purposes. Finally, they may be applied to telephony for the hard of hearing, where perception of talker-specific features such as identity, gender, and age, might be supported through a supplementary vibrotactile channel.

## References

- Bekésy, G. von. (1959). Similarities between hearing and skin senses. *Psychological Review*, 66, 1–22. doi:10.1037/h0046967
- Birnbaum, D. M., & Wanderley, M. M. (2007). A systematic approach to musical vibrotactile feedback. In *Proceedings of the International Computer Music Conference (ICMC)* (Vol. 2, pp. 397–404). Copenhagen: ICMA.
- Bolanowski, S. J., Jr., Gescheider, G. A., Verrillo, R. T., & Checkosky, C. M. (1988). Four channels mediate the mechanical aspects of touch. *Journal of the Acoustical Society of America*, 84, 1680–1694. doi:10.1121/1.397184
- Caetano, G., & Jousmaki, V. (2006). Evidence of vibrotactile input to human auditory cortex. *Neuroimage*, 29, 15–28. doi:10.1016/j.neuroimage.2005.07.023
- Franzén, O., & Nordmark, J. (1975). Vibrotactile frequency discrimination. *Perception & Psychophysics*, 48, 480–484. doi:10.3758/BF03203298
- Fritz, L. (1997). *University of Iowa musical instrument samples*. Retrieved from <http://theremin.music.uiowa.edu>
- Gescheider, G., Bolanowski, S. J., & Hardick, K. R. (2001). The frequency selectivity of information-processing channels in the tactile sensory system. *Somatosensory & Motor Research*, 18, 191–201. doi:10.1080/01421590120072187
- Gescheider, G. A., & Niblette, R. K. (1967). Cross-modality masking for touch and hearing. *Journal of Experimental Psychology*, 74, 313–320. doi:10.1037/h0024700
- Goff, G. D. (1967). Differential discrimination of frequency of cutaneous mechanical vibration. *Journal of Experimental Psychology*, 74, 294–299.
- Karam, M., Russo, F. A., & Fels, D. (2009). Designing the model human cochlea: An ambient crossmodal audio tactile display. *IEEE Transactions on Haptics*, 2, 160–169.
- Makous, J. C., Friedman, R. M., & Vierck, C. J. (1995). A critical band filter in touch. *Journal of Neuroscience*, 15, 2808–2818.
- Marks, L. E. (1979). Summation of vibrotactile intensity: An analog to auditory critical bands? *Sensory Processes*, 3, 188–203.
- Marshall, M. T., & Wanderley, M. M. (2006). Vibrotactile feedback in digital musical instruments. In *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression* (pp. 226–229). Paris: IRCAM.
- Orr, A. W., Helmke, B. P., Blackman, B. R., & Schwartz, M. A. (2006). Mechanisms of mechanotransduction. *Developmental Cell*, 10, 11–20. doi:10.1016/j.devcel.2005.12.006
- Russo, F. A., Sandstrom, G. M., & Maksimowski, M. (2011). Mouth versus eyes: Gaze fixation during perception of sung interval size. *Psychomusicology: Music, Mind & Brain*, 21, 98–107.
- 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, 559–568. doi:10.3758/BF03193514
- Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, 36, 888–897. doi:10.1068/p5635
- Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, 156, 203–227. doi:10.1515/semi.2005.2005.156.203
- Thompson, W. F., & Russo, F. A. (2007). Facing the music. *Psychological Science*, 18, 756–757. doi:10.1111/j.1467-9280.2007.01973.x
- Thompson, W. F., Russo, F. A., & Quinto, L. (2008). Audio-visual integration of emotional cues in song. *Cognition and Emotion*, 22, 1457–1470. doi:10.1080/02699930701813974
- Verrillo, R. (1992). Vibration sensation in humans. *Music Perception*, 9, 281–302.
- Verrillo, R., & Gescheider, G. A. (1975). Enhancement and summation in the perception of two successive vibrotactile stimuli. *Perception & Psychophysics*, 18, 128–136. doi:10.3758/BF03204100
- Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross modal interactions in the perception of musical performance. *Cognition*, 101, 80–113. doi:10.1016/j.cognition.2005.09.003
- Warrier, C. M., & Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception & Psychophysics*, 64, 383–396. doi:10.3758/BF03195786
- Zwislocki, J. J., & Sokolich, W. G. (1974). On loudness enhancement of a tone burst by a preceding tone burst. *Perception & Psychophysics*, 16, 87–90. doi:10.3758/BF03203256

Received December 16, 2011

Revision received April 23, 2012

Accepted April 25, 2012 ■