Music Notation, But Not Action on a Keyboard, Influences Pianists' Judgments of Ambiguous Melodies

Bruno H. Repp
Haskins Laboratories

Robert M. Goehrke
Yale University

PITCH INCREASES FROM LEFT TO RIGHT ON PIANO keyboards. When pianists press keys on a keyboard to hear two successive octave-ambiguous tones spanning a tritone (half-octave interval), they tend to report hearing the tritone go in the direction consistent with their key presses (Repp & Knoblich, 2009). This finding has been interpreted as an effect of action on perceptual judgment. Using a modified design, the present study separated the effect of the action itself from that of the visual stimuli that prompt the action. Twelve expert pianists reported their perception of octave-ambiguous three-note melodies ending with tritones in two conditions: In the active condition, they saw a notated melody and played it on a keyboard to hear it, while in the passive condition they viewed the notation while the melody was played to them. Participants tended to report hearing the tritone as it appeared in the notation, but action had no additional effect. We discuss whether the "action direction effect" described by Repp and Knoblich may have been caused by the visual action prompts, not by the action itself.

Received April 16, 2010, accepted August 9, 2010.

Key words: tritone, tritone paradox, action, perception, ambiguous melodies

A RAPIDLY INCREASING VOLUME OF RESEARCH documents tight relationships between perception and action. While the dependence of action on perception has long been recognized, many recent studies demonstrate previously unsuspected effects of action on perception (e.g., Hamilton, Wolpert, & Frith, 2004; Miall et al., 2006; Schubö, Aschersleben, & Prinz, 2001; Wohlschläger, 2000; Wühr & Müsseler, 2001; Zwickel, Grosjean, & Prinz, 2007, 2010; for reviews, see Hommel, Müsseler, Aschersleben, & Prinz, 2001; Proffitt, 2006; Schütz-Bosbach & Prinz, 2007; Wilson & Knoblich, 2005). Almost all of these studies have been concerned with visual perception. However, Repp and Knoblich (2007, 2009) recently claimed to have found effects of action on auditory perception.

Their research was based on the fact that on a piano keyboard, the pitch of produced tones increases from left to right. Pianists are continuously exposed to this relationship, whereas other musicians and some nonmusicians merely know it but do not have sensorimotor experience with it. On each trial of their experiment, Repp and Knoblich (2009) showed participants a pair of two-digit numbers that corresponded to two keys on a labeled piano keyboard (a silent MIDI controller). The keys always spanned the musical interval of a tritone (half an octave), with the second key being either higher (i.e., to the right of) or lower than (to the left of) the first key. When participants pressed the keys in succession according to the numerical action prompts, they heard two octave-ambiguous tones (Shepard, 1964) spanning a tritone (Deutsch, 1986, 1987) and had to report whether the interval was rising or falling. They perceived the tritone as matching the direction of their action more often than not. This "action direction effect" (ADE) was much stronger for pianists than for non-pianists, and pianists also showed it even when they sat still and observed someone else pressing the keys.

One interpretation of the ADE is that performed or observed keyboard actions automatically evoke strong auditory associations in pianists, and that these images interact with the perception of ambiguous auditory stimuli. Indeed, there is evidence from behavioral and neuroscience studies that auditory imagery occurs when pianists press silent piano keys (Bangert, Jürgens, Häusler, & Altenmüller, 2006) and that, conversely, perception of piano tones can lead to activation of the motor system in pianists (Drost, Rieger, Brass, Gunter, & Prinz, 2005a, 2005b; Drost, Rieger, & Prinz, 2007; Haueisen & Knösche, 2001; Lahav, Saltzman, & Schlaug, 2007). Other interpretations of the ADE are possible, however, as Repp and Knoblich (2009) acknowledge. The effect of action may have been post-perceptual, affecting just the response decision about an ambiguous auditory stimulus, rather

Music Perception VOLUME 28, ISSUE 3, PP. 315–320, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 DOI:10.1525/MP.2011.28.3.315

than pushing the percept in one or another direction. It is also possible that the ADE was not due to the action itself but to the visual stimuli that prompted the action. Although Repp and Knoblich varied the direction of the number labels on the keyboard and found no effect of ascending versus descending numbers, the numbers that served as action prompts had to be translated into an action plan by the participants. The realization that a left-right or right-left movement was to be carried out may have been sufficient to bias responses to the auditory stimuli, with the action itself being redundant.

The purpose of the present experiment was to address this last possibility by dissociating the action from the visual prompt that gave rise to it. This aim could be realized more effectively by using music notation instead of numbers, because notation retains its action relevance in the absence of a keyboard. (There is no reason to expect a pair of numbers to affect auditory perception unless it is associated with an action plan.) In the experiment, pianists either played a notated melody on a keyboard and reported how the resulting sequence of octave-ambiguous tones sounded to them, or they just viewed the notation while listening passively to the tones and then reported their impression. Whether the notation would influence their judgments was an interesting question in itself. The primary question, however, was whether carrying out the action would have an effect on perceptual judgment above and beyond any effect that the notation (the action prompt) might have. If it did, there should be a larger effect on tritone perception in the active condition (prompt + action) than in the passive condition (prompt alone).

Our experiment also aimed to make the experimental paradigm less transparent. In the Repp and Knoblich (2009) study, it must have been rather obvious to the participants what the research was about, and they may have either given in to or resisted a bias induced by action and/or action prompts. Indeed, there were large individual differences in the size of the ADE, which are difficult to explain on the basis of sensorimotor experience alone, as all pianists were highly trained. In the present study we preceded each tritone pair with a third tone, thus forming a three-tone melody, which was intended to divert participants' attention from the tritone intervals as such. It is known from previous research that a preceding context tone influences tritone perception according to a principle of pitch proximity that minimizes the range of the pitches perceived in successive octave-ambiguous tones (Repp, 1997; Shepard, 1983). Thus, for example, if the tritone C-F# is preceded by D#, it tends to be heard as rising, whereas when it is preceded by A, it tends to be heard as falling. These percepts keep the total pitch range at six semitones, whereas it would be nine semitones otherwise. Replication of this context effect was also of interest, as the earlier data were rather limited.

Furthermore, we used two sets of octave-ambiguous tones with different spectral envelopes, which, according to some previous findings (e.g., Deutsch, 1987) should yield similar results, but according to others might show different response patterns (Repp, 1994, 1997) or differences in degree of ambiguity (Repp & Thompson, 2010). Finally, we also varied the playing hand (left or right) in the active condition, though we did not expect it to make a difference.

Method

PARTICIPANTS

Twelve pianists were paid to participate. They included three graduate students from the Yale School of Music and nine undergraduate students from Yale University. They all played advanced repertoire, had 10-23 years of training, and currently took lessons with Yale faculty pianists.

MATERIALS

Octave-ambiguous tones were created using a program written in MAX/MSP (version 4.0.9). Each tone consisted of six octave-spaced partials whose relative amplitudes were governed by a fixed convex spectral envelope function centered on either 262 Hz (C4) or 370 Hz (F#4) (see Deutsch, Kuyper, & Fisher, 1987; Repp & Thompson, 2010: Figure 1). For each envelope there were 12 tones representing the 12 musical pitch classes (A, A#,…, G#). From each set of 12 tones, 12 pairs were formed whose members were separated by the interval of a tritone (A-D#, A#-E, …, G#-D). Then each of those pairs was preceded with a third tone that was three semitones away from the initial tone, either higher or lower (e.g., C-A-D# and F#-A-D#). This resulted in 24 three-tone melodies for each envelope set.

Two notations were created for each of these melodies, one in which the tritone was ascending and another in which it was descending. As an example, the alternative notations for the two melodies ending with the B-F tritone are shown in Figure 1. Enharmonic equivalents (flats and sharps) were used such that the first and second notes were always a minor third apart (e.g., Ab-Cb instead of Ab-B, which would be considered an augmented second in music theory; note that B = Cb on the piano).

DESIGN AND PROCEDURE

The experiment was divided into 8 blocks, each containing 48 randomly ordered trials resulting from the combination of 24 melodies (12 tritones preceded by one of two initial

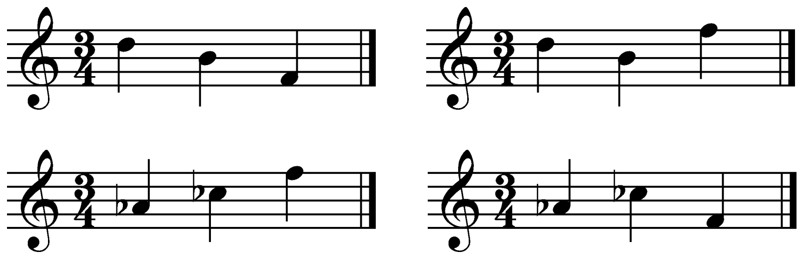

FIGURE 1. Four notations of melodies containing the B-F tritone, varying with regard to initial tone and rising/falling tritone.

tones) and two notations (tritone rising or falling). All tones in a block had the same spectral envelope. The 8 blocks resulted from the crossing of three independent variables: condition (active or passive), envelope (centered on C4 or F#4), and playing hand (left or right) in the active condition; in the passive condition, playing hand was a dummy variable, and participants just listened to two blocks of the same stimuli. The order of the blocks was counterbalanced across participants, such that condition varied most slowly (first versus second half of session), envelope varied within condition, and playing hand varied within envelope in the active condition. The six participants who started with the passive condition did not know they would have to play the melodies later.

Participants came for a single session lasting about 75 minutes. They were seated in front of the computer and listened over Sennheiser HD 540 reference II headphones at a comfortable intensity. In each trial, one of the 48 notations first appeared on the computer screen. In the passive condition, participants heard the corresponding tones after a delay of 2 s. Each tone lasted 500 ms, with interonset intervals of 750 ms. Participants held on to the mouse with their right hand. In the active condition, participants played the displayed melody on a three-octave Fatar Studio 37 MIDI controller that was placed in front of them. They were instructed to use appropriate fingering and to play non-legato at a tempo of approximately 80 quarter notes per minute (= interonset intervals of 750 ms), as demonstrated by the experimenter. Each correct key depression made the computer play an octave-ambiguous tone of the corresponding pitch class with a fixed duration of 500 ms. Depression of an incorrect key did not result in a tone. Participants could play the melody repeatedly if they wished.

After hearing each octave-ambiguous melody, participants judged whether it matched the notation on the screen. If they thought so, they clicked a "Yes" box on the screen using the mouse and went on to the next trial. If not, they clicked "No" and selected the pitch contour that matched the heard melody from a window that popped up (see Figure 2). One of the six response choices (either

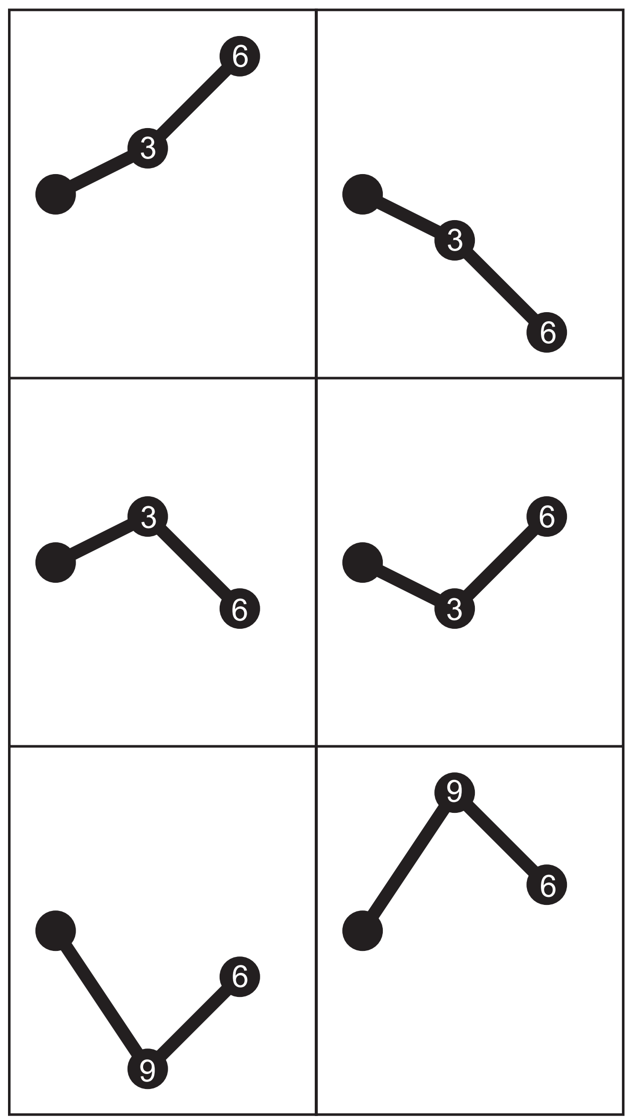

FIGURE 2. Possible contours that a participant could select. The dots represent the three notes, while the numbers inside the dots represent the preceding interval in semitones. (This was explained to the participants.)

contour 3 or 4) was the one that had just been rejected and thus was unlikely to be selected. The other contour spanning a range of six semitones (either contour 3 or 4) was considered a likely response because it follows the pitch proximity principle, whereas other contours (spanning a range of nine semitones) were considered less likely choices. A preference for contours 3 and 4 over contours 1 and 2 represents the expected context effect of the initial tone on tritone perception (Repp, 1997). Responses 5 and 6 were considered least likely because they reflect perception of the initial interval as a major sixth instead of a minor third, which violates the pitch proximity principle for pairs of octave-ambiguous tones (Shepard, 1964). Two theoretically possible contours, each representing a major sixth followed by a tritone in the same direction, were not included in the response panel because they span 15 semitones, which is perceptually impossible (or at least extremely unlikely) given the octave circularity of Shepard tones.

ANALYSIS

Each response was coded according to whether the participant had heard the tritone as rising or falling. The percentage of trials in which the tritone was heard as falling ("percent falling") was calculated in each cell of the design for each participant. (The complementary "percent rising" could have been chosen just as well.) In an initial ANOVA on the active condition only, there were no significant effects involving playing hand, so the data were collapsed over this variable. A 5-way mixed-model ANOVA was then conducted, with the within-participants variables of condition (active or passive), notation (rising or falling tritone), first tone (lower or higher than the second tone according to pitch class proximity), and envelope (C4 or F#4), and the between-participants variable of order of conditions (active first or passive first).

Results

The ANOVA yielded three significant main effects and no significant interactions. The most reliable effect was that of the first tone, $F(1, 10) = 31.81, p < .001$. Consistent with the pitch proximity principle, participants heard the tritone more often as falling when the preceding tone was a minor third below the initial tone of the tritone (68.0% "falling" responses) than when it was a minor third above (44.0%). There was also an effect of notation, $F(1, 10) = 5.64$, $p = .039$. The tritone was more often reported as falling when the notation showed a falling tritone (63.0%) than when it showed a rising tritone (49.0%). Finally, there was an unexpected effect of envelope, $F(1, 10) = 5.68$, $p = .038$, because participants identified the tritone more often as falling when the tones had the C4 envelope (57.4%) than when they had the F#4 envelope (54.6%), a small difference for which we have no explanation. If action had had an effect on tritone perception, the effect of notation should have been more pronounced in the active than in the passive condition, leading to a significant Condition × Notation interaction. However, there was no sign of such an interaction, $F(1, 10) = 0.33$, $p = .896$. No other interaction involving condition came even close to significance. Moreover, order of conditions did not interact significantly with any of the within-participants variables.

It is well known from earlier research on the "tritone paradox" (e.g., Deutsch, 1986, 1987; Deutsch et al., 1987; Deutsch, North, & Ray, 1990) that the identification of a tritone as falling or rising depends strongly on the pitch classes that define the interval, that this pattern is subject to considerable individual differences, and that it may also vary with envelope to a greater (Repp, 1994, 1997) or lesser (Deutsch, 1987) extent. We will not dwell on these pitch class effects here; suffice it to say that their average pattern resembled that found in a recent study using the same tones (Repp & Thompson, 2010), where a more detailed description can be found. One effect peculiar to the present study, however, was the occasional perception of the initial three-semitone interval of the melodies as a major sixth. These responses, which were more frequent than expected (11% of all responses), occurred almost exclusively with certain pitch class combinations that could be predicted from the overall pattern of responses to the tritones: If the pitch class circle was bisected so as to maximize the percentage of "falling" responses in one (the "high") half, occasional perception of the initial interval of the three-tone melody as a major sixth occurred precisely when an initial tone from the low half was followed by a tone from the high half whose pitch class was three semitones lower (e.g., "low" C followed by "high" A, yielding a rising sixth), or the reverse. Thus, the pitch class effect discovered by Deutsch and colleagues extends to intervals other than the tritone, though with diminished strength because these intervals are less ambiguous.

Discussion

The two main findings of this study are that (1) music notation influences perceptual judgment of octave-ambiguous melodies, and (2) playing the melodies on a keyboard according to the notation does not have an additional effect on perception. In other words, there was an effect of the action prompts but not of the actions themselves. This immediately raises the question of whether the "action direction effect" (ADE) reported by Repp and Knoblich (2009) could also have been due to the action prompts rather than the pianists' actions on the keyboard. The answer seems to be negative because Repp and Knoblich found an ADE when pianists merely observed keyboard actions of the experimenter, without seeing the numeric prompts. In that condition, the ADE cannot have been due to the prompts, and even an explicit action was not necessary. The ADE in that case was attributed to internal simulation of observed actions.

However, the information conveyed by an observed action was similar to that conveyed by a numeric prompt in that both referred to a movement in the left-right or right-left direction. Perhaps all that mattered was the directional information, not the manner in which it was conveyed. In that case, numbers referring to a labeled keyboard, in a condition where corresponding actions had to be performed, could have been as effective as observed actions, making the self-performed actions redundant. The directional information, via its strong association with the increase in pitch from left to right on the piano, may have influenced tritone perception or response decisions, or both.

The action prompts in the present study (music notation) were different from those used by Repp and Knoblich (2009) because we wanted to be able to present them divorced from action without sacrificing their action relevance. Numeric prompts would lose their connection with action if they were presented in a passive listening condition, without the labeled keyboard in view. (Although it could be argued that seeing two numbers in ascending order might bias responses to tritones towards "rising," we mentioned already that Repp and Knoblich used both ascending and descending keyboard labels, and found no effect of that variable.) By contrast, music notation does not require the presence of a keyboard to be action-relevant. For musicians, notation is an instruction either to play on an instrument or to imagine sounds that result from such playing, which may amount to imagining the actions along with the sounds. Indeed, one possible interpretation of the absence of an action effect in the present study is that participants' imagined keyboard actions in the passive condition were as effective in creating an ADE as were their actual actions in the active condition.

An alternative explanation is that the notation conveyed directional information that exerted a bias on perception or responses, without any mediating role of action. A purely perceptual bias would have to be mediated by some other internal process that can interact with auditory perception, with auditory imagery being the prime candidate. Auditory imagery in the present paradigm is facilitated by the fact that as soon as the first tone of a tritone is heard, the pitch class of the second tone is certain; only the direction of the interval needs to be imagined, and it is often predictable, too, from preceding context that delineates the total pitch range of the ambiguous tones. A response bias could arise from more abstract cognitive expectations of "rising" or "falling." In the case of our experiment, the bias could have been as simple as a tendency to say "yes" to the question of whether the heard melody matched the notation. A possible counterargument to this suggestion is that Repp and Thompson (2010) found no significant effect of unambiguous auditory primes (rising or falling tritones) on tritone judgment, and no relationship of individual priming effects to the perceived ambiguity of tritones. Whether the effect of notation represents a perceptual or response bias could be tested further in a passive condition by presenting the music notation after the test melody has been heard: A response bias should persist, whereas a perceptual bias should not.

There was also a difference in the nature of the actions between the present study and that of Repp and Knoblich (2009). In their experiment, participants pressed two different keys successively with the same finger and thus had to move their hand horizontally in each trial. In our study, the pianists' hand was stationary in each trial, as different fingers were used to depress three keys in succession. Could it be that lateral movement of the hand and arm is necessary for an ADE to be obtained? We assumed that the spatial relationship of the depressed piano keys was the relevant factor, but we could have been mistaken.

In summary, our results raise questions about the interpretation of the ADE found by Repp and Knoblich (2007, 2009) but cannot answer these questions conclusively. On the one hand, we may have replicated the ADE by evoking action imagery by means of musical notation. On the other hand, we may have shown that the ADE was not really due to keyboard action but represented a perceptual or response bias induced by the informational content of prompts and of observed actions. All we can say with certainty is that there was no effect of overt action in the present experiment. We hope this finding will stimulate further research on the ADE.

Author Note

This research was supported by National Science Foundation grant BCS-0924206. Author contributions: BHR designed the study, did the programming, collected the data, did statistical analyses, and prepared the manuscript for publication. RMG helped design the study, prepared the materials, analyzed the data, and wrote a senior thesis that was submitted to Yale University in April 2009. We are grateful to Peter Pfordresher, whose comments on the Repp and Knoblich (2009) study helped stimulate the present experiment.

Correspondence concerning this article should be addressed to Bruno H. Repp, Haskins Laboratories, 300 George Street, New Haven, CT 06511-6624. E-MAIL: repp@haskins.yale.edu

References

Bangert, M., Jürgens, U., Häusler, U., & Altenmüller, E. (2006). Classical conditioned responses to absent tones. *BMC Neuroscience, 7*, 60. Open access doi: 10.1186/1471-2202-7-60

Deutsch, D. (1986). A musical paradox. *Music Perception, 3,* 275–280.

Deutsch, D. (1987). The tritone paradox: Effects of spectral variables. *Perception and Psychophysics, 41,* 563–575.

Deutsch, D., Kuyper, W. L., & Fisher, Y. (1987). The tritone paradox: Its presence and form of distribution in a general population. *Music Perception, 5,* 79–92.

Deutsch, D., North, T., & Ray, L. (1990). The tritone paradox: Correlate with the listener's vocal range for speech. *Music Perception, 7,* 371–384.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005a). Action-effect coupling in pianists. *Psychological Research, 69,* 233–241.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005b). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *Quarterly Journal of Experimental Psychology, 58A,* 1376–1389.

Drost, U. C., Rieger, M., & Prinz, W. (2007). Instrument specificity in experienced musicians. *Quarterly Journal of Experimental Psychology, 60,* 527–533.

Hamilton, A., Wolpert, D. M., & Frith, C. D. (2004). Your own action influences how you perceive another person's action. *Current Biology, 14*, 493–498.

Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience, 13,* 786–792.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24,* 849–937.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired action. *Journal of Neuroscience, 27,* 308–314.

Miall, R. C., Stanley, J., Todhunter, S., Levick, C., Lindo, S., & Miall, J. D. (2006). Performing hand actions assists the visual discrimination of similar hand postures. *Neuropsychologia, 44,* 966–976.

Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science, 1,* 110–122.

Repp, B. H. (1994). The tritone paradox and the pitch range of the speaking voice: A dubious connection. *Music Perception*, *12*, 227–255.

Repp, B. H. (1997). Spectral envelope and context effects in the tritone paradox. *Perception, 26,* 645–665.

Repp, B. H., & Knoblich, G. (2007). Action can affect auditory perception. *Psychological Science, 18,* 6–7.

Repp, B. H., & Knoblich, G. (2009). Performed or observed keyboard actions affect pianists' judgments of relative pitch. *Quarterly Journal of Experimental Psychology, 62,* 2156–2170.

Repp, B. H., & Thompson, J. M. (2010). Context sensitivity and invariance in perception of octave-ambiguous tones. *Psychological Research, 74,* 437–456.

Schubö, A., Aschersleben, G., & Prinz, W. (2001). Interaction between perception and action in a reaction task with overlapping S-R assignments. *Psychological Research, 65,* 45–57.

Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: Action-induced modulation of perception. *Trends in Cognitive Sciences, 11,* 349–355.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America, 36,* 2345–2353.

Shepard, R. N. (1983). Demonstrations of circular components of pitch. *Journal of the Audio Engineering Society, 31,* 641–649.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin, 131*, 460–473.

Wohlschläger, A. (2000). Visual motion priming by invisible action. *Vision Research, 40,* 925–930.

Wühr, P., & Müsseler, J. (2001). Time course of the blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 1260–1270.

Zwickel, J., Grosjean, M., & Prinz, W. (2007). Seeing while moving: Measuring the online influence of action on perception. *Quarterly Journal of Experimental Psychology, 60,* 1063–1071.

Zwickel, J., Grosjean, M., & Prinz, W. (2010). On interference effects in concurrent perception and action. *Psychological Research, 74,* 152–171.