Cognitive Processing (2019) 20:495–506
https://doi.org/10.1007/s10339-019-00929-8

RESEARCH ARTICLE

Reaching for the high note: judgments of auditory pitch are affected by kinesthetic position

Autumn B. Hostetter[1] · **Christina M. Dandar**[1] · **Gabrielle Shimko**[1] · **Colin Grogan**[1]

Received: 22 May 2019 / Accepted: 12 August 2019 / Published online: 21 August 2019

Abstract

Auditory pitch is represented on a vertical continuum (e.g., high vs. low). In three experiments, we examined whether the kinesthetic experience of moving in a particular direction (e.g., walking up vs. down stairs; reaching up vs. down) would affect judgments of auditory pitch. Participants listened to three tones twice each, once while moving upward and once while moving downward, and estimated the pitch of each tone. In all experiments, participants' judgments of the tones' pitch were biased in the direction of their movement. Moreover, this effect is not due to visibility of the movement or to using a numerical response method. Our results suggest that kinesthetic information from one's own bodily movements biases pitch estimation, and several possible mechanisms for the effect are discussed.

Keywords Auditory pitch · Kinesthetic information · Cross-modal priming

Introduction

The perceived pitch of a tone is determined by the wavelength of the tone's soundwave. A soundwave with a long wavelength that oscillates relatively few times per second is perceived as producing a low pitch, whereas a soundwave with a short wavelength that oscillates many more times per second is perceived as producing a high pitch. These terms ("low" vs. "high") are indicative of the vertical spatial metaphor that is often used to think about and describe pitch (see Fernández-Prieto et al. 2017). This vertical system for representing pitch is also evident in modern musical notation, where the vertical position of a note on the scale indicates its intended pitch. In the present study, we consider whether activating representations of higher or lower vertical space—specifically through bodily experience—might affect the perceived pitch of a tone.

Decades of research suggest that people represent pitch on a vertical continuum (e.g., Bonetti and Costa 2017; Occelli et al. 2009; Roffler and Butler 1968; Rusconi et al. 2006; Spence 2011). For example, Pratt (1930) explicitly asked participants to position tones in space and found that tones with a higher pitch were positioned spatially higher than tones with a lower pitch. More recently, Evans and Treisman (2010) used a more indirect measure and found that participants were faster to make a visual judgment about a stimulus that appeared on the top part of the screen when the stimulus was paired with a high pitch than when the stimulus was paired with a low pitch (and vice versa for when the stimulus appeared on the bottom part of the screen). Similarly, Mossbridge et al. (2011) found that ascending pitches primed participants to look up, whereas descending pitches primed participants to look down.

In addition to priming participants to look in a particular vertical location, pitch can also create a visual illusion of vertical movement. Maeda et al. (2004) found that participants perceived ambiguous motion as moving up when it was accompanied by an ascending pitch, but perceived the same ambiguous motion as moving down when it was accompanied by a descending pitch. This effect appears to occur at intermediate stages of neural integration between auditory and visual cortex, rather than low-level areas associated with integrating auditory motion and vision or higher-level areas associated with integrating linguistic labels and

Handling editor: Riccardo Brunetti (European University of Rome).

Reviewers: Claudia del Gatto (European University of Rome), Fernando Marmolejo-Ramos (Stockholm University).

✉ Autumn B. Hostetter
autumn.hostetter@kzoo.edu

[1] Department of Psychology, Kalamazoo College, 1200 Academy Street, Kalamazoo, MI 49006, USA

vision (Sadaghiani et al. 2009). The cross-modal association between pitch and vision is thus not the sole result of linguistic labeling with a shared term (e.g., Dolscheid and Casasanto 2015), and even infants under 1 year of age have the association (e.g., Wagner et al. 1981).

This cross-modal priming between visual experience and pitch also works the other way, with visual experience looking in a particular direction biasing the pitch that is heard or produced (e.g., Dolscheid et al. 2013). For example, Connell et al. (2013) found that participants' pitch discriminations were biased by vertical movements in the singer's gesture (e.g., motioning upward or downward). When participants viewed a video of a singer producing musical notes, they judged pitches to be higher when the singer gestured upward and lower when the singer gestured downward. The effect of gesture observed by Connell et al. persisted under a verbal load, suggesting that the origin of the effect is not in verbal labeling of the stimuli as high or low. Similarly, Dolscheid et al. found that viewing a stimulus positioned higher on a computer screen resulted in the sung production of a higher pitch than viewing a stimulus positioned lower, but not for Farsi speakers (who do not have a linguistic vertical metaphor for pitch). Taken together, these results show that cross-modal priming between visual experience of vertical space and the heard or produced pitch of an accompanying sound is not dependent on verbal labeling of the pitches or locations, but may be dependent on exposure to a language that codifies pitch in vertical terms.

There are similar cross-modal effects between auditory pitch and motor responses in a particular direction, where hearing a high versus low tone primes arm movements in an upward or downward direction [see Spence (2011) for a review]. Salgado-Montejo et al. (2016) asked participants to move a mouse to represent various tones and found that participants tended to move the mouse in upward trajectories while listening to a high pitch and in downward trajectories while listening to a low pitch. A similar effect occurs on more indirect tasks, where participants are not explicitly asked to represent the tone in their movements. Rusconi et al. (2006) demonstrated that participants were faster to respond to a high tone by pressing a key located higher in space than by pressing a key located lower in space. Conversely, they were faster to respond to a low tone by pressing a key that was lower in space. This effect occurs even for speakers of languages that do not linguistically label pitch in vertical terms, but the effect is smaller than for English speakers (see also Fernández-Prieto et al. 2017).

The cross-modal mapping between pitch and movement is often embodied when a pitch is produced. For example, Lemaitre et al. (2017) found that participants gesture upwards when imitating a higher pitch and gesture downwards when imitating a lower pitch. Huron et al. (2009) found that the facial expressions adopted by participants as they sang a high pitch were judged as friendlier (with more upturned mouths and raised eyebrows) than the facial expressions adopted by participants as they sang a low pitch. Further, it appears that adopting a particular facial posture can influence the pitch of sung notes. Huron and Shanahan (2013) found that participants who were instructed to hold their eyebrows in a raised, neutral, or low position produced a pitch that positively correlated with the position of their eyebrows (i.e., raised eyebrows correlated with a higher produced pitch). Such studies suggest that production of a particular pitch is often accompanied by a facial expression or gesture that embodies the vertical "height" of the pitch, and the effect appears to be bidirectional.

The research to date has established the bidirectionality of cross-modal priming between auditory pitch and visual experience (e.g., Dolscheid et al. 2013; Evans and Treisman 2010). Further, there is some evidence that the priming between bodily movement and auditory pitch *production* is also bidirectional (e.g., Huron and Shanahan 2013; Lemaitre et al. 2017). However, although there is evidence that perception of a pitch primes movements in the corresponding direction (e.g., Rusconi et al. 2006), no studies have yet tested whether this priming effect between bodily movement and pitch *perception* is also bidirectional. Does movement in a particular vertical direction affect the perceived pitch of a tone? We examine this question in three experiments.

Experiment 1

Participants were instructed to walk up and down two flights of stairs while listening to a tone through headphones. When the participants arrived at the intended destination (the bottom of the stairs or the top of the stairs), participants estimated the pitch (in Hz) of the tone they had heard. We hypothesized that estimations of auditory pitch would be influenced by the direction that participants walked on the stairs. Specifically, we predicted that participants would judge a pitch as higher (i.e., a greater number of Hz) when they walked up the stairs and judge the same pitch as lower when they walked down the stairs.

A fair test of this hypothesis requires that participants make their judgments after having sufficient experience walking up and down the stairs. That is, we did not want participants to make a judgment immediately upon hearing each tone. To distract participants from the pitch judgment task as they walked up or down the stairs, we inserted a number of random pulses into each tone and asked participants to count the number of pulses they heard. In addition to directing participants' attention away from the pitch judgment task, the act of counting also occupies participants' verbal working memory (see Imbo and Vandierendonck 2007), making it difficult for them to label each tone covertly to themselves

before being asked for their judgment when they reached the end of the stairs.

Methods

Participants

A convenience sample of 60 students (62% female, 35% male, 3% non-binary; $M_{age} = 20.6$ years, $SD_{age} = .96$) was recruited by verbal request to participate in a study about the effects of exercise on memory. The majority of participants self-identified as Caucasian (70%), with smaller percentages identifying as Asian (10%), African American (7%), Indian (5%), Chinese (3%), multiracial (3%), or Hispanic (2%). All participants were proficient in English, and their musical experience ranged from zero to 18 years ($M = 5$ years, $SD = 4.38$).

Tones

Three tones (1200 Hz; 1400 Hz; 1800 Hz) were created using a tuning function from *Logic Pro X* (Apple Inc 2013). Each tone lasted 1 min, which was determined during pilot testing to be an adequate duration to allow participants to complete the stair walking task. Each tone contained a fixed number of irregular and intermittent pulses (i.e., 1200 Hz had 25 pulses, 1400 Hz had 17 pulses, and 1800 Hz had 16 pulses). The placement of the pulses in each tone was determined by a random generator.

Procedure

These methods were approved by the Kalamazoo College Institutional Review Board, and all participants signed a written document consenting to participate in the study before beginning. The consent document included information about the aerobic demands of the experiment, and this information was also verbally relayed to participants by the experimenter. The experiment involved a single session located on the staircase of a 4-story academic building. During each trial of the experiment, participants walked up or down two flights of stairs (e.g., 1st floor to 3rd floor) which involved 42 steps. Each flight was divided by a landing at the midpoint between the floors. Each step had a tread width of 12 in and a rise of 8 in. Participants were randomly assigned to begin the experiment on either the top or bottom floor of the building. Those beginning on the top floor were taken there in an elevator and then began the experiment by walking down two flights in the first trial. Those who began on the bottom floor walked up two flights in their first trial.

Once at their starting location, participants were given a pair of Bose SoundTime Headphones Around-Ear Style™ connected to an iPhone. Participants were able to adjust the volume to their comfort level. The task was introduced by playing two reference tones for the participants for 5 s each. The lower of the two reference tones was played first, and participants were told that its frequency was 1000 Hz. The higher of the two tones was then played, and participants were told that its frequency was 2000 Hz. Participants were told that the pitch of all tones they heard in the experiment would fall somewhere between these two pitches and that, for each tone, they should report their best estimate of its pitch in exact Hz (e.g., 1250 Hz; 1872 Hz). Further, participants were told that, unlike the reference tones, the tones in the experiment would be broken up with intermittent pulses and they would be asked to report the number of pulses they heard in each tone in addition to its pitch at the end of each trial.

Participants then began walking either up or down the stairs at their own pace. The experimenter walked slightly behind them and controlled the iPhone connected to the participants' headphones. As soon as the participants began their ascent or descent of the stairs, the experimenter began playing one of the three tones, as determined by a predetermined counterbalanced order.

When the participant reached the intended destination (i.e., top of the stairs or bottom of the stairs), the participant removed their headphones, judged the Hz of the tone, and reported the number of pulses they heard. Participants then put the headphones back on and listened to 30 s of an excerpt of classical music (Bach 1720) to clear the previous tone from their short-term memory. After 30 s, the music was turned off and participants were instructed to begin walking the opposite direction for the next trial. In this way, trials alternated between walking up and down the stairs, with each participant completing six trials (three walking up and three walking down). Each of the three tones was played twice for each participant, once on an ascending trial and once on a descending trial, and each tone was played once in the first set of three trials and once in the second set of three trials. Participants were debriefed and thanked for their participation.

Results and discussion

Our hypothesis was that participants would judge a pitch as higher when they heard it while walking up the stairs and judge the same pitch as lower when they heard it while walking down the stairs. The average estimated pitch (Hz) across the three tones is shown in Fig. 1. As predicted, participants judged pitches as higher when they walked up the stairs and judged pitches as lower when they walked down the stairs. The data were analyzed with a 2 (Direction: Up vs. Down) × 3 (Tone: 1200 Hz vs. 1400 Hz vs. 1800 Hz) within-subjects repeated measures factorial analysis of variance (ANOVA) with tone and direction as repeated measures.

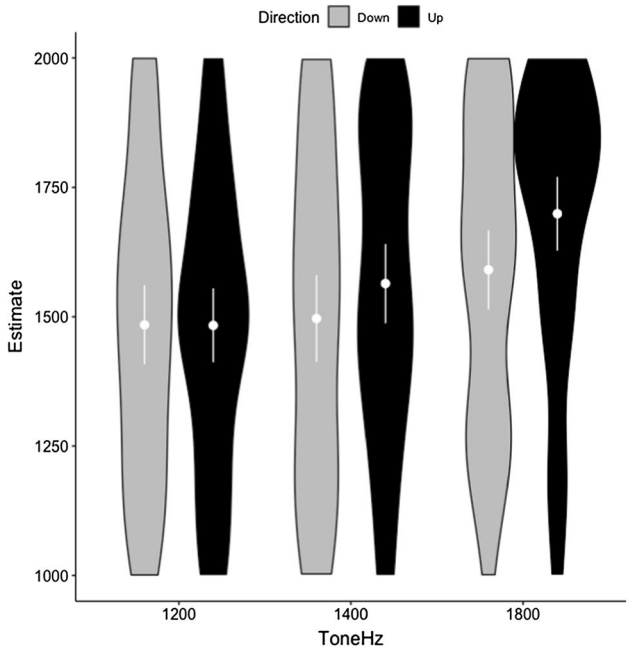

Fig. 1 The estimated pitch (in Hz) of each tone as participants walked up and down the stairs in Experiment 1. White dots and lines represent means and 95% confidence intervals for estimates in each condition. As noted in the text, there is a main effect of direction, such that overall participants made higher estimates when they walked up the stairs (black violins) than when they walked down the stairs (gray violins)

There was a significant main effect of direction, $F(1, 59) = 5.36$, $p = .024$, $\eta_p^2 = .083$. Participants estimated pitches as being higher in pitch (i.e., they reported a greater number of Hz) when the pitch had been heard while walking up the stairs ($M = 1585.11$ Hz, SD = 163.98) than when it had been heard while walking down the stairs ($M = 1527.81$ Hz, SD = 166.02). This effect of direction did not interact with tone, $F(2, 118) = 1.04$, $p = .356$, $\eta_p^2 = .017$.

There was a significant main effect of tone, $F(2118) = 8.42$, $p < .001$, $\eta_p^2 = .125$, suggesting that participants' estimates of the tones' pitches were affected by the actual pitch of each tone. Participants judged the 1800 Hz tone ($M = 1644.88$ Hz, SD = 204.50) as significantly higher in pitch than the 1400 Hz tone ($M = 1540.67$ Hz, SD = 247.21), paired-$t(59) = 2.66$, $p = .01$, $d = -.345$, or the 1200 Hz tone ($M = 1483.83$ Hz, SD = 214.95), paired-$t(59) = 3.87$, $p < .001$, $d = .500$. Participants did not reliably differentiate the pitch of the 1200 Hz and 1400 Hz tones, $t(59) = 1.47$, $p = .146$, $d = .191$. These conclusions regarding statistical significance hold both with a traditional alpha of .05 as well as with a Bonferroni-adjusted alpha (.05/3 = .016).

We also explored whether there were differences in the effect of stair direction on pitch estimation across musicians versus non-musicians. Some previous work has suggested that the correspondence between vertical space and musical pitch is stronger for those with musical experience than for those without musical experience (e.g., Eitan and Granot 2006). The median number of years of musical experience reported by this sample was 4 years, with 34 participants reporting 4 or fewer years of musical experience and 26 reporting more than 4. We did a median split on the sample and entered the factor into a 2 (Musical Experience: 4 or fewer years vs. more than 4 years) × 2 (Stair Direction: Up vs. Down) × 3 (Tone) mixed analysis of variance. There were no main effects or interactions involving musical experience, all F's < 3. The bias in pitch estimations that came from walking in a particular direction was not stronger for those with musical experience compared to those without.

The results support the hypothesis that the kinesthetic experience of walking up or down the stairs would bias people's judgments of auditory pitch in the direction they were moving. However, there are several confounds in the present experiment that prevent us from concluding with certainty that the observed effect on pitch estimates is due to kinesthetic experiences. First, in Experiment 1, kinesthetic experience was confounded with visual experience. As participants walked up the stairs, the experience of getting higher was also apparent in their visual field. When participants made their judgments at the top or bottom of the staircase, there were visual cues in the environment that they were located high or low in space. When at the top of the stairs, they could see the railing and drop over the stairwell. The bias in pitch perception could have been due to visual cues of being high versus low, as has been shown previously (e.g., Dolscheid et al. 2013).

Second, it is unclear whether participants in the present study were aware of the hypothesized relation between pitch estimates and direction of stair walking. Although they were given a generic cover story about the effects of exercise on memory, we did not ask them whether they believed this story or what they thought the hypothesis of the study was. Thus, it is possible that the effect we observed was a demand characteristic of participants' guessing the hypothesis of the study. We conducted Experiment 2 to eliminate these confounds and to examine the effect under a different manipulation of kinesthetic experience.

Experiment 2

In Experiment 2, participants adopted either an upward or downward stretching posture as they listened to the tones (see Fig. 2). In addition to determining whether the effect observed in Experiment 1 would replicate under a different manipulation of kinesthetic experience, we also aimed to exclude the possibility that the effect observed in Experiment 1 was due to visual experience or demand characteristics

Fig. 2 The down (left) and up (right) postures adopted by participants in Experiments 2 and 3

related to the participants guessing the hypothesis. To eliminate the possibility that visual experience of moving up and down created the effect, we blindfolded half of the participants in Experiment 2 as they adopted their stretching postures. If pitch estimation is biased by kinesthetic information, pitch should be equally biased by bodily position regardless of whether participants are blindfolded. Further, to exclude the possibility that participants responded as they did because they guessed the hypothesis of the study, we administered a post-study questionnaire to identify (and exclude) participants who had guessed the hypothesis of the study. If the kinesthetic experience of stretching biases pitch estimation in the direction of the reach, then even participants who cannot accurately state the hypothesis should judge the tones as higher when they have heard the tone while adopting an upward stretching posture than while adopting a downward stretching posture. We again included the pulse counting task, as in Experiment 1, in order to occupy participants' verbal working memory and prevent them from self-labeling the pitch of each tone immediately after hearing it.

Method

Participants

A convenience sample of 110 participants was recruited and compensated with a piece of candy. Before analysis, we excluded data from eight participants who indicated on the post-study questionnaire that they thought the hypothesis of the study was that direction of stretching would affect pitch estimation and data from three participants who reported a hearing impairment. To assure that this was a unique, non-overlapping sample with Experiment 1, we also excluded data from five participants who indicated that they had completed Experiment 1. The final sample thus consisted of 94 participants (51 female; 42 male; 1 who did not identify as either) with an average age of 22.5 years (SD = 9.78). The majority of the sample identified as Caucasian (81%), with some participants identifying as African American (5%), Asian (5%), Hispanic (5%), Middle Eastern (1%), or other (2%). The majority of the sample (81%) reported having some musical experience ($M = 4.19$ years, SD = 4.03). All participants reported speaking English, with 84% being monolingual. The remaining 16% reported knowing a range of other languages in addition to English, including Spanish, French, Korean, Russian, and Hindi.

Procedure

The methods were approved by the Kalamazoo College Institutional Review Board, and all participants signed a written consent form describing the study before beginning. Participants arrived individually to a quiet space on campus. The experimenter informed them that the study involved the relationship between exercise and memory and that they would be asked to adopt various stretching postures while listening to musical tones. The experimenter demonstrated the two stretching positions that the participant would perform (see Fig. 2). The upward stretching position entailed standing with one's legs shoulder-width apart and extending the arms straight above the head. The experimenter specified that in this position, the participant's feet should remain flat on the ground (i.e., not be on their toes). The downward stretching position involved participants standing with their legs shoulder-width apart and bending over to attempt to touch their toes as best as they could.

After demonstrating these two positions, participants were randomly assigned to either the blindfold or no blindfold condition and to begin the experiment stretching upwards or downwards. Participants in the blindfold condition ($n = 46$) were given a black eye mask with elastic straps and asked to place it over their eyes in a comfortable position. Participants in the no blindfold condition ($n = 48$) were instructed to keep their eyes open throughout the study.

All participants were then given a pair of Bose SoundTime Headphones Around-Ear Style™ connected to an iPhone. As in Experiment 1, the task was introduced by playing the two reference tones for 5 s each and labeling each with its exact Hz (1000 Hz for the lower tone and 2000 Hz for the higher tone). Participants were told that all tones in the experiment would fall somewhere between these two pitches, and they should report their best estimate of the tone in exact Hz when asked (e.g., 1250 Hz; 1872 Hz). Further,

to distract participants from over-focusing on the pitch of the tones and from making their judgment immediately after hearing it, participants were again asked to count the number of intermittent pulses they heard in each tone. The first 30 s of the three tones from Experiment 1 were used in this study; we reduced the duration of the trials from 60 to 30 s because 60 s seemed like an uncomfortably long time to have to maintain the stretching postures. The 1200 Hz tone contained 16 pulses, the 1400 Hz tone contained 13 pulses, and the 1800 Hz tone contained 10 pulses.

After indicating that they understood the instructions, participants adopted either the upward or downward stretching posture for the first trial, as instructed by the experimenter. The experimenter stood next to the participant and began playing the first tone as soon as the participants began to stretch. The order of the tones was counterbalanced, such that each of the three tones occurred equally often in each of the six trials across all participants, but for each participant, each tone occurred once on an upward stretch trial and once on a downward stretch trial. The participant held the stretching posture for 30 s, while the tone played. After 30 s, the participant came back to a neutral stance (i.e., standing) and removed their headphones. Participants in the blindfold condition kept their blindfold on. The experimenter then asked participants to report both their estimate of the Hz of the tone and the number of pulses they heard. The order of the two questions was counterbalanced across participants. After reporting these two measures, participants put the headphones back on and listened to 30 s of an excerpt of classical music (Bach 1720) to clear the previous tone from their short-term memory. After 30 s, the music was turned off and participants were instructed to adopt the opposite stretching posture for the next trial. In this way, trials alternated between stretching upwards and downwards, with each participant completing six trials (three stretching upwards and three stretching downwards).

After the sixth trial, participants completed a questionnaire that included questions about their musical experience, languages spoken, hearing impairments, and whether they had participated in any similar study in the past. Further, they were asked to write a short description of what they thought the hypothesis of the study was. Participants who stated that they thought direction of stretching was supposed to affect estimated pitch were excluded from analysis ($n = 8$).

Results and discussion

Participants' judgments of the pitch of each tone (in Hz) were analyzed in a 3 (Tone: 1200, 1400, 1800) × 2 (Stretching Direction: up vs. down) × 2 (Blindfold: yes vs. no) mixed factorial ANOVA with tone and stretching direction as within-subjects factors. As shown in Fig. 3, there were main effects of both tone, $F(2, 184) = 33.96$, $p < .001$, $\eta_p^2 = .27$, and stretching direction, $F(1, 92) = 6.46$, $p = .013$, $\eta_p^2 = .066$, but no main effect of blindfold and no significant 2- or 3-way interactions, all F's < 1.0. Paralleling the effect observed in Experiment 1, participants' estimates of the frequency of all three tones were higher when made in an upward stretching posture than when made in a downward stretching posture, suggesting that participants' estimates were biased by their bodily position. This effect was not dependent on participants' ability to see their surroundings, as there was no interaction with the presence of the blindfold, $F(1, 92) = 0.61$, $p = .44$, $\eta_p^2 = .007$. It appears that the kinesthetic experience of stretching upwards or downwards biased estimates of heard pitch in the same direction that participants were stretching.

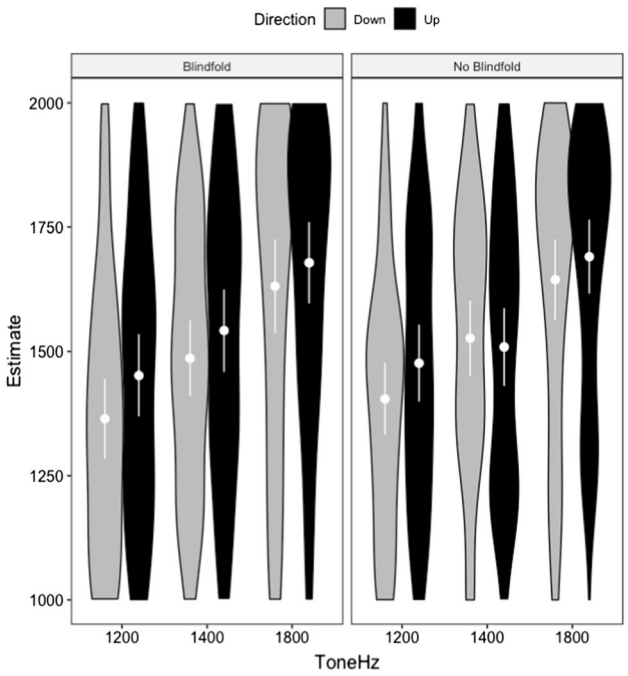

Fig. 3 The estimated Hz of each tone as participants who wore a blindfold (left) or who did not wear a blindfold (right) adopted the up and down stretching postures in Experiment 2. White dots and lines represent means and 95% confidence intervals for estimates in each condition. There is a main effect of direction, such that overall participants made higher estimates when in an upward stretch (black violins) than when in a downward stretch (gray violins)

As in Experiment 1, participants' judgments of the tones' pitch approximated the actual differences in the pitch of the tones, with participants estimating the pitch of the 1200 Hz tone as significantly lower ($M = 1424.62$ Hz, SD = 204.44) than the frequency of either the 1400 Hz tone ($M = 1515.66$ Hz, SD = 195.70) or the 1800 Hz tone ($M = 1660.86$ Hz, SD = 237.01), all pairwise t tests significant at $p = .001$. Thus, participants' pitch estimations were the product of the actual pitch of the tone they heard as well as the posture they were in as they heard it.

We again examined whether the effect of direction on pitch estimates might be affected by musical experience. The median years of reported musical experience in this sample was 3 years, with 50 participants reporting 3 years or less of musical experience and 44 reporting more than 3 years of musical experience. We analyzed the effect of musical experience in a 2 (Musical Experience: 3 years or less vs. more than 3 years) × 2 (Stretching Direction: up vs. down) × 3 (Tone) mixed ANOVA. There were no significant main effects or interactions involving musical experience, all F's < 3.

In summary, adopting an upward stretching posture resulted in higher pitch estimates of the tones' frequency than adopting a downward stretching posture. This aligns with the finding in Experiment 1. Further, in Experiment 2 we find that this effect does not depend on whether participants were also experiencing visual vertical cues, as the effect was not diminished for participants wearing a blindfold, and did not depend on musical experience.

Although the findings of Experiments 1 and 2 are consistent using two different manipulations of kinesthetic experience (stair climbing and stretching), there remains the possibility that the effect in both studies is due to the numerical response method used. Specifically, participants in the first two experiments estimated the pitch they heard by reporting a number between 1000 and 2000, and this response method may have introduced a confound. Because magnitude is often conceptualized in vertical terms (e.g., numbers that are greater in magnitude are "high" and numbers that are lesser in magnitude are "low" (see Ito and Hatta 2004), it is possible that an upward kinesthetic experience biased participants, not in how they heard the pitch, but in their conversion of that pitch to a numerical scale. For example, when asked to report their estimate of the tone's pitch, perhaps participants recalled the exact pitch of the tone they heard, but converted it to a larger number when they had heard the pitch in an upward posture than when they had heard the pitch in a downward posture. To eliminate this confound, we conducted Experiment 3, which did not require participants to make a numerical estimation of the pitch they heard.

Experiment 3

Experiment 3 extended the findings of Experiments 1 and 2 by eliminating the need for a numerical judgment of pitch. Participants in Experiment 3 estimated the pitch of each tone on a horizontal analog scale. Two sample tones were introduced as x and y before the experiment began, and participants marked their estimation for each tone on a horizontal line with endpoints labeled x and y. We measured the physical distance between the endpoint on the scale indicating the lower reference tone and the participant's mark. In this way, we derived a quantitative estimate of the pitch perceived for each tone, but the participants were not making this numerical conversion themselves. This should exclude the possibility that kinesthetic experience biased participants' numerical ratings of each pitch, rather than their perception of the pitch itself. If kinesthetic experience of stretching in a particular direction biases pitch perception in the same direction, then we should see the same effect that emerged in the first two experiments, with participants indicating a higher pitch when tones are heard in an upward posture than in a downward posture.

Method

Participants

Data were collected in three Psychology courses with 20–30 students present in each class. After excluding participants who guessed the hypothesis of the study ($n = 3$), those who skipped the question asking them to make a guess about the hypothesis ($n = 2$), and those who failed to complete the response scale for the tones correctly ($n = 2$), the final sample included 62 participants (37 women; 25 men) with an average age of 19.18 years (SD = 1.25). The majority of the sample (74%) were monolingual English speakers, although some reported also speaking Spanish (13%), or one of a variety of other languages (including Chinese, Japanese, French, Amharie, and Hindi). Most participants identified as Caucasian (66%), with Hispanic/Latinx (13%), Black/African American (10%), Asian (8%), and Middle Eastern (1%) students also represented.

We asked for more detailed information about participants' musical experience in Experiment 3 than in the previous experiments, including number of instruments (including voice) for which they had received formal training and the number of years of training they had received for each instrument. Participants reported a range of cumulative years (0–50) of musical experience across all instruments they knew, and most participants (80%) reported at least one year of formal musical training ($M = 7.92$ years, SD = 9.00) and reported formal training on at least one musical instrument ($M = 1.89$, SD = 1.50). We also asked participants in Experiment 3 if they had ever been told that they had perfect pitch; all participants said no.

Tones

We used the tuning function from *Logic Pro X* (Apple Inc 2013) to create three new tones (600 Hz, 800 Hz, and 1000 Hz) and two new reference tones (500 Hz and 1500 Hz) for use in Experiment 3. We decided to use tones with a lower pitch in Experiment 3 than in the previous experiments because several participants in the first two

experiments commented that the pitch of some of the tones was uncomfortably high. Each tone lasted 30 s and contained between 50 and 70 irregular and intermittent pulses. We inserted more pulses into the tones used in Experiment 3 in order to make the pulses more difficult for participants to count, thereby occupying more of their attention and further reducing the possibilities that they would make their pitch estimation immediately after hearing the tone or that they would verbally label the tone to themselves.

Procedure

Following a lesson about experimental methodology during the first week of each course, the experimenter entered the classroom and invited all students present to participate in a study about the effects of exercise on memory. After signing a consent form, participants were given a response booklet that included an assignment to Group A or Group B on the front as well as instructions not to open the booklet until instructed. The experimenter then demonstrated the two postures that participants would adopt as they heard the tones in the study. Those in Group A were asked to adopt the upward stretching posture for the first tone (see right panel of Fig. 2; those in Group B were asked to adopt the downward stretching posture for the first tone (see left panel of Fig. 2). The two groups then alternated postures for each tone that was played, so that across participants, the order of stretching direction was counterbalanced for each tone.

After explaining the two postures that would be adopted during the experiment, the experimenter introduced the rating scale. A horizontal line was projected on a white screen at the front of the room with the endpoints labeled *X* (on the left) and *Y* (on the right). The experimenter instructed participants to think of the *X* tone "as this" and played the first reference tone. For two of the three classes, this was the lower (500 Hz) reference tone, and for the other class it was the higher (1500 Hz) reference tone. The participants were then instructed to think of the *Y* tone "as this," and the other reference tone was played. Note that the exact Hz of these two reference tones was not given to participants. Participants were then told that for the actual tones in the experiment, they could mark the line anywhere between *X* and *Y* that they thought was appropriate. If they thought the tone was identical to the *X* reference tone, they should mark the far-left end of the line. If they thought the tone was identical to the *Y* reference tone, they could mark the far-right end of the line. They should place their mark anywhere along the line between *X* and *Y* they thought was most appropriate to indicate the pitch of each tone.

Similar instructions were given for the pulse ratings. A horizontal line with the far-left end marked 1 and the far-right end marked 100 was projected on the screen. Participants were told that the tones they heard in the study would have a series of pulses that they should try to count as they listened to each one. Each tone would contain somewhere between 0 and 100 pulses, and participants should mark the line to indicate their best guess for the number contained in each.

After receiving the instructions, participants were asked to stand up and adopt their starting posture (Group A stretching up; Group B stretching down). The first tone was then played for 30 s. When it was finished, the experimenter instructed participants to relax, and complete their ratings for the first tone in their response booklets. The response booklet included one page for each of the six tones that participants heard in the experiment. On each page, there were two questions, one instructing participants to rate the pitch of the tone and one instructing them to indicate the number of pulses heard in the tone. The pitch question was worded as follows: "Please place a mark on the line below to indicate where you think the tone you just heard would fall in terms of its similarity to the two reference tones you heard at the beginning of the study." Below this sentence, there was a 175 mm horizontal line with *X* at the far-left end and Y at the far-right end. The pulse question was worded as: "The tone contained a number of pulses or breaks in the sound. Please estimate how many pulses you think there were in the tone by placing a mark on the line below." Beneath these instructions, there was a second 175 mm horizontal line labeled with 1 at the far-left and 100 at the far-right. The order of these two questions on each page was counterbalanced across the answer booklets.

After completing their ratings for the first tone, a 30-s excerpt of classical music (Bach 1720) was played to clear participants' memory of the previous tone. Participants were then asked to stand and adopt the next posture (stretching down for Group A; stretching up for Group B). The second tone was played for 30 s, and participants were then instructed to relax and make their ratings for Tone 2 in their response booklet. This procedure repeated six times, until all three tones had each been played twice. The order of tones was counterbalanced across the three classes, so that each tone was played in each ordinal position for approximately one-third of participants, and all participants heard each tone twice—once while stretching up and once while stretching down.

Following the sixth tone, the participants turned to the final pages of their response booklet and answered questions about their demographics, languages spoken, musical experience, history of hearing impairment, perfect pitch, suspected hypothesis of the experiment, and prior experience participating in similar experiments. The experimenter collected the response booklets and debriefed the participants regarding the true hypothesis of the study. The students' instructor then used the experiment as a starting point for

a discussion about experimental methods. These methods were approved by the Kalamazoo College IRB.

Data coding

A coder who was unaware of which tone each participant had been listening to or which posture they were in examined the mark made on each line for each tone. The coder used a ruler to measure the distance between the participant's mark and the end point indicating the lower tone to the nearest mm. For participants in two of the classes, the low end was the left end marked *X*, and for participants in the third class, the low end was the right end marked *Y*. Scores could range from 0 mm (the participant marked the endpoint of the line) to 175 mm (the participant marked the opposite end of the line), with a shorter distance indicating a lower estimated pitch than a longer distance. Two participants always marked the *X* or *Y* endpoint, rather than placing their mark along the line. We assumed that these participants did not understand the response system and excluded their data from analysis.

Results and discussion

Figure 4 shows the average distance from the low reference point indicated by participants for each tone while in an upward and downward stretch. Data were analyzed in a 3 (Tone: 600 Hz vs. 800 Hz vs. 1000 Hz) × 2 (Stretching Direction: Up vs. Down) repeated measures ANOVA. As expected, there was a main effect of tone, $F(2, 122) = 133.40$, $p < .001$, $\eta_p^2 = .69$, with participants marking the line further away from the low reference point as the pitch of the tone increased. More importantly, there was also a main effect of posture, $F(1, 61) = 11.83$, $p = .001$, $\eta_p^2 = .16$, replicating the effect found in the first two experiments. Hearing a tone while in an upwards stretch led participants to mark the line further from the low reference point than hearing the same tone while in a downwards stretch. There was no tone x direction interaction, $F(2, 122) = 1.82$, $p = .17$, $\eta_p^2 = .03$.

As in the first two experiments, we explored whether the effect of direction on tone might be different for those with and without musical experience. The median cumulative years of reported musical experience in this sample was 6 years, and a median split yielded 33 participants with 6 or fewer years of cumulative musical experience and 29 participants with more than 6 years. Note that the median years of experience is higher in Experiment 3 than in Experiments 1 or 2 because we asked the question differently; instead of simply asking about total years of experience, in Experiment 3 we asked about cumulative years of experience, such that a single year playing two different instruments counted as two cumulative years. We conducted a 2 (Musical Experience: 6 or fewer vs. more than 6 years) × 2 (Stretching Direction: Up vs. Down) × 3 (Tone) mixed ANOVA. There was

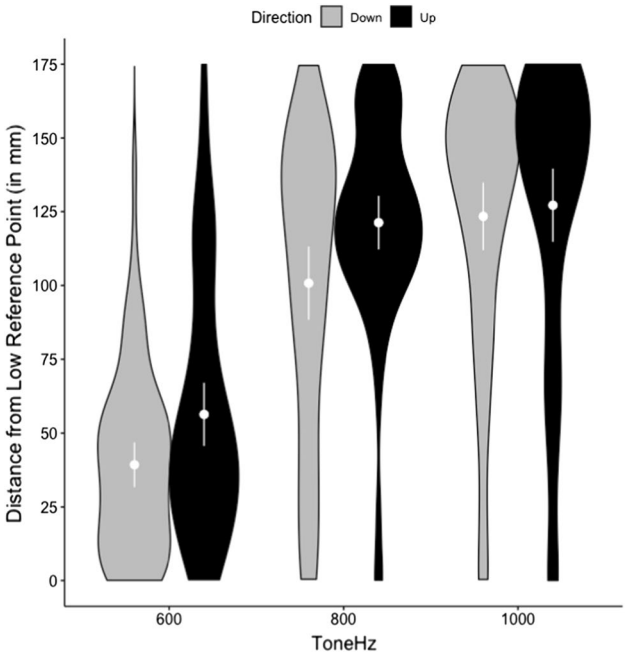

Fig. 4 The average distance from the low reference point (in mm) that participants marked the line to indicate the pitch of each tone while in the up and down stretching postures in Experiment 3. White dots and lines represent means and 95% confidence intervals for estimates in each condition. There is a main effect of direction, such that participants made higher estimates when in an upward stretch (black violins) than when in a downward stretch (gray violins)

a three-way interaction between tone, direction, and musical experience, $F(2, 120) = 6.49$, $p = .002$, $\eta_p^2 = .098$. To understand this interaction, we conducted 6 paired-samples t tests comparing judgments of each tone in the up versus down stretch for participants with low musical experience and for participants with high musical experience. We used Bonferroni-corrected alpha for multiple comparisons of .008 (.05/6). Participants with 6 or fewer years of musical experience showed an effect of reaching direction for the 800 Hz tone, paired-$t(32) = 3.35$, $p = .002$ only, whereas participants with more than 6 years of musical experience showed an effect of reaching direction for the 600 Hz tone only, paired-$t(28) = 2.94$, $p = .006$. In both cases, the effect is such that estimates were further from the low reference tone when participants adopted the upward stretch than when they adopted the downward stretch. It is unclear why this complex interaction emerged, but we think it is likely to be an artifact of this particular sample, given that there was no evidence of it in the previous two experiments. Importantly, the direction and size of the effect of stretching direction on pitch estimation is not changing with musical experience.

As in the first two experiments, participants' estimations of pitch were biased by the position their body was in as they heard the pitch. When participants were stretching upward, they reported hearing the tone as having a higher pitch than

when they were stretching downward. Importantly, this is not the result of participants assigning a higher numerical value to their pitch estimation when they were in an upward posture, because in Experiment 3, participants did not estimate the pitch with a numerical rating. Rather, participants indicated their best estimate of the pitch by marking a horizontal line—a response which should not be, by itself, biased by the participant's posture. It appears that participants' perception of the pitch they had heard was higher when they had heard it in an upward posture, and this higher pitch was translated to the rating scale by marking the line further away from the low reference point.

A notable limitation of Experiment 3 is that we administered the task in a group setting rather than individually. This raises the possibility that participants were affected by other participants in the room. However, it is important to note that participants reported their estimates privately in their individual answer booklets and were not aware of the estimates given by other participants in the room. Further, we counterbalanced the order of the two stretching positions within each group; for each tone, half the participants in the room adopted the upward posture and half the downward posture. As a result, we would argue that group administration may have weakened the effect because participants' ability to see other participants adopting the opposite posture from their own might have interfered with the effect of their own bodily position. In sum, we doubt that group administration of the task in Experiment 3 produced the effect we observed, particularly when considered in light of the effects observed in the first two experiments where the task was administered individually.

General discussion

Across three experiments, we found evidence that participants' kinesthetic position affected their estimates of auditory pitch. Specifically, engaging in an ascending movement resulted in perceiving a pitch as higher than engaging in a descending movement. We replicated this basic effect over a variety of conditions and situations: with two different types of movement (walking up/down stairs in Experiment 1 vs. stretching up/down in Experiments 2 and 3), with visual experience available and blocked (Experiment 2), with a numeric (Experiments 1 and 2) and non-numeric rating scale (Experiment 3), with two different sets of auditory tones (Experiments 1 and 2 vs. 3), and with individual and group administration (Experiments 1 and 2 vs. 3). We also found in Experiments 2 and 3 that the effect did not differ across those with more versus less formal musical training. Further, the effect is unlikely to be due to participants guessing the hypothesis of the study, as fewer than 7% of participants in Experiments 2 and 3 were able to state the hypothesis explicitly when asked, and the finding emerged without those participants included in the analysis. Thus, this appears to be a robust effect that emerges across a variety of conditions and stimuli.

There are several mechanisms by which movement in a particular direction could affect perceived pitch. First, such an effect could occur through verbal labeling; because both pitch and bodily position can be described in English with a shared term (e.g., "high" vs. "low"), the verbal label could mediate the relationship between the two. For example, perhaps participants self-labeled their position as "high" versus "low" and the activation of this label was then applied to the tone they heard. Indeed, some research has shown that the cross-modal correspondence between auditory pitch and vertical space is weaker for speakers of languages that do not describe pitch in vertical terms (e.g., Dolscheid et al. 2013; Fernández-Prieto et al. 2017). We doubt that this explanation can account for the results in the present study because the simultaneous pulse counting task that participants were engaged in should have occupied their verbal working memory and prevented them from self-labeling their position or the tones to themselves. However, our participants were speakers of English, and it is unclear whether the effect would be as strong for speakers of another language that does not use a shared label for pitch and vertical position.

A second possibility is that the effect is mediated by emotion or arousal. Previous research has shown that positive versus negative emotional memories can be primed by bodily movement in a particular direction, with upward movement priming happier memories than downward movement (Casasanto and Dijkstra 2010). In addition, music with a higher pitch is associated with increased pleasantness ratings than music with a lower pitch (e.g., Jaquet et al. 2012), raising the possibility that upward movement may lead to more positive affect, which then biases the pitch ratings in an upward direction. Similarly, it has been shown that reported arousal is greater with lower pitches than higher pitches (e.g., Jaquet et al. 2012), leaving open the possibility that the downward movements used here led to more arousal, which biased the pitches downward, compared to the upward movements. Although we cannot rule this possibility out, it does seem questionable whether the downward movements used here would have been *more* arousing than the upward movements, particularly in Experiment 1 where walking up the stairs was certainly more physiologically demanding than walking down the stairs. Future research could more carefully examine these possibilities, for example by measuring participants' emotional and physiological arousal in each position.

A third possible mechanism for the effect is that kinesthetic experience of moving in a particular direction activated the concept of vertical space, which was then metaphorically applied to the pitch of the tones. According to

embodied accounts of cognition (see Glenberg et al. 2013; Körner et al. 2015; Wilson 2002), human understanding of abstract ideas is enabled by metaphorical comparison with the more concrete, physical world (Lakoff and Johnson 1980). Under this view, abstract ideas that do not inherently have physical properties come to be thought about and described as possessing physical properties. Because speakers of English have spent a lifetime describing pitch in vertical terms, they have come to think about pitch as a vertical construct, and this conceptual metaphor can be activated through kinesthetic experience. Participants' sensorimotor experiences of moving up or down activated the concept of *high* or *low*, which they then applied to the pitch judgments. This would be similar to other modal priming effects, in which a particular sensorimotor experience affects judgments of a concurrently presented stimulus. For example, holding a heavy book increases judgments of its importance (e.g., Chandler et al. 2012). Such effects likely occur because the bodily experience has made the associated concept more accessible and more likely to be used in the rating task (see Körner et al. 2015); in the present paradigm, moving up made the concept of *high* readily accessible during the pitch judgment task, and biased the judgment in the direction of the activated concept.

Some limitations of the present set of experiments are that we have not established the timescale of the effect, how much bodily experience in a particular direction is necessary for the priming to occur, or whether the cross-modal priming that occurs is relative or absolute (see Brunetti et al. 2018; Spence 2011). We tried to prevent participants from making their judgments immediately after hearing the tones by giving them the extraneous task of counting pulses. However, we do not know how effective this was at preventing participants from making their judgments before they were asked to report them. It is possible that participants made a judgment immediately upon hearing each tone as they first began the stair climb (Experiment 1) or stretch (Experiments 2 and 3) and then just held that judgment in mind until they were asked to report it. However, if participants actually did settle on their judgment early in each trial, it is even more surprising that the effects emerged as they did, and it would suggest that bodily position primes the concept of space quite readily. Future research could investigate how much bodily experience in a particular direction is needed to bias pitch estimation. Similarly, it would be interesting in future work to design a means of examining this question with a more indirect task, in which participants are not asked to explicitly rate the pitch of the tones they heard. For example, engaging participants in a speeded categorization task where each tone must be categorized as either low or high may shed light on whether the correspondence is absolute or relative to the other pitches and positions used in the experiment (see Brunetti et al. 2018).

In conclusion, although previous research has shown that perception of high versus low tones primes participants to move up versus down in space (e.g., Mossbridge et al. 2011; Rusconi et al. 2006), this is the first demonstration that the effect also works in reverse; moving one's body in a particular way affects the perception of a pitch that is heard. This finding suggests that the cross-modal priming effect between perception of auditory pitch and vertical movement is bidirectional, as has been shown for the cross-modal effect between auditory pitch and visual orientation to a particular direction (e.g., Dolscheid et al. 2013; Evans and Treisman 2010). Further, it has been suggested that bidirectionality of this type is an important characteristic of cross-modal priming effects that rely on conceptual metaphors as their mechanism (Körner et al. 2015). It appears that when judging the pitch of a tone, our ears are not the only part of our body that matters.

Acknowledgements We thank Ethel Mogilevsky for her assistance with data collection and Chelsea Miller for her assistance with data coding. We also thank Siu-Lan Tan for her advice regarding the design of Experiment 3.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest regarding the results of these studies.

Ethical standards All procedures performed in the studies described were in accordance with ethical standards of the Institutional Review Board at Kalamazoo College and with the 1964 Declaration of Helsinki and its later amendments.

Informed consent Informed consent was obtained from all individuals who participated in the study, and there is no identifying information included about any individual participant.

References

Apple Inc (2013) Logic Pro X (Computer software). Retrieved from https://www.apple.com/logic-pro/

Bach JS (1720) Unaccompanied cello suite no. 1 in G Major, BWV 1007: Prélude [Recorded by Yo-Yo Ma]. On *The Six Unaccompanied Cello Suites* [MP3 file]. Sony Classical (1983)

Bonetti L, Costa M (2017) Pitch-verticality and pitch-size cross-modal interactions. Psychol Music 46:340–356. https://doi.org/10.1177/0305735617710734

Brunetti R, Indraccolo A, del Gatto C, Spence C, Santangelo V (2018) Are crossmodal correspondences relative or absolute? Sequential effects on speeded classification. Atten Percept Psychophys 80:527–534. https://doi.org/10.3758/s13414-017-1445-z

Casasanto D, Dijkstra K (2010) Motor action and emotional memory. Cognition 115:179–185. https://doi.org/10.1016/j.cognition.2009.11.002

Chandler JJ, Reinhard D, Schwarz N (2012) To judge a book by its weight you need to know its content: knowledge moderates the use of embodied cues. J Exp Soc Psychol 48:948–952. https://doi.org/10.1016/j.jesp.2012.03.003

Connell L, Cai ZG, Holler J (2013) Do you see what I'm singing? Visuospatial movement biases pitch perception. Brain Cognit 81:124–130. https://doi.org/10.1016/j.bandc.2012.09.005

Dolscheid S, Casasanto D (2015) Spatial congruity effects reveal metaphorical thinking, not polarity correspondence. Front Psychol 6:1–11. https://doi.org/10.3389/fpsyg.2015.01836

Dolscheid S, Shayan S, Majid A, Casasanto D (2013) The thickness of musical pitch: psychophysical evidence for linguistic relativity. Psychol Sci 24:613–621. https://doi.org/10.1177/0956797612457374

Eitan Z, Granot RY (2006) How music moves: musical parameters and listeners' images of motion. Music Percept Interdiscip J 23:221–247. https://doi.org/10.1525/mp.2006.23.3.221

Evans KK, Treisman A (2010) Natural cross-modal mappings between visual and auditory features. J Vis 10:1–12. https://doi.org/10.1167/10.1.6

Fernández-Prieto I, Spence C, Pons F, Navarra J (2017) Does language influence the vertical representation of auditory pitch and loudness? i-Perception. https://doi.org/10.1177/2041669517716183

Glenberg AM, Witt JK, Metcalfe J (2013) From the revolution to embodiment: 25 years of cognitive psychology. Perspect Psychol Sci 8:573–585. https://doi.org/10.1177/1745691613498098

Huron D, Shanahan D (2013) Eyebrow movements and vocal pitch height: evidence consistent with an ethological signal. Acoust Soc Am 133:2947–2952. https://doi.org/10.1121/1.4798801

Huron D, Dahl S, Johnson R (2009) Facial expression and vocal pitch height: evidence of an intermodal association. Empir Musicol Rev 4:93–100

Imbo I, Vandierendonck A (2007) The role of working memory in simple-arithmetic strategies. Eur J Cognit Psycholgy 19:910–933. https://doi.org/10.1080/09541446010515571

Ito Y, Hatta T (2004) Spatial structure of quantitative representation of numbers: evidence from the SNARC effect. Mem Cognit 32:662–673

Jaquet L, Danuser B, Gomez P (2012) Music and felt emotions: how systematic pitch level variations affect the experience of pleasantness and arousal. Psychol Music 42:51–70. https://doi.org/10.1177/0305735612456583

Körner A, Topolinski S, Strack F (2015) Routes to embodiment. Front Psychol 6:940. https://doi.org/10.3389/fpsyg.2015.00940

Lakoff G, Johnson M (1980) The metaphorical structure of the human conceptual system. Cognit Sci 4:195–208. https://doi.org/10.1207/s15516709cog0402_4

Lemaitre G, Scurto H, Françoise J, Bevilacqua F, Houix O, Susini P (2017) Rising tones and rustling noises: metaphors in gestural depictions of sounds. PLoS ONE 12:1–30. https://doi.org/10.1371/journal.pone.0181786

Maeda F, Kanai R, Shimojo S (2004) Changing pitch induced visual motion illusion. Curr Biol 14:R990–R991. https://doi.org/10.1016/j.cub.2004.11.018

Mossbridge JA, Grabowecky M, Suzuki S (2011) Changes in auditory frequency guide visual-spatial attention. Cognition 121:133–139. https://doi.org/10.1016/j.cognition.2011.06.003

Occelli V, Spence C, Zampini M (2009) Compatibility effects between sound frequency and tactile elevation. NeuroReport 20:793–797. https://doi.org/10.1097/WNR.0b013e32832b8069

Pratt CC (1930) The spatial character of high and low tones. J Exp Psychol 13:278–285. https://doi.org/10.1037/h0072651

Roffler SK, Butler RA (1968) Localization of tonal stimuli in the vertical plane. J Acoust Soc Am 43:1260–1266. https://doi.org/10.1121/1.1910977

Rusconi E, Kwan B, Giordano BL, Umiltà C, Butterworth B (2006) Spatial representation of pitch height: the SMARC effect. Cognition 99:113–129. https://doi.org/10.1016/j.cognition.2005.01.004

Sadaghiani S, Maier JX, Noppeney U (2009) Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. J Neurosci 29:6490–6499. https://doi.org/10.1523/JNEUROSCI.5437-08.2009

Salgado-Montejo A, Marmolejo-Ramos F, Alvarado JA, Arboleda JC, Suarez DR, Spence C (2016) Drawing sounds: representing tones and chords spatially. Exp Brain Res 234:3509–3522

Spence C (2011) Crossmodal correspondences: a tutorial review. Atten Percept Psychophys 73:971–995. https://doi.org/10.3758/s13414-010-0073-7

Wagner S, Winner E, Cicchetti D, Gardner H (1981) "Metaphorical" mapping in human infants. Child Dev 52:728–731. https://doi.org/10.1111/j.1467-8624.1981.tb03106.x

Wilson M (2002) Six views of embodied cognition. Psychon Bull Rev 9:625–636. https://doi.org/10.3758/BF03196322