Music can elicit a visual motion aftereffect

Stephen C. Hedger · Howard C. Nusbaum ·
Olivier Lescop · Pascal Wallisch · Berthold Hoeckner

Abstract Motion aftereffects (MAEs) are thought to result from the adaptation of both subcortical and cortical systems involved in the processing of visual motion. Recently, it has been reported that the implied motion of static images in combination with linguistic descriptions of motion is sufficient to elicit an MAE, although neither factor alone is thought to directly activate visual motion areas in the brain. Given that the monotonic change of musical pitch is widely recognized in music as a metaphor for vertical motion, we investigated whether prolonged exposure to ascending or descending musical scales can also produce a visual motion aftereffect. After listening to ascending or descending musical scales, participants made decisions about the direction of visual motion in random-dot kinematogram stimuli. Metaphoric motion in the musical stimuli did affect the visual direction judgments, in that repeated exposure to rising or falling musical scales shifted participants' sensitivity to visual motion in the opposite direction. The finding that music can induce an MAE suggests that the subjective interpretation of monotonic pitch change as motion may have a perceptual foundation.

Keywords Music cognition · Sound recognition · Motion integration · Visual perception

Visual motion can be described as a space–time correlation (Watson & Ahumada, 1983), which in the real world is typically produced by consistent displacement of an object in space over time. The perception of visual motion yields much survival-relevant information to an organism,

S. C. Hedger (✉) · H. C. Nusbaum · O. Lescop · B. Hoeckner
University of Chicago, Chicago, IL, USA
e-mail: shedger@uchicago.edu

P. Wallisch
New York University, New York, NY, USA

enabling cognitive functions such as image segmentation, breaking of camouflage (Nakayama & Loomis, 1974), heading perception, and navigation (Warren, Blackwell, Kurtz, Hatsopoulos, & Kalish, 1991), among many others.

However, we can infer visual motion even when none is present in the stimulus. For example, observers infer motion even if they are only presented with two subsequent static images (Freyd, 1983). The sense of motion experienced from these static stimulus displays is similar to that evoked by visual motion, and some evidence has indicated shared neural pathways in real and implied motion perception (Kourtzi & Kanwisher, 2000; Lorteije et al., 2006; Peuskens, Vanrie, Verfaillie, & Orban, 2005; Senior et al., 2000). These listed cases all featured variation in contrast across static images, but even that is not necessary. In a study by Winawer, Huk, and Boroditsky (2008), participants viewed static images that each individually implied motion in a specific direction (e.g., from left to right), such as a runner pictured frozen in motion. After prolonged exposure to a sequence of similarly static pictures, each depicting motion in the same direction, participants showed systematically biased forced choice responses to random-dot kinematogram (RDK) motion in the *opposite* direction, which is similar to the well-known motion aftereffect (MAE; Anstis, Verstraten, & Mather, 1998; Purkinje, 1820, 1825; Wohlgemuth, 1911). Motion is not literally perceived in the static adaptors, yet the consequences for judgments of visual motion are consistent with the perception of motion.

Beyond this effect of static images suggesting motion, there is evidence that metaphoric movement—patterns that are not necessarily correlated with visual motion but that provide information about motion—can produce similar effects. Dils and Boroditsky (2010) reported that verbal descriptions of motion can produce contrastive changes in judgments of visual motion, suggesting that even symbolic descriptions of motion can affect motion perception in much

the same way as real or implied visual motion. It thus seems that nonvisual information can affect motion judgments in a way that is similar to physical visual motion.

In music, composers have long used the metaphoric mapping of pitch change to vertical motion. For example, Symphony No. 6 by Franz Joseph Haydn (1761), known as *Le Matin* or "Morning," uses an ascending melodic line to convey the rising of the sun. In fact, the metaphor that equates frequency changes with vertical movement is so strong in Western culture that it often pervades the everyday lexicon used to describe musical events (e.g., the *falling* bass line/the *rising* melody). When referring to a rising melody played on a keyboard instrument, nothing physically rises; the hands move in a left-to-right fashion. An even more striking example of the pitch–verticality metaphor is found in the language describing the movements of large stringed instruments, such as the cello or double bass. For these instruments, playing a rising melody requires one to move *down* in vertical space, yet the gesture is often called "moving up the fingerboard."

While the origin of these musical metaphors remains unclear, likely incorporating cultural norms and knowledge (e.g., Eitan & Tubul, 2010) as well as embodied experience (Turner, 1987), the fact remains that these conceptual metaphors are pervasive in the understanding of musical meaning. Thus, scales and melodies that increase and decrease in frequency are classified as rising and falling, respectively, yet musical pitch changes are correlates neither of movement in the vertical axis nor necessarily of physical consequences of the movement. Since understanding musical change as movement is an important way in which music has been thought to convey information, auditory metaphors in descriptions of music are used extensively to convey vast amounts of information about dynamic processes, such as movement (Gjerdingen, 1994; Johnson & Larson, 2003; Todd, 1992). Moreover, listeners can certainly understand some of these musical devices as intended: Eitan and Granot (2006) asked participants to imagine movement as they heard several short music motifs, and they found consistent mappings between the manipulation of certain musical parameters (e.g., pitch, acceleration) and bodily movement.

Given the strength and pervasiveness of the pitch–verticality metaphor, one might hypothesize that specific frequency changes can cross-modally produce changes in motion judgments. Maeda, Kanai, and Shimojo (2004) demonstrated that continuously ascending or descending pitch glides could disambiguate visual motion, in that ambiguous visual motion presented with an ascending auditory stimulus was more likely to be judged as moving up, and ambiguous visual motion presented with a descending auditory stimulus was more likely to be judged as moving down. Furthermore, metaphoric auditory signals have been shown to facilitate the comprehension of implied visual motion (Hedger et al.,

2011) as well as real motion (Sadaghiani, Maier, & Noppeney, 2009). Thus, if congruent metaphoric auditory motion can facilitate comprehension of both implied and real visual motion, it also may be possible that prolonged exposure to a particular implied direction of motion in either audition or vision could shift sensitivity to congruent motion in the opposite modality, in a manner similar to an MAE. An auditorily induced MAE would more directly implicate a basic visual mechanism than it would top-down changes in visual judgment.

This is the question that Kitigawa and Ichihara (2002) addressed by adapting participants to either auditory or visual motion, and then measuring the degree to which adaptation in one modality elicited an MAE in the other modality. The authors found that an auditory aftereffect is elicited from adaptation by visual motion, but not the converse, providing evidence for visual dominance in cross-modal aftereffects. However, this asymmetry might have had to do with the particular stimuli, for the authors investigated expanding/contracting visual stimuli and perceived dynamics for auditory stimuli. It is thus possible that the pitch–verticality correspondence is a stronger metaphoric mapping, and thus can elicit a visual MAE from purely auditory adaptation. This is one of the hypotheses tested by Jain, Sally, and Papathomas (2008). They found that after adaptation to continuously rising or falling auditory motion, participants were more likely to judge ambiguous sinusoidal gratings as moving in the opposite direction. This was the only auditory-to-visual adaptation that produced a significant MAE (adaptation to rightward/leftward panned auditory motion, as well as looming/receding auditory motion, did not result in a visual MAE). This suggests that experiencing a *continuous* pitch increase as a pitch rise can affect our perception of visual motion.

In music, the relationship between pitch changes and vertical motion might not be as straightforward as it might seem from the foregoing discussion. More specifically, in the previously outlined studies, the auditory stimuli consisted of continuous acoustic frequency changes, whereas notes are generally presented as discrete steps in music. In this manner, pitch changes in music are more similar to apparent visual motion (e.g., a dot being flashed in opposite sides of a screen, with no physical signal linking the perceived motion). Furthermore, representing pitch change as music might also shift attention away from motion information (e.g., one might attend to melodic, timbral, or scalar information, without directly attending to contour).

To test whether an MAE could be elicited by music, we measured the effects of prolonged listening to ascending and descending musical scales on visual perception. If the perceptual metaphor of pitch change in music selectively adapts direction-selective neurons in motion-processing areas, as physical motion does, we would predict similar changes in

perception. Thus, taking the results obtained by Winawer et al. (2008) to their logical conclusion, we predict a cross-modal MAE, in that repeated exposure to ascending musical scales would result in a systematic shift that would favor downward motion, while repeated exposure to descending musical scales would result in a systematic shift favoring upward motion. The present study expands upon the previous literature by testing whether auditory signals that are constrained by the parameters of Western music are sufficient to elicit cross-modal MAEs.

Method

Participants

A group of 43 University of Chicago undergraduates were recruited to participate in the experiment. All of the participants were naive to the purposes of the experiment, did not have any hearing problems, and had normal or corrected-to-normal vision. Three of the participants did not perform adequately on the baseline task (see the Baseline Task section for details), and thus did not participate in the main task.

Materials

The experiment was displayed on a cathode-ray-tube (CRT) monitor with a screen resolution of 1,024×768 pixels and a refresh rate of 75 Hz. The music stimuli, which consisted of a sine-wave timbre, included ten different musical scales (Ionian, Aeolian, Dorian, Phrygian, Mixolydian, chromatic, whole-tone, octatonic, hexatonic, and pentatonic) starting on every note within an octave range, resulting in 120 music stimuli for both upward and downward motion. Each scale consisted of eight notes, meaning that some scales (e.g., chromatic) did not span an octave range, while others (e.g., whole-tone) spanned slightly more than an octave range. The duration of each musical scale was 500 ms (thus, each note lasted 62.5 ms), and all of the stimuli were normalized to an average of 70 dB SPL. While discrete notes lasting only 62.5 ms might seem too fast to be perceived musically, there are numerous examples of musical events (e.g., trills) that occur at this speed. Moreover, in a separate discrimination study, we found that participants were able to discriminate between types of scales with over 95 % accuracy.

The RDK stimuli used in this task were based on those used by Newsome and Paré (1988) and had previously been utilized to measure MAEs from adaptation to real (Blake & Hiris, 1993) as well as implied (Winawer et al., 2008) motion. Using RDK stimuli marks an important deviation from the previous literature on cross-modal MAEs

(e.g., Jain et al., 2008), as RDK stimuli lack any trackable visual features and have a much richer spatiotemporal frequency content. Each RDK stimulus consisted of 100 moving dots. This stimulus display took up approximately one-third of the screen, or about 12º×9º of visual angle. Participants were seated approximately 70 cm away from the computer screen. On each trial, a subset of the dots moved coherently either up or down by approximately 0.05º per frame. All of the other dots disappeared and reappeared at random points within the rectangular window. Each trial lasted 1 s, consisting of 25 stimulus frames of 40 ms each.

Procedure

Baseline task Before participating in the main experiment, participants completed a baseline measure to estimate their motion detection thresholds. The baseline task consisted of 390 randomly ordered 1-s RDK stimuli (195 with upward motion, and 195 with downward motion). The dot-motion coherence levels in the baseline task ranged from 5 % to 65 %, in 5 % increments. A logistic function was fit to the responses, in which "downward" was arbitrarily coded as negative, and "upward" as positive. Asymptotic performance in both directions was calculated for each participant on the basis of the fitted logistic function. The point at which participants reached asymptotic performance was defined as the maximal coherence unit for each participant. In the main task, we tested motion judgments for RDKs with the maximal coherence unit, 50 % of the maximal coherence unit, and 25 % of the maximal coherence unit as test stimuli, resulting in six types of visual stimuli per participant (three levels of upward motion and three levels of downward motion). For instance, if a participant reached asymptotic performance at 40 % coherence, their coherence levels would be 40 %, 20 %, and 10 % for the main task. Three participants were likely unable to detect the motion contained within the RDK stimuli, as they performed so poorly on the baseline task that their normalized coherence unit would have been over 100 %. Consequently, they did not continue on to perform the main task.

Main task Following the baseline task, participants completed the main portion of the experiment, which consisted of four blocks. Each of these blocks consisted of 30 trials. A trial consisted of a presentation of music followed by a visual RDK display for which an "up" or "down" decision was made. The presentation of each block was structured so that participants listened to a set of either ascending or descending musical scales (one direction only per block) for either 60 s (prior to the first RDK test trial) or 6 s (all subsequent RDK test trials). Thus, two ascending and two descending music blocks were presented, with no adjacent

blocks featuring the same musical direction, and the first block being randomly determined across participants. As each musical scale lasted 500 ms, before the first RDK test trial, participants heard all of the musical scales (120) in a randomized order, and for each of the subsequent trials they heard a randomly selected subset of the musical scales (12), also presented in a randomized order, prior to each visual direction judgment.

Following each presentation of the music stimuli, participants made a forced choice judgment about the direction of motion represented by the coherent dot motion within an RDK, each of which lasted for 1 s and had a coherence level of 100 %, 50 %, or 25 % of the participant's maximal coherence unit. The participants received all six RDK stimuli (three coherence levels apiece of upward and downward motion) five times per block, in a randomized fashion. Figure 1 provides a diagram of the experimental design.

Questionnaires After completing the main task, each participant filled out two questionnaires (see the Appendix). The first questionnaire addressed musical background (e.g., how many instruments one played and for how long), while the second questionnaire addressed previous knowledge about the MAE (e.g., whether participants had ever heard of MAEs). After completing the questionnaires, participants were debriefed and compensated.

Results

Responses

Since participants' responses were binary ("up" or "down"), we ran a generalized linear mixed-effects model with a logistic link (e.g., Agresti, 2007; Baayen, Davidson &

Bates, 2008). The fixed effects of the model were coherence level (an ordered factor) and music direction (ascending or descending), while video number (trial) and participants were treated as random effects.

To measure model fits, we used the corrected Akaike information criterion for finite sample sizes (AICc). The AICc provides a goodness-of-fit measurement for models while providing a penalty for extra parameters. To compare between models, we used a delta AICc measure (Δ_i), which is obtained by subtracting the AICc of the best model and the AICc of the tested model. As a general rule of thumb, a Δ_i value less than 2 suggests substantial evidence for the model, while a Δ_i between 3 and 7 indicates that the model has considerably less support, while a Δ_i over 10 suggests that the model is highly unlikely (Burnham & Anderson, 2002).

The best model that fit the data included an interaction between music direction and the quadratic component of a polynomial regression. In other words, removing the quadratic component resulted in a worse model (Δ_i=6). Similarly, removing the interaction resulted in a worse model (Δ_i=11). Finally, removing music direction resulted in a much worse model (Δ_i=15). Thus, it appears that the participants' responses differed as a function of whether or not they were listening to ascending or descending music, and the magnitude of this difference between ascending and descending music depended on the coherence level. Moreover, this interaction is consistent with the MAE, in that descending music trials resulted in more "upward" RDK decisions, and ascending music trials resulted in fewer "upward" RDK decisions. Figure 2 shows this interaction.

These results converge with more traditional psychophysical measurements, such as points of subjective equality (PSEs). In this study, the PSE was defined as the coherence level that was required to obtain the 50 % point on the ordinate (i.e., where the fitted logistic function reached the
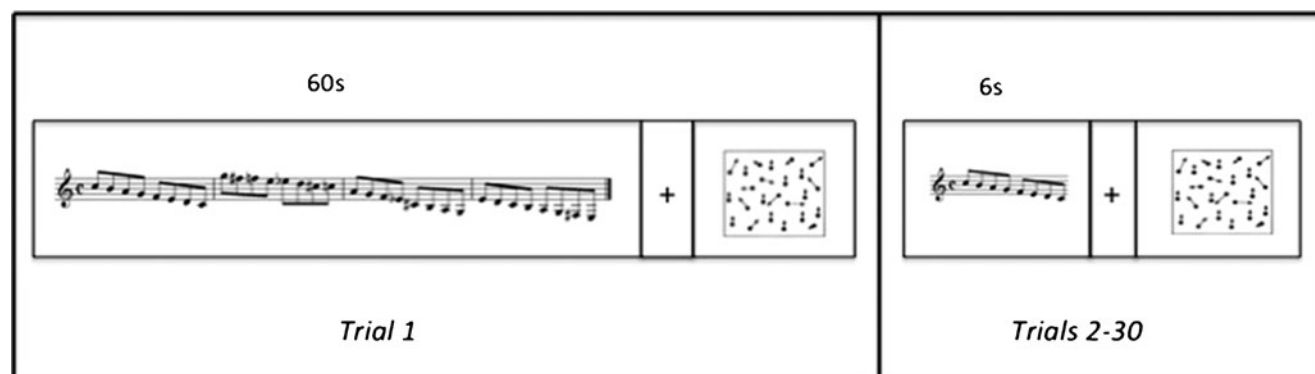

Fig. 1 Sample block from the experiment. For the first trial, participants heard 60 s (120 scales) of either ascending or descending motion, after which they judged the direction (up or down) of coherently moving dots. In the subsequent trials, participants only heard 6 s (12 scales) of music before making their decision. The music direction was consistent within a block
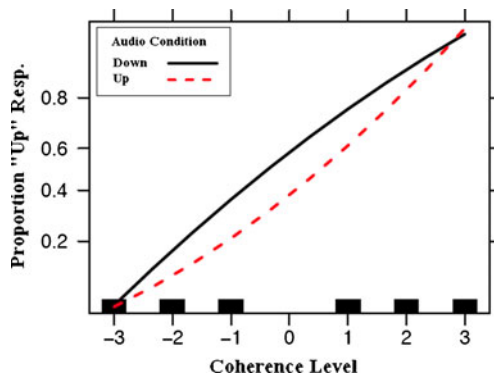
Fig. 2 Mean proportions of "up" responses as a function of coherence level and music direction, plotted on a logistic scale. For the more difficult-to-detect coherence levels, the music direction influenced participants' responses in a manner that was congruent with the motion aftereffect, in that adaptation to descending scales resulted in more upward random-dot kinematogram (RDK) responses, and adaptation to ascending scales resulted in fewer upward RDK responses

point at which participants were equally likely to judge the visual stimulus as ascending or descending). Since the coherence levels were personally calculated for each participant, if the music had no effect on visual perception, we would expect both the ascending and descending music conditions to yield a PSE of 0 normalized coherence units. However, contrary to this null hypothesis, we found a significant shift in the PSE based on the metaphoric direction of the music (i.e., ascending or descending). Averaged across participants, the PSEs were $-.15$ normalized units of coherence for ascending musical stimuli, and $.23$ normalized units of coherence for descending musical stimuli (with downward and upward visual motion being arbitrarily coded as negative and positive, respectively). This difference in PSEs was significant [$t(39)=2.87$, $p<.01$]. In other words, after prolonged exposure to metaphorically descending music, participants' PSEs were shifted toward making more ascending responses for the RDK stimuli, and after prolonged exposure to metaphorically ascending music, participants' PSEs were shifted toward making more descending responses.

Music experience

In order to assess whether the impact of the music adaptor was moderated by individual musical experience, we added instrument experience as a covariate in the model. We added this covariate because of the possibility that certain instruments might support the metaphoric mapping of pitch to vertical space more than others. Specifically, we looked at the participants' years of experience with instruments that have a congruent metaphoric mapping between pitch and vertical space (e.g., clarinet, oboe)—in which

ascending physical motions produce higher pitches—as we thought that participants who had extensive experience on these instruments might show a larger MAE. An exhaustive search for the best generalized logistic mixed-effects models, however, did not include music experience as a covariate.

Discussion

While the metaphoric mapping between pitch and vertical space is pervasive in Western culture, the genesis of this metaphor and the mechanism that mediates its interpretation are less clear. Specifically, the understanding of this metaphor could either be an inferential effect, based on cultural norms and knowledge, or it could be determined by perceptual effects that are rooted in the same mechanism by which we perceive visual motion. The present results provide evidence for the latter interpretation.

Paradigms such as the visual motion aftereffect have long been interpreted in terms of the nature of visual neural circuits (e.g., Barlow & Hill, 1963). Specifically, the visual MAE is taken as reflecting the reduction in neural sensitivity of direction-tuned opponent mechanisms from prolonged exposure to one direction of motion, thus increasing sensitivity to the other direction. However, the present study shows that visual motion perception is affected after listening to metaphoric musical motion, in a manner consistent with real and static visual motion adaptation. This suggests that the metaphoric motion information conveyed by music does not merely affect postperceptual reasoning. Taken in the context of the general interpretation of the MAE, this cross-modal transfer of adaptation from metaphoric auditory motion to real visual motion provides evidence that ascending and descending musical scales may well be stimulating visual direction-selective neurons. This conclusion is supported by the fact that no participant correctly identified the purpose of the study, and only a quarter of the participants had even heard of the MAE. Furthermore, previous knowledge of the MAE did not predict the observed effect, and upon debriefing, most participants described the music as "video game music," rather than rising or falling.

Even though we have provided evidence that the perception of metaphoric motion in music may rely on the same perceptual mechanisms by which we process visual motion, this does not necessarily mean that the auditory frequency–vertical space metaphor is innate. Specifically, although the auditory frequency–vertical space metaphor is pervasive throughout Western music, it is by no means the only metaphoric mapping used to understand auditory frequency (e.g., Eitan & Timmers, 2010). Indeed, the prevalent metaphoric

mapping in Bali and Java equates auditory frequency with physical size, where "low" frequencies are "large" and "high" frequencies are "small" (Van Zanten, 1986), while the Suyá in the Amazon equate auditory frequency with age, with "low" frequencies being "old" and "high" frequencies "young" (Seeger, 1987). As we assume that the effects that we report are indeed driven by metaphors, we predict that such cultures would not experience a visual aftereffect in the same fashion that we observed.

Indeed, previous studies have demonstrated that musical metaphors—as culturally driven phenomena—are not as strongly developed in children (e.g., Eitan & Tubul, 2010). Yet, other research has provided evidence for a pitch–verticality correspondence in infants (e.g., Walker et al., 2010), suggesting that this metaphoric mapping is not necessarily a byproduct of verbal metaphor or experienced correlations. Further studies should thus address the developmental and cultural components of this effect, as this would help disambiguate the extent to which the present findings were culturally derived.

The results of this study can perhaps best be explained by means of cross-modal associations that, over time, have become internalized and automatized. From the Western musical notation system, which uses vertical space to represent notes, to the common use of pitch changes to metaphorically convey object motion, it is likely that through repeated exposure in the environment, we have come to automatically associate decreasing and increasing musical pitch with descending and ascending objects, respectively. Indeed, the importance of statistical regularities on perception has long been discussed, and it is clear that observers do take advantage of this information (e.g., Barlow, 1959; Craik, 1943; Weiss, Simoncelli, & Adelson, 2002). From a neural perspective, the repeated pairing of specific auditory events with real motion information has possibly resulted in neurons that respond to both real and metaphoric motion in a direction-selective manner.

A related possibility is that the auditory metaphors did not act through an automatized representation, but rather through visual mental imagery. It is conceivable that the metaphor evoked a visual image in the minds of the observers, which then produced the observed adaptation effects. This is not inconceivable, as something like this has recently been demonstrated (Winawer, Huk, & Boroditsky, 2010). One way to distinguish these possibilities would be to administer a test of visual mental imagery (e.g., the VVIQ) to participants and see whether visual imagery scores predicted the magnitude of the effect.

Recent work on the role of analog acoustic expression (AAE) speech comprehension may also be instructive in interpreting these results. Indeed, the first empirical investigation of AAE demonstrated that people modulate the pitch of their voice while speaking in order to analogically convey vertical motion information about an object (Shintel, Nusbaum, & Okrent, 2006). While it is unclear whether music mimics the acoustic patterns in speech, or vice versa, the evidence for this metaphor in speech as well as music demonstrates the strength of this association in Western culture.

Finally, our results fit nicely into the wider recent literature on multimodal interactions between auditory and visual motion. For instance, it has been shown that some synesthetes immediately perceive specific sounds when presented with particular kinds of visual motion (Saenz & Koch, 2008) or that the perceived speed of visual motion influences judgments of auditory tempo (Su & Jonikaitis, 2011). These couplings do seem to go both ways, as according to some reports music can influence visual perception in a top-down fashion, either directly (Hidaka et al., 2011) or indirectly—for instance, via emotion (Jolij & Meurs, 2011) or by communicative gestures that have been shown to improve motion perception (Manera, Becchio, Schouten, Bara, & Verfaillie, 2011). Sounds have even been shown to induce motion perception for static displays (Teramoto et al., 2010). As in our study, the meaning of the auditory signals does matter; for instance, in tap dancing, auditory interacts with visual information in a matter that depends on how synchronized the information sources are (Arrighi, Marini, & Burr, 2009). Taken together, these results suggest a plethora of interactions between the auditory system and the perception of visual motion.

In conclusion, we found a cross-modal MAE in which direction-selective motion adaptation resulted from listening to metaphorically moving musical scales. These behavioral results suggest that extracting motion information from music involves the use of direction-selective motion mechanisms, and is not solely the result of top-down or inferential effects. While it remains unclear whether or not these metaphoric auditory events recruit the same neural networks involved in the processing of auditory and visual motion, these results provide empirical evidence that the processing of metaphoric musical motion shares perceptual mechanisms with the perception of visual motion.

Appendix: Postexperimental questionnaires

Music Experience Questionnaire

Age: _____

Gender: _____

Hand (Circle one):   RIGHT   LEFT

    1. Do you play (or have you played) any musical instruments?   YES   NO

    1a. At what age (approximately) did you first begin musical instruction?   _____

    1b. Do you actively play a musical instrument?   YES   NO

    2. If you circled YES in Question 1, please list all the musical instruments you have played, including the number of years you have played them:

Instrument_____   Number of Years_____

Instrument_____   Number of Years_____

Instrument_____   Number of Years_____

Instrument_____   Number of Years_____

Instrument_____   Number of Years_____

    Voice_____   Number of Years_____

    3. Do you know how to read music?   YES   NO

    4. Have you ever played in any ensembles or sung in any choirs?   YES   NO

    5. If you answered YES to Question 4, please list the ensembles and/or choirs with which you played/sung.

Ensemble/Choir_____   Number of Years_____

Ensemble/Choir_____   Number of Years_____

    6. Have you taken any music courses at the high school or collegiate level?   YES   NO

    7. If you answered YES to Question 6, please list the approximate number of music courses you have completed   _____

    8. Do you have perfect pitch (the ability to name any note without a reference)?   YES   NO

    9. Approximately how many hours do you spend making music per week? (circle one):

    Less than one hour   1-2 hours   3-4 hours   5-6 hours   More than 6 hours

    10. Approximately how many hours do you spend listening to music per day (circle one):

    Less than one hour   1-2 hours   3-4 hours   5-6 hours   More than 6 hours

    11. Approximately what percentage of music you listen to falls under the following genres?

    Pop/Rock   _____

    Dance/Techno   _____

    Jazz   _____

    R&B   _____

    Classical   _____

    Non-Western/Ethnic   _____

    Other (please specify) _____

    12. Approximately how many concerts do you attend per year?   _____

Knowledge of Motion Aftereffect (MAE) Questionnaire (To be orally administered during debriefing)

1. Have you ever heard of something called the motion aftereffect, or "waterfall illusion?"
2. If your answer to the first question was YES, could you please try to explain the motion aftereffect to the best of your ability?
3. (For those who could not explain the MAE): If you had to guess, would you say that looking at downward motion for a prolonged period of time would make stationary objects appear to move DOWN or UP?

References

Agresti, A. (2007). *An introduction to categorical data analysis*. New York, NY: Wiley.

Anstis, S., Verstraten, F. A. J., & Mather, G. (1998). The motion aftereffect. *Trends in Cognitive Sciences, 2,* 111–117. doi:10.1016/S1364-6613(98)01142-5

Arrighi, R., Marini, F., & Burr, D. (2009). Meaningful auditory information enhances perception of visual biological motion. *Journal of Vision, 9*(4), 25. doi:10.1167/9.4.25. 1–7.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. doi:10.1016/j.jml.2007.12.005

Barlow, H. B. (1959). Sensory mechanisms, the reduction of redundancy, and intelligence. *NPL Symposium on the Mechanization of Thought Process, 10,* 535–539.

Barlow, H. B., & Hill, R. M. (1963). Evidence for a physiological explanation of the waterfall illusion. *Nature, 200,* 1345–1347.

Blake, R., & Hiris, E. (1993). Another means for measuring the motion aftereffect. *Vision Research, 33,* 1589–1592. doi:10.1016/0042-6989(93)90152-M

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodal inference: A practical information-theoretic approach*. New York, NY: Springer.

Craik, K. J. W. (1943). *The nature of explanation*. Cambridge, UK: Cambridge University Press.

Dils, A. T., & Boroditsky, L. (2010). Processing unrelated language can change what you see. *Psychonomic Bulletin & Review, 17,* 882–888. doi:10.3758/PBR.17.6.882

Eitan, Z., & Granot, R. Y. (2006). How music moves: Musical parameters and listeners' images of motion. *Music Perception, 23,* 221–247. doi:10.1525/mp.2006.23.3.221

Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition, 114,* 405–422.

Eitan, Z., & Tubul, N. (2010). Musical parameters and children's images of motion. *Musicae Scientiae, 2010,* 89–111.

Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics, 33,* 575–581.

Gjerdingen, R. O. (1994). Apparent motion in music. *Music Perception, 11,* 335–370.

Haydn, F. J. (1761). *Symphony No. 6 in D major ("Le Matin")*, Hob. I:6.

Hedger, S. C., Nusbaum, H. C., & Hoeckner, B. (2011). *Music can imply motion like prosody in speech*. Rochester, NY: Lecture given at the Society for Music Perception and Cognition.

Hidaka, S., Teramoto, W., Sugita, Y., Manaka, Y., Sakamoto, S., & Suzuki, Y. (2011). Auditory motion information drives visual motion perception. *PLoS One, 6,* e17499. doi:10.1371/journal.pone.0017499

Jain, A., Sally, S. L., & Papathomas, T. V. (2008). Audiovisual short-term influences and aftereffects in motion: Examination across three sets of directional pairings. *Journal of Vision, 8*(15), 7. doi:10.1167/8.15.7. 1–13.

Johnson, M. L., & Larson, S. (2003). "Something in the way she moves"—Metaphors of musical motion. *Metaphor and Symbol, 18,* 63–84. doi:10.1207/S15327868MS1802_1

Jolij, J., & Meurs, M. (2011). Music alters visual perception. *PLoS One, 6,* e18861. doi:10.1371/journal.pone.0018861

Kitigawa, N., & Ichihara, S. (2002). Hearing visual motion in depth. *Nature, 416,* 172–174.

Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience, 12,* 48–55. doi:10.1162/08989290051137594

Lorteije, J. A. M., Kenemans, J. L., Jellema, T., van der Lubbe, R. H. J., de Heer, F., & van Wezel, R. J. A. (2006). Delayed response to animate implied motion in human motion processing areas. *Journal of Cognitive Neuroscience, 18,* 158–168. doi:10.1162/jocn.2006.18.2.158

Maeda, F., Kanai, R., & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology, 14,* R990–R991.

Manera, V., Becchio, C., Schouten, B., Bara, B. G., & Verfaillie, K. (2011). Communicative interactions improve visual detection of biological motion. *PLoS One, 6,* e14594.

Nakayama, K., & Loomis, J. M. (1974). Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis. *Perception, 3,* 63–80.

Newsome, W. T., & Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience, 8,* 2201–2211.

Peuskens, H., Vanrie, J., Verfaillie, K., & Orban, G. A. (2005). Specificity of regions processing biological motion. *European Journal of Neuroscience, 21,* 2864–2875. doi:10.1111/j.1460-9568.2005.04106.x

Purkinje, J. E. (1820). Medicinische Jahrbucher des kaiserlich–koniglichen osterreichischen Staates. *Medicinische Jahrbucher des kaiserlich–koniglichen osterreichischen Staates, 6,* 79–125.

Purkinje, J. E. (1825). *Beobachtungen und Versuche zur Physiologie der Sinne. Neue Beitrage zur Kenntniss des Sehens in subjektiver Hinsicht*. Berlin, Germany: Reimer.

Sadaghiani, S., Maier, J. X., & Noppeney, U. (2009). Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *Journal of Neuroscience, 29,* 6490–6499.

Saenz, M., & Koch, C. (2008). The sound of change: Visually induced auditory synesthesia. *Current Biology, 18,* R650–R651.

Seeger, A. (1987). *Why Suyá sing: A musical anthropology of an Amazonian people*. Cambridge, UK: Cambridge University Press.

Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., et al. (2000). The functional neuroanatomy of implicit-motion perception or "representational momentum. *Current Biology, 10,* 16–22.

Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language, 55,* 167–177. doi:10.1016/j.jml.2006.03.002

Su, Y. H., & Jonikaitis, D. (2011). Hearing the speed: Visual motion biases the perception of auditory tempo. *Experimental Brain Research, 214,* 357–371.

Teramoto, W., Manaka, Y., Hidaka, S., Sugita, Y., Miyauchi, R., Sakamoto, S., et al. (2010). Visual motion perception induced by sounds in vertical plane. *Neuroscience Letters, 479,* 221–225.

Todd, N. P. (1992). The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America, 91,* 3540–3550. doi:10.1121/1.402843

Turner, M. (1987). *Death is the mother of beauty: Mind, metaphor, criticism.* Chicago, IL: University of Chicago Press.

Van Zanten, W. (1986). The tone material of the kacapi in tembang sunda in West Java. *Ethnomusicology, 30,* 84–112.

Walker, P., Gavin Bremner, J., Mason, U., Spring, J., Mattock, K., Slater, A., et al. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science, 21,* 21–25. doi:10.1177/0956797609354734

Warren, W. H., Jr., Blackwell, A. W., Kurtz, K. J., Hatsopoulos, N. G., & Kalish, M. L. (1991). On the sufficiency of the velocity field for perception of heading. *Biological Cybernetics, 65,* 311–320.

Watson, A. B., & Ahumada, A. J. (1983). A look at motion in the frequency domain. *NASA Technical Memorandum, 84352,* 1–10.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5,* 598–604.

Winawer, J., Huk, A. C., & Boroditsky, L. (2008). A motion aftereffect from still images depicting motion. *Psychological Science, 19,* 276–283. doi:10.1111/j.1467-9280.2008.02080.x

Winawer, J., Huk, A. C., & Boroditsky, L. (2010). A motion aftereffect from visual imagery of motion. *Cognition, 114,* 276–284. doi:10.1016/j.cognition.2009.09.010

Wohlgemuth, A. (1911). *On the after-effect of seen movement* (*British Journal of Psychology* Monograph Suppl. No. 1). Cambridge, UK: Cambridge University Press.