

ALL EYES ON ME: BEHAVING AS SOLOIST IN DUO PERFORMANCES LEADS TO INCREASED BODY MOVEMENTS AND ATTRACTS OBSERVERS' VISUAL ATTENTION

MATS B. KÜSSNER

Humboldt-Universität zu Berlin, Berlin, Germany

EDITH VAN DYCK

Ghent University, Ghent, Belgium

BIRGITTA BURGER

University of Jyväskylä, Jyväskylä, Finland

DIRK MOELANTS, & PIETER VANSTEENKISTE

Ghent University, Ghent, Belgium

DUO MUSICIANS EXHIBIT A BROAD VARIETY OF bodily gestures, but it is unclear how soloists' and accompanists' movements differ and to what extent they attract observers' visual attention. In Experiment 1, seven musical duos' body movements were tracked while they performed two pieces in two different conditions. In a congruent condition, soloist and accompanist behaved according to their expected musical roles; in an incongruent condition, the soloist behaved as accompanist and vice versa. Results revealed that behaving as soloist, regardless of the condition, led to more, smoother, and faster head and shoulder movements over a larger area than behaving as accompanist. Moreover, accompanists in the incongruent condition moved more than soloists in the congruent condition. In Experiment 2, observers watched videos of the duo performances with and without audio, while eye movements were tracked. Observers looked longer at musicians behaving as soloists compared to musicians behaving as accompanists, independent of their respective musical role. This suggests that visual attention was allocated to the most salient visuo-kinematic cues (i.e., expressive bodily gestures) rather than the most salient musical cues (i.e., the solo part). Findings are discussed regarding auditory-motor couplings and theories of motor control as well as auditory-visual integration and attention.

Received: October 8, 2018, accepted July 13, 2020.

Key words: bodily gestures in musical duo performances, motion capture, eye tracking, auditory-motor coupling, audio-visual perception and attention

MUSICIANS DISPLAY A BROAD VARIETY OF expressive bodily gestures, both in solo and ensemble performances (Davidson & Broughton, 2016). The types and extent of these bodily gestures can attract visual attention and influence our perception of a musical performance (e.g., Broughton & Stevens, 2009; Vuoskoski, Thompson, Clarke, & Spence, 2014). In ensemble performances, musicians use bodily gestures to communicate with both co-performers and audience members. However, it is not clear how a musician's role within an ensemble affects their bodily gestures and whether audience members attend to these gestures. This study examines the extent to which soloists' and accompanists' bodily gestures in musical duos differ and how their respective movements may capture observers' visual attention.

Gestures in Music Performance

Bodily gestures play a crucial role in music performance (Dahl & Friberg, 2007; Davidson & Broughton, 2016) and have been functionally divided into sound-producing, communicative, sound-facilitating, and sound-accompanying gestures (Jensenius, Wanderley, Godøy, & Leman, 2010). Communicative gestures constitute an integral part of musical performances and convey important information to co-performers and audiences alike. For example, a head-nod from the leader of an ensemble may function as the signal for a synchronous start (Palmer & Deutsch, 2012), and also draw observers' attention to the stage and raise their receptiveness for attending to performers' expressive bodily gestures. Research has shown that ensemble musicians use different kinds of communicative gestures for various purposes; for instance, they may be used to change to a different tempo (King & Ginsborg, 2011), to highlight particular phrases (Williamson & Davidson, 2002), and to indicate structural boundaries (Davidson & Coulam, 2006; King & Ginsborg, 2011). The types of gestures utilized by musicians vary inter-individually and are constrained by instrument (Davidson, 2012) as well as a musician's role within an ensemble. In general, soloists move first (Keller & Appel,

2010), make more exaggerated movements (Goebel & Palmer, 2009), and choose from a highly idiosyncratic set of gestures (Davidson, 2005), which seems to be relatively stable over time (Davidson, 2007).

Furthermore, bodily gestures synchronize sound as well as actions in musical ensembles. Investigating the influence of visual information and vividness of auditory imagery on synchronization in piano duos, Keller and Appel (2010) showed that the quality of coordination between two pianists playing primo and secondo parts is not so much dependent on visual contact between the two musicians, but more so on internal models able to simulate one's own as well as others' actions during a performance. Bishop and Goebel (2015) found that pianists rely primarily on auditory cues when synchronizing their secondo part with the (pre-recorded) primo part of another pianist or violinist. However, pianists attended more to visual cues when synchronizing parts during moments where auditory information was absent (e.g., after long pauses).

Despite the wide array of different gestures and body parts used in musical performance, Davidson (1994, 2002, 2012) identifies body sway as a core element of communicative and expressive bodily gestures in both solo and ensemble performances. This view is echoed by Chang and colleagues (2017), who showed that randomly assigned leaders in a string quartet affected the body sway of followers most when auditory and visual information were present, highlighting the importance of nonverbal bodily gestures for social communication during ensemble performances. Still, specific bodily regions may be especially important in facilitating communication of musical information and meaning. Head movements can convey a variety of expressive and emotional intentions (Castellano, Mortillaro, Camurri, Volpe, & Scherer, 2008; Thompson & Luck, 2012) and may provide an audience with important cues regarding musical structure (Timmers, Marolt, Camurri, & Volpe, 2006). Glowinski and colleagues (2013) compared violinists playing the same piece of music alone vs. in their usual quartet and found that head movements are more predictable in the ensemble condition, which may be related to the communicative function of head movements in ensemble playing. Observers rating these excerpts were able to distinguish the two kinds of performances (solo vs. ensemble), whereas ratings of emotional content and expressivity showed no difference. However, it is not clear how leadership intention affects the movement style or quantity of motion of specific body parts (e.g., head, shoulders) and whether this embodied intention is influenced by an individual's

musical role within an ensemble (e.g., being a soloist versus an accompanist).

Related research in psychology suggests that observers can detect intended (deceitful) action from kinematic cues. In a classic study, Runeson and Frykholm (1983) showed that, based on preparatory movements conveyed via point-light displays, people can identify whether a person intends to lift a light or heavy box. What is more, individuals are able to detect deceitful preparatory actions, such as a person pretending to lift a heavy box although she knows it is light. This suggests that deceitful actions in music performance, such as overly expressive gestures or incongruent behavior (an accompanist behaving as a soloist), may be detected by concert goers. Still, it remains unclear what kinds of kinematic cues are used when behavioral roles are swapped. Knowing how leaders and followers in musical ensembles use bodily gestures to convey information to co-performers or the audience would provide valuable insights into how more complex auditory-motor couplings necessary to coordinate and communicate with others are generated. The underlying principles of these bodily gestures can be explained with theories of motor control such as optimal feedback control (Todorov & Jordan, 2002). Our motor system is able to achieve goals reliably and repeatedly, while at the same time allowing for highly flexible detailed movements. As Todorov and Jordan (2002) have shown, these two properties do not necessarily need to be in conflict if a dynamical systems view of motor coordination is postulated. In this approach, there is no distinction between trajectory planning and execution, and decisions regarding movement details can be postponed until the last moment. Redundancy (or variability) in the motor system is therefore a crucial condition to perform tasks well and could explain why musicians performing unusual(ly) expressive gestures are still able to achieve a certain musical goal (i.e., produce an adequately expressive sound).

The Impact of Visual and Auditory Cues on Music Perception

Another important question regarding the use of body movement in music ensembles is the extent to which visual and auditory cues interact in the perception of music performance. Recent evidence suggests that visual information constitutes a central part of musical performance from an observer's perspective. After analyzing fifteen studies on the perception of auditory and visual components of music, Platz and Kopiez (2012, p. 75) concluded that "the visual component

is not a marginal phenomenon in music perception, but an important factor in the communication of meaning." Still, whether or not audience members visually attend to musicians' bodily gestures and if so, how the combination of salient auditory and visual stimuli captures their attention are both questions that remain unanswered. In general psychology, there is a substantial amount of literature suggesting that vision takes precedence over audition in various situations where auditory and visual information is present (Bertelson & Aschersleben, 1998; McGurk & MacDonald, 1976). An example of the dominance of visual input during audio-visual processing is the McGurk effect, which demonstrates that visual input from speech gestures (i.e., lip movements) alters auditory perception (McGurk & MacDonald, 1976). However, Alais and Burr (2004) provided evidence that a similar phenomenon, the ventriloquist effect—which is experienced when the perceived location of a sound source deviates from the actual sound source—can be manipulated by increasing or decreasing the clarity of visual stimuli. They showed that the integration of audio-visual stimuli largely depends on the clarity of visual (spatial) information. If the visual localization is adequate, vision dominates or “captures” sound. If the visual stimulus is severely blurred, the ventriloquist effect disappears and sound captures vision.

These robust effects, as demonstrated from related research in cognitive science, have also been applied in performance psychology, where it has been shown that visual input influences the perception of different aspects of musical performances, including rubato (Juchniewicz, 2008), note duration (Schutz & Lipscomb, 2007), tension and phrasing (Vines, Krumhansl, Wanderley, & Levitin, 2006) or emotional expression (Broughton & Stevens, 2009; Vuoskoski et al., 2014). Indeed, expression embedded in performance can be successfully decoded from visual kinematic information (Dahl & Friberg, 2007) and has been shown to influence felt (Chapados & Levitin, 2008) and perceived emotion (Vines, Krumhansl, Wanderley, Dalca, & Levitin, 2011). Vuoskoski et al. (2014) investigated how pianists' body movements and the sound of their performances contribute to the perceived expressiveness of piano performances. Drawing on a classic design by Davidson (1993), they asked pianists to perform Chopin's *Prelude in E minor* with three types of expression: deadpan, normal, and exaggerated. Creating both matching (e.g., exaggerated audio + exaggerated movement) and mismatching (e.g., exaggerated audio + deadpan movement) stimuli, the authors showed that although both auditory and visual kinematic cues contributed to the

perception of overall expressivity, the effects of the visual kinematic cues were more pronounced. Furthermore, Tsay (2013) discovered that musical experts and novices alike were able to successfully identify the winners of international classical music competitions through silent videos. They did not succeed to do so through audio recordings alone and even failed the task when using recordings with both video and sound; Tsay (2014) subsequently provided similar evidence for judging musical ensembles. However, the extent to which these findings are generalizable and applicable to different musical situations is currently debated (Mehr, Scannel, & Winner, 2018). What these findings emphasize is the importance of visual kinematic cues in musical performances and visual dominance in a domain supposedly defined by auditory information.

Visual vs. Auditory Attention

Nevertheless, visual dominance can only occur if visual cues capture attention, or allow for prioritization of specific visual stimuli over others. For instance, humans frequently attend to faces, thereby neglecting other visual stimuli (Bindemann, Burton, Langton, Schweinberger, & Doherty, 2007). Our visual attention is generally shifted by two orienting mechanisms (Connor, Egeth, & Yantis, 2004). On the one hand, top-down attention shifts are controlled endogenously by voluntary, goal-directed behavior. For example, a person may decide to look at a specific musician during an ensemble performance because they are being evaluated in an audition or a competition. On the other hand, bottom-up processes refer to instances where visual attention is captured by the saliency of the visual scene such as color or motion (Koch & Ullman, 1987; Rosenholtz, 1999). Several studies have also highlighted the impact of scene semantics in gaze allocation. For example, Loftus and Mackworth (1978) reported that subjects fixated earlier, more often, and with longer durations when there was an incongruent object in the scene (e.g., an octopus on a farm). Both endogenous and exogenous attention shifts can occur either overtly (i.e., with eye movements) or covertly (i.e., without eye movements, using peripheral vision) and may affect performance differently in a variety of cognitive tasks including temporal order judgments, texture segmentation and visual search tasks (Liu, Abrams, & Carrasco, 2009).

Similar attention-related processes, including shifts, are apparent within the auditory domain. While the auditory scene may consist of many complex sound patterns, neural resources for processing are limited. Auditory saliency maps have been proposed as

a mechanism for allocating auditory attention (Kayser, Petkov, Lippert, & Logothetis, 2005). In this model, different features of sound stimuli (intensity, frequency contrast, temporal contrast) are divided into separate processing streams and analyzed in parallel to detect features that deviate significantly from the background noise. In music performance, salient auditory and visuo-kinematic stimuli compete using either different unisensory pathways or the same saliency map in multisensory areas.

In the absence of a specific task or set of instructions, eye-tracking research revealed that observers generally direct their attention towards the point that is the most visually salient (Hayhoe & Ballard, 2005). However, salient auditory cues such as the solo part in an ensemble performance may also attract visual attention. In an experiment on duo singing performance, the melody part (soprano) was shown to attract more visual attention than the accompaniment part (alto) throughout the piece (Kawase & Obata, 2016). Since singers were instructed to “maintain a stationary position to avoid any significant movements that might attract visual attention” (p. 18), it is unknown whether body movements would have diverted the visual attention from a soloist showing few expressive bodily gestures to an accompanist displaying more expressive bodily gestures. As dynamic factors such as motion have been shown to contribute significantly to the saliency of visual cues (Koch & Ullman, 1987; Rosenholtz, 1999), gesturing and movement are highly likely to attract overt visual attention, possibly taking precedence over auditory cues.

Aims of the Study

The purpose of this study is thus two-fold. First, we aim to investigate how musicians’ roles in a duo performance affect their bodily gestures and second, whether expressive bodily gestures attract observers’ visual attention, leading to longer dwell times. Each duo performed two pieces: one in which performer A played the solo and performer B played the accompaniment, and another piece in which performer B played the solo and performer A the accompaniment. Both pieces were played twice, resulting in four performances per duo: first, each piece was performed in a natural way, with each musician behaving according to their respective roles; then, the same piece was repeated, but with the soloist instructed to behave as if they were playing the accompaniment and the accompanist instructed to behave as if they were playing the solo.

Since it has been shown that more expression leads to increased body movements (Davidson, 1994; Thompson

& Luck, 2012), we expected musicians behaving as soloists to display more extensive head and shoulder movements than musicians behaving as accompanists. Thus, not the musical material itself but the embodied *intention* of being a soloist or accompanist should influence performers’ movements while playing.

In the second experiment, we employed eye-tracking technology to investigate observers’ gaze during musical duo performance. Observers were invited to watch a set of duo performances encompassing a range of different instruments. It was hypothesized that observers’ gaze would be directed towards salient visuo-kinematic cues (i.e., expressive bodily gestures of a soloist), independent of their coincidence with salient auditory cues (i.e., the musical solo part).

Experiment 1

METHOD

Participants. Ten gender-homogeneous duos of professional musicians took part in Experiment 1. We analyzed data for seven duos ($N = 14$; 10 males; $M_{age} = 27.36$ years, $SD = 5.90$; $M_{music\ training} = 18.79$ years, $SD = 5.91$) due to issues with the motion capture data for the other three duos who used music stands interfering with reflective markers. All remaining seven duos played from memory and were recruited from professional music organizations and music conservatories. The aim was to have a broad, general view of music performance, with a range of different instruments (including voice). Participants were trained in various music genres such as classical, pop, jazz, and folk music. To control for familiarity between the performers, only duos who had at least two years of experience playing together were invited. The musicians did not rehearse their behavioral roles to avoid possible learning effects. All participants signed a form to declare that they: 1) participated voluntarily, 2) gave clearance for video recording of the performances (for scientific and educational purposes only), 3) received sufficient information concerning the tasks, the procedures, and the technologies used, and 4) were aware of being able to ask questions at any time throughout the experiment.

MUSIC MATERIALS

Each piece had a maximum duration of 60 s and was played twice, resulting in four performances per duo. The pieces were self-selected and rehearsed during the week prior to the experiment as well as just before the start of the experiment. Musicians were instructed to choose pieces which had a clearly identifiable solo part that differed in musical complexity (pitch, rhythm, dynamics)

TABLE 1 Overview of Musical Duos, Genres, Pieces, Composers, and Instrumentation

Duo (sex)	Music genre	Pieces and composers	Instrumentation
1 (f)	Classical / jazz	Piece 1: Wiegenlied Op. 49 No. 4 (Johannes Brahms) Piece 2: The Lion Sleeps Tonight (Solomon Linda)	Mezzo-soprano voice Soprano voice
2 (m)	Jazz / latin	Piece 1: 7/4 song (own composition) Piece 2: Latin song (own composition)	Acoustic guitar Alto saxophone
3 (m)	Blues	Piece 1: Funny Hi (Modern Art) Piece 2: Cage (Modern Art)	Acoustic guitar Bass guitar
4 (m)	Swing / spiritual	Piece 1: Jesus on the Mainline (traditional) Piece 2: Joshua Fit The Battle Of Jericho (traditional)	Tenor saxophone Electric guitar
5 (m)	Hiphop / funk	Piece 1: Tha Shiznit (Snoop Dogg) Piece 2: Low Rider (War)	Melodica Trombone
6 (m)	Jazz	Piece 1: Bye Bye Blackbird (Jerome H. Remick) Piece 2: Autumn Leaves (Johnny Mercer)	Acoustic guitar B \flat clarinet
7 (f)	Folk	Piece 1: Scottish Urbaine (Grégory Jolivet) Piece 2: Swedish Polka (traditional)	Diatonic accordion Bagpipes

from the accompaniment part. More information on music genres, pieces (including composers) and instruments played by each duo is reported in Table 1.

PROCEDURE

Experiments took place in a laboratory room with black curtains surrounding the test area to ensure that participants would be shielded from the experimenter and possible outside influences. All experiments took place

in the morning. Upon arrival, the participants were dressed in a black suit and hat with reflective markers required to facilitate motion capturing (see Figure 1). Instructions were given to enable the participants to continue the rest of the procedure independently, without further experimenter intervention.

Over the course of the experiment, performer A played the solo in the first piece and the accompaniment in the second, while performer B played the



FIGURE 1. Example of musical duo (No. 4) performing during the experiment.

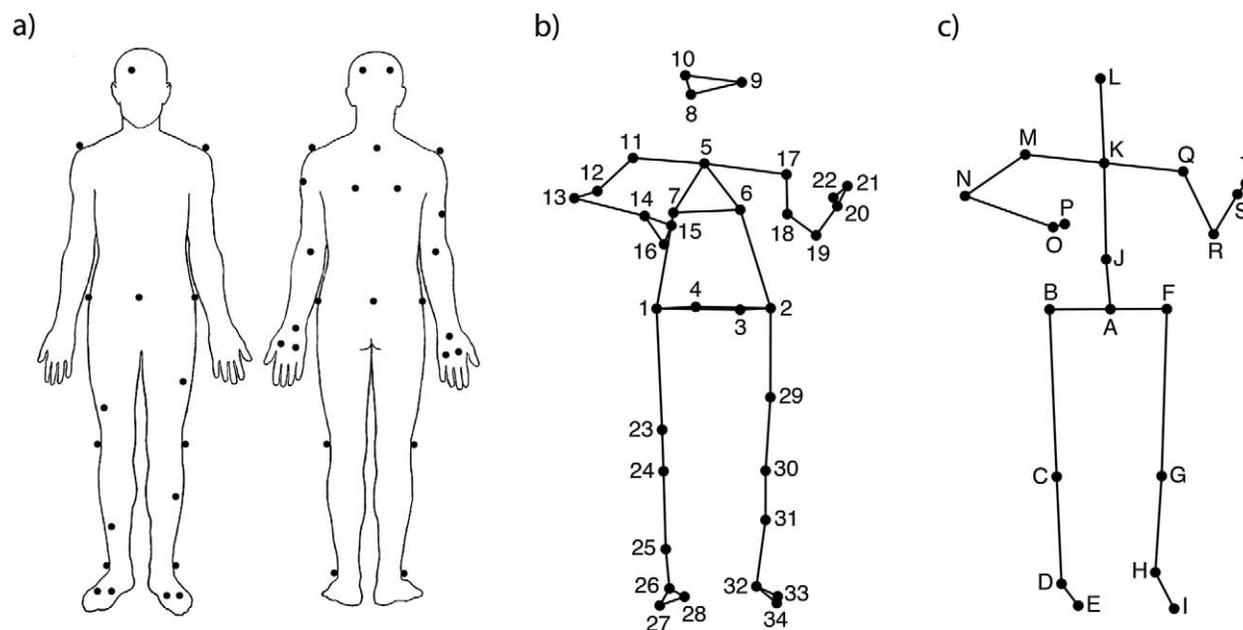


FIGURE 2. Marker and joint locations. a) Anterior and posterior views of the marker placement on the participants' bodies; b) Anterior view of the marker locations as stick figure illustration; c) Anterior view of the locations of the secondary markers/joints used in the analysis.

accompaniment in the first piece and the solo in the second. Each piece was played in two conditions: 1) the congruent condition in which the soloist (musical solo-behavioral solo) and accompanist (musical accompaniment-behavioral accompaniment) behaved according to their respective roles, and 2) the incongruent condition in which the accompanist behaved as a soloist (musical accompaniment-behavioral solo) and the soloist behaved as an accompanist (musical solo-behavioral accompaniment). The corresponding instructions were as follows: 1) Play the piece and behave according to your respective musical roles; the soloist behaving as a soloist and the accompanist behaving as an accompanist. 2) Play the piece once more, but now switch roles; the soloist behaving as an accompanist and the accompanist behaving as a soloist. At the end of the experiment, participants were asked to fill out a questionnaire, which contained questions concerning their age and musical background.

MOVEMENT RECORDING

An infrared optical motion capture system (OptiTrack) consisting of 12 synchronized cameras with related ARENA motion capture software (<http://www.naturalpoint.com>) was used to record participants' body movements in three dimensions. Each participant was equipped with reflective markers (see Figure 2a and 2b for reference). The locations of the markers

were as follows: 1-4: hip; 5-7: upper back; 8-10: head (attached to a baseball cap); 11: right shoulder; 12: right upper arm; 13: right elbow; 14-16: right hand (rigid body); 17: left shoulder; 18: left upper arm; 19: left elbow; 20-22: left hand (rigid body); 23: right thigh; 24: right knee; 25: right shank; 26: right ankle; 27-28: right foot (rigid body); 29: left thigh; 30: left knee; 31: left shank; 32: left ankle; 33-34: left foot (rigid body). The motion capture data were tracked at a sampling rate of 100 Hz. All performances were exported into BioVision Hierarchy (BVH) files. Additionally, performances were audio-visually recorded using a Canon Legria HF S100 camera while audio was captured with a Zoom H4n portable recorder.

MOVEMENT FEATURE EXTRACTION

Using the MATLAB motion capture toolbox (<http://www.cs.man.ac.uk/~neill/mocap/>), the three-dimensional position and displacement of the markers were calculated. That data were read into MoCap Toolbox (Burger & Toiviainen, 2013) to extract movement features used in subsequent analyses. The first step was to trim the data to the duration of the performances. Following this, a set of 20 secondary markers, subsequently referred to as joints, was derived from the original data (see Figure 2c). The locations of joints B, C, D, E, F, G, H, M, N, P, Q, R, and T are identical to the locations of the respective ones in the original marker setup, while

the remaining joints were obtained by averaging the locations of two or more markers: joint A: midpoint of all four hip markers; E: mid-point of both right foot markers; I: mid-point of both left foot markers; J: mid-point of the four hip and the three back markers; K: mid-point between both shoulder markers; L: mid-point of the three head markers; O: mid-point of the three right hand markers; S: mid-point of the three left hand markers. Subsequently, the data were transformed to a local coordinate system by rotating them on a frame-by-frame basis, so that the shoulder joints (M and Q) were parallel to the x -axis and defining the mid-point of the shoulders (joint K) as the new origin, to express the data relative to this point and to ensure that all performers are aligned along the x -axis and face in the same direction.

In line with our hypothesis related to head and shoulder movement in particular (Castellano et al., 2008; Dahl & Friberg, 2007; Glowinski et al., 2013; Thompson & Luck, 2012), movement feature extraction focused on these body parts. Since most performers' lower bodies did not move during playing, leg movement was not analyzed. A full-body marker setup was nevertheless chosen for data collection in order to create full-body animations and to have full-body data available if needed. The following nine movement features were extracted:

Quantity of motion: an overall measurement of the amount of detected motion, operationalized by the cumulative distance travelled, divided by the duration of the performance (Burger & Toiviainen, 2020 [called 'Linear Speed' there]; Camurri, Lagerlöf, & Volpe, 2003; Eerola, Jakubowski, Moran, Keller, & Clayton, 2018; Jensenius, Zelechowska, & Gonzalez Sanchez, 2017). The feature is calculated for the head joint (L) and the averaged shoulder joints (M and Q).

Fluidity: overall movement fluidity/smoothness measure based on the ratio of velocity to acceleration. The combination of high velocity and low acceleration reflects fluid movement, whereas the combination of low velocity and high acceleration reflects non-fluid movement. The feature is calculated for the head joint (L) and the averaged shoulder joints (M and Q).

Bounding rectangle: the smallest rectangle that fits the movement of either the head (joint L) or the shoulders (average of joints M and Q) as a two-dimensional projection on the horizontal plane (i.e., floor), averaged across four-second analysis windows with a two-second overlap. Both fluidity and bounding rectangle have been used in Burger, Saarikallio, Luck, Thompson and Toiviainen (2013) for studying emotional expression when moving to music.

Shoulder speed comprises three features: the speed of both shoulder joints for each dimension separately, indicating the amount of movement in each movement direction (i.e., medio-lateral/sideway [X], antero-posterior/back and forth [Y], and superior-inferior/vertical [Z]). For this feature, the data were rotated to a frontal view to align the coordinate system, but was not transformed to the local coordinate system.

RESULTS

For head and shoulder movements, the following features were computed: quantity of motion, fluidity, and bounding rectangle. Additionally, the speed of the shoulders was computed along the three axes x (side-ways), y (back and forth), and z (up and down). A 2×2 repeated-measures ANOVA was performed with the independent variables Musical Role (soloist vs. accompanist) and Behavioral Role (soloist vs. accompanist) to investigate the effects of musical and behavioral role on these movement features. All follow-up t -tests were computed within a performer and corrected with the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). When behavioral roles were compared, this occurred within one piece of a duo; when musical roles were compared, this occurred across two pieces of a duo.

Quantity of motion. For head movements, results showed a significant main effect of Behavioral Role, $F(1, 13) = 36.041, p < .001, \eta_p^2 = .735$. Soloists moved their head more when they were behaving as a soloist ($M = 0.935, SEM = 0.103$) compared to behaving as an accompanist ($M = 0.553, SEM = 0.073$), $t(13) = 4.676, p < .001, r = .792$. Similarly, accompanists moved their head more when they were behaving as a soloist ($M = 1.263, SEM = 0.171$) compared to behaving as an accompanist ($M = 0.669, SEM = 0.101$), $t(13) = 5.055, p < .001, r = .814$. Interestingly, accompanists behaving as soloists moved their head more than soloists behaving as soloists, $t(13) = 2.460, p = .029, r = .564$.

For shoulder movements, there was also a significant main effect of Behavioral Role, $F(1, 13) = 27.997, p < .001, \eta_p^2 = .683$. Soloists moved their shoulders more when they were behaving as a soloist ($M = 0.639, SEM = 0.069$) compared to behaving as an accompanist ($M = 0.380, SEM = 0.067$), $t(13) = 5.185, p < .001, r = .821$. Likewise, accompanists moved their shoulders more when they were behaving as a soloist ($M = 0.863, SEM = 0.130$) compared to behaving as an accompanist ($M = 0.403, SEM = 0.060$), $t(13) = 4.182, p = .001, r = .757$. The difference between accompanists behaving as soloists and soloists behaving as soloists was not

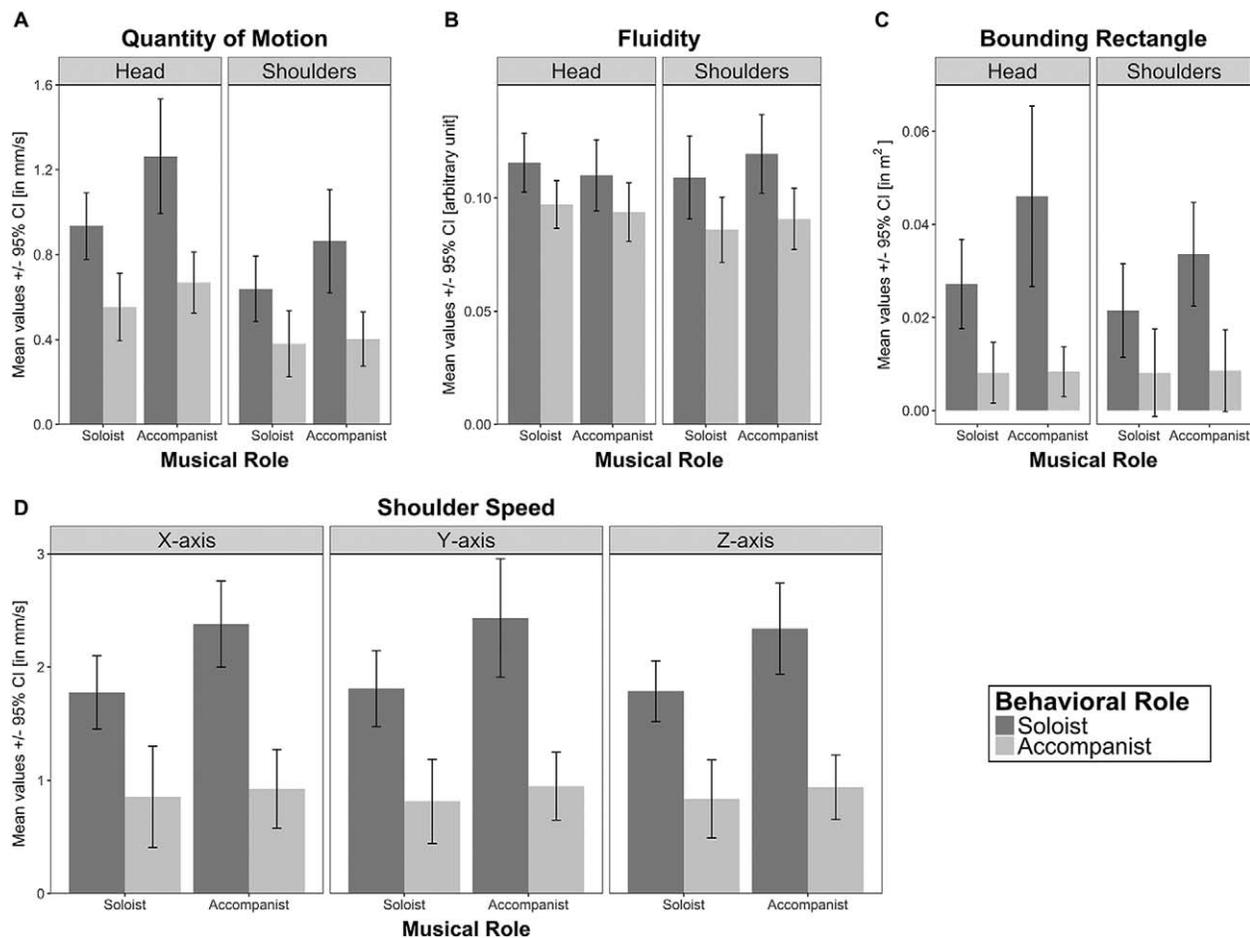


FIGURE 3. Movement features of duo musicians playing vs. acting the solo and accompaniment part. A. Quantity of Motion (head and shoulders). B. Fluidity (head and shoulders). C. Bounding Rectangle (head and shoulders). D. Shoulder Speed along x-, y-, and z-axes.

significant, $t(13) = 1.762$, $p = .102$, $r = .439$. For an overview, see Figure 3A.

Fluidity. For head fluidity, results showed a significant main effect of Behavioral Role, $F(1, 13) = 23.896$, $p < .001$, $\eta^2_p = .648$. Soloists showed smoother head movements when they were behaving as a soloist ($M = 0.116$, $SEM = 0.008$) compared to behaving as an accompanist ($M = 0.097$, $SEM = 0.007$), $t(13) = 3.781$, $p = .002$, $r = .724$. Similarly, accompanists displayed smoother head movements when they were behaving as a soloist ($M = 0.110$, $SEM = 0.008$) compared to behaving as an accompanist ($M = 0.094$, $SEM = 0.008$), $t(13) = 3.030$, $p = .010$, $r = .643$. The difference between accompanists behaving as soloists and soloists behaving as soloists was not significant, $t(13) = 0.580$, $p = .572$, $r = .159$.

For shoulder fluidity, results also showed a significant main effect of Behavioral Role, $F(1, 13) = 23.425$,

$p < .001$, $\eta^2_p = .643$. Soloists showed smoother shoulder movements when they were behaving as a soloist ($M = 0.109$, $SEM = 0.008$) compared to behaving as an accompanist ($M = 0.086$, $SEM = 0.008$), $t(13) = 3.426$, $p = .005$, $r = .689$. Similarly, accompanists displayed smoother shoulder movements when they were behaving as a soloist ($M = 0.120$, $SEM = 0.011$) compared to behaving as an accompanist ($M = 0.091$, $SEM = 0.010$), $t(13) = 4.031$, $p = .001$, $r = .745$. The difference between accompanists behaving as soloists and soloists behaving as soloists was not significant, $t(13) = 1.433$, $p = .175$, $r = .369$. For an overview, see Figure 3B.

Bounding rectangle. For head movements, results showed a significant main effect of Behavioral Role, $F(1, 13) = 28.352$, $p < .001$, $\eta^2_p = .686$. Soloists moved their head over a larger area when they were behaving as a soloist ($M = 0.027$, $SEM = 0.004$) compared to

behaving as an accompanist ($M = 0.008$, $SEM = 0.003$), $t(13) = 4.348$, $p = .001$, $r = .770$. Likewise, accompanists moved their head over a larger area when they were behaving as a soloist ($M = 0.046$, $SEM = 0.009$) compared to behaving as an accompanist ($M = 0.008$, $SEM = 0.002$), $t(13) = 4.069$, $p = .001$, $r = .748$. The difference between accompanists behaving as soloists and soloists behaving as soloists was not significant, $t(13) = 2.162$, $p = .050$, $r = .514$.

For shoulder movements, results also showed a significant main effect of Behavioral Role, $F(1, 13) = 32.291$, $p < .001$, $\eta^2_p = .713$. Soloists moved their shoulders over a larger area when they were behaving as a soloist ($M = 0.022$, $SEM = 0.004$) compared to behaving as an accompanist ($M = 0.008$, $SEM = 0.004$), $t(13) = 3.378$, $p = .005$, $r = .684$. Similarly, accompanists moved their shoulders over a larger area when they were behaving as a soloist ($M = 0.034$, $SEM = 0.006$) compared to behaving as an accompanist ($M = 0.009$, $SEM = 0.004$), $t(13) = 4.264$, $p = .001$, $r = .764$. The difference between accompanists behaving as soloists and soloists behaving as soloists was not significant, $t(13) = 2.130$, $p = .053$, $r = .509$. For an overview, see Figure 3C.

Shoulder speed. For the x -axis, results showed a significant main effect of Behavioral Role, $F(1, 13) = 54.784$, $p < .001$, $\eta^2_p = .808$. Soloists moved their shoulders faster sideways when they were behaving as a soloist ($M = 1.776$, $SEM = 0.198$) compared to behaving as an accompanist ($M = 0.855$, $SEM = 0.208$), $t(13) = 4.668$, $p < .001$, $r = .791$. Likewise, accompanists moved their shoulders faster sideways when they were behaving as a soloist ($M = 2.379$, $SEM = 0.218$) compared to behaving as an accompanist ($M = 0.926$, $SEM = 0.184$), $t(13) = 6.544$, $p < .001$, $r = .876$. Accompanists behaving as soloists moved their shoulder faster sideways than soloists behaving as soloists, $t(13) = 2.500$, $p = .027$, $r = .570$.

For the y -axis, results also showed a significant main effect of Behavioral Role, $F(1, 13) = 51.854$, $p < .001$, $\eta^2_p = .800$. Soloists moved their shoulders faster back and forth when they were behaving as a soloist ($M = 1.811$, $SEM = 0.189$) compared to behaving as an accompanist ($M = 0.814$, $SEM = 0.167$), $t(13) = 4.995$, $p < .001$, $r = .811$. Similarly, accompanists moved their shoulders faster back and forth when they were behaving as a soloist ($M = 2.433$, $SEM = 0.275$) compared to behaving as an accompanist ($M = 0.949$, $SEM = 0.175$), $t(13) = 5.893$, $p < .001$, $r = .853$. The difference between accompanists behaving as soloists and soloists behaving as soloists was not significant after Benjamini-Hochberg correction, $t(13) = 2.219$, $p = .045$, $r = .524$.

For the z -axis, results showed a significant main effect of Behavioral Role, $F(1, 13) = 60.883$, $p < .001$, $\eta^2_p = .824$. Soloists moved their shoulders faster up and down when they were behaving as a soloist ($M = 1.787$, $SEM = 0.184$) compared to behaving as an accompanist ($M = 0.836$, $SEM = 0.164$), $t(13) = 5.398$, $p < .001$, $r = .832$. Similarly, accompanists moved their shoulders faster up and down when they were behaving as a soloist ($M = 2.339$, $SEM = 0.237$) compared to behaving as an accompanist ($M = 0.940$, $SEM = 0.170$), $t(13) = 6.591$, $p < .001$, $r = .877$. The difference between accompanists behaving as soloists and soloists behaving as soloists was not significant, $t(13) = 2.100$, $p = .056$, $r = .503$. For an overview, see Figure 3D.

DISCUSSION

The first study revealed that musicians behaving as soloists—regardless of whether they played the musical solo or accompaniment—showed increased values in a number of movement features in comparison to musicians behaving as accompanists. Specifically, musicians behaving as soloists displayed more head and shoulder movements (quantity of motion), moved these body parts more smoothly (fluidity) over a larger area (bounding rectangle), and moved their shoulders faster along all three axes (shoulder speed) in comparison to musicians behaving as accompanists.

Our findings are in line with studies reporting that leaders of musical duos show larger bodily gestures than accompanists (Goebel & Palmer, 2009) and that musicians increase the amount of movements—especially of head and shoulders (Thompson & Luck, 2012), but also of the entire body and instrument if possible (Wanderley, Vines, Middleton, McKay, & Hatch, 2005)—when asked to perform more expressively. More generally, our results resonate with literature showing that musicians are able to adapt their behavior depending on their role in a musical ensemble (Glowinski et al., 2013). In a regular concert setting, a soloist's role may be emphasized by a number of attention-capturing visual cues such as facial expression (Thompson, Russo, & Livingstone, 2010), clothing (Griffiths, 2008) or physical appearance (Wapnick, Mazza, & Darrow, 1998, 2000). We provide evidence that—in addition to these visual cues—the intentions driving body movements are important factors for distinguishing a soloist's musical performance from that of an accompanist. Interestingly, our results also show that accompanists in the incongruent condition (i.e., behaving as soloists) tend to move their head more and over a larger area—as well as moving their shoulders faster—than soloists in the congruent condition. These exaggerated body movements may indicate

difficulties in adjusting auditory-motor couplings during music performance (for a review, see Zatorre, Chen, & Penhune, 2007), possibly because of the increased cognitive load associated with the production of this unusual musical performance. The conflict between sound-producing and expressive gestures—which are normally integrated into one overarching motor program for the performance of a specific piece—may therefore lead to an “overshooting” of communicative, expressive bodily gestures. Even though motor theories based on optimal feedback control allow the integration of higher-order goals and real-time sensorimotor control (Todorov, 2004; Todorov & Jordan, 2002), several trials of motor learning may be needed to re-adjust sound-producing and expressive gestures.

While differences between soloist and accompanist performance may arise depending on musical genre, style, and instrument (Davidson & Broughton, 2016), averaging across these factors—as we did in our study—still yields significant differences between soloists’ and accompanists’ head and shoulder movements. The change of behavior depending on the musical role in an ensemble may be dictated by physical constraints of the instrument (Davidson, 2012), the need for sound-producing movements (Jenselius et al., 2010), the underlying musical structure (MacRitchie, Buck, & Bailey, 2013), as well as the style of the music (Huang & Krumhansl, 2011). A soloist’s behavior is furthermore constrained by cultural norms. For example, Western audiences generally expect expressive movements from soloists, whereas Japanese audiences may be irritated by expressive bodily gestures, as they may detract from the sonic expressivity of the music itself (Malm, 2000).

Despite these constraints, our participants, who had received some form of Western music training, seemed to have had a similar understanding of the connection between behavioral role, amount of expressiveness, and extent of movement. In most Western musical traditions, the role of a soloist is associated with more expressiveness which, in turn, is related to more extensive bodily gestures (Davidson, 1994; Thompson & Luck, 2012). The musicians in our study indicated that it was easier to move more expressively than less expressively. In five out of the seven cases, the duo reported that it was more difficult to behave as if they were playing the accompaniment while actually playing the solo than when asked to behave as if they were playing the solo while actually playing the accompaniment. When asked further about this discrepancy, they mentioned that it felt more natural to “move more” than to constrain themselves from moving. It thus seems that musicians themselves reflect on the behavioral distinction

between soloist and accompanist in terms of movement, further supporting the argument that music performance and movement are inherently intertwined. Moreover, Davidson and Broughton (2016, p. 14) have argued that it is nearly impossible to eliminate expressive gestures completely, suggesting that “(i) it is difficult to inhibit a learned expressive motor program, (ii) naturally expressive bodily movements and gestures are crucial to the practicalities of generating performance as well as communicating expression, (iii) expressive bodily movement naturally occurs in reaction to the sounds the body is producing, or (iv) perhaps some combination of the three.” What seems clear is that, in an audio-visual music performance, auditory information alone is not sufficient for conveying all the features associated with a soloist performance. Being a soloist, or rather, the defining characteristics of being a soloist, is determined by a *confluence* of auditory, visual, and visuo-kinematic cues, and as accumulating evidence shows, the latter two may be equally as powerful as the former (e.g., Griffiths, 2010; Kawase & Obata, 2016; Platz & Kopiez, 2012; Waddell & Williamson, 2017; Wapnick et al., 1998).

Experiment 2

In the first experiment, it was shown that behaving as soloist, regardless of the musical role (soloist or accompanist), led to more, smoother, and faster head and shoulder movements over a larger area than behaving as accompanist. Moreover, accompanists behaving as soloists (incongruent condition) moved more than soloists behaving as soloists (congruent condition). In our second experiment, we investigated whether the differences in performances between conditions demonstrated in the first experiment would influence the gaze of an audience observing the performance. In addition, the impact of the auditory input on observers’ gaze was explored.

METHOD

Participants. In total, 34 participants (henceforth referred to as “observers”) were tested (16 male; $M_{age} = 31.26$ years; $SD = 6.71$). Their music training varied between 0 and 30 years ($M = 9.13$, $SD = 8.27$, $Mdn = 9.0$); 9 observers had received no music training.

Materials. Video excerpts consisted of the 28 duo performances recorded in the previous experiment. Each excerpt had a duration of 20 s, and only one excerpt per performance was created. The duration was kept constant to exclude possible duration effects (e.g., impact on boredom, possibly resulting in different

TABLE 2 Mean Tempo and Mean Sound Pressure Level in Congruent and Incongruent Performance Conditions

Duo	Piece	Tempo (BPM)		Sound Pressure Level (dB)	
		Congruent	Incongruent	Congruent	Incongruent
1	1	93.50	97.50	70.46	70.29
1	2	95.00	98.00	74.87	73.82
2	3	97.50	93.50	79.78	82.15
2	4	99.00	97.50	78.54	77.53
3	5	93.50	94.50	79.28	80.05
3	6	90.50	95.50	82.99	82.26
4	7	94.50	93.00	76.00	75.97
4	8	95.50	91.50	70.20	71.57
5	9	94.50	94.00	69.78	71.80
5	10	97.00	93.00	72.81	71.55
6	11	96.00	96.00	72.71	72.27
6	12	98.00	96.00	64.35	64.69
7	13	95.00	93.00	72.62	73.14
7	14	98.00	99.50	72.58	72.83

Note: BPM = beats per minute; dB = decibel.

gazing strategies). To avoid potential start-up effects, none of the excerpts included material from the initial 10 s of a performance. Each excerpt started with the beginning of a musical phrase, and a fade-out was added at the end to conceal possible disruptive cut-offs of the musical phrasing. To obtain an equal level of perceived loudness for all clips, Audacity (Version 2.3.0) was used to normalize the audio across two performances of the same piece (see Table 2). Additionally, a paired-samples *t*-test was carried out to check for possible differences in mean tempo between the different performances of each piece. The analysis showed no significant difference between the congruent ($M = 95.536$, $SEM = 0.606$) and incongruent ($M = 95.179$, $SEM = 0.626$) condition, $t(13) = 0.455$, $p = .657$, $r = .125$ (see Table 2). All observers wore Sennheiser HD60 headphones during the experiment.

Eye tracking. Eye tracking was chosen as a noninvasive method of measuring behavior because it is an unobtrusive strategy to gather information with regard to observers' gaze, providing more objective evidence of behavior than self-reports, for instance. The relationship between observers' gaze and their visual attention is complex, yet the direction of gaze was interpreted to coincide with the focus of visual attention in our experimental setup (see also Discussion).

Eye movements were recorded using a remote eye-tracking device (RED) by SensoMotoric Instruments (SMI), operating at 120 Hz. A five-point calibration was used and three validations were performed during the experiment. To calculate the time participants watched

the different regions of the performers' bodies, dynamic areas of interest (AOIs) were coded on the video excerpts using BeGaze 3.2 (SMI), each of them representing the full body of a musician in a specific condition. Once the AOIs (Musical solo-Behavioral solo, Musical solo-Behavioral accompaniment, Musical accompaniment-Behavioral solo, and Musical accompaniment-Behavioral accompaniment) were coded, dwell-time percentages (percentage of time the eyes were directed towards the AOI) were retrieved. Validation tests showed an average accuracy (i.e., the horizontal average distance from the actual gaze point to the one measured by the eye tracker) of 0.69° ($SD = 0.35$) and the average tracking ratio (percentage of time eye movement was actually measured) was 95.96% ($SD = 3.86$). Eye tracking data of all observers proved to be accurate enough to be analyzed statistically. To investigate relationships between dwell time and displayed bodily gestures, the same movement features as above (quantity of motion, fluidity, bounding rectangle and shoulder speed) were calculated for the 20-s motion capture excerpts matching the respective video recordings shown in the perceptual experiment.

Procedure. Observers were seated in front of a projector screen attached to a wall, the RED was placed on a table in front of them, and they were equipped with Sennheiser HD60 headphones. The distance between all items and the observers was fixed at 3.5 m, and they were not able to adjust the loudness during the experiment. Observers watched each video excerpt twice, once with and once without audio, resulting in 56 trials in total. Visual-only and audio-visual stimuli were presented in two blocks, with visual-only stimuli always being presented first. The order of the video excerpts within each block was randomized, while ensuring that a particular sequence could occur only once. After watching the video excerpts, observers were asked to fill out a short questionnaire on the main reason for normally attending music performances and their general focus of attention (visual vs. musical information) during concerts.

RESULTS

Questionnaire. Among all observers, 73.53% reported that the musical aspects are the main reason for attending music performances, while only 2.94% opted for visual aspects; 23.53% indicated that both aspects contribute roughly equally as motives for attending music performances. When attending music performances, 20.59% of all observers declared that their attention is mainly directed towards musical information, 20.59% reported that their attention is generally

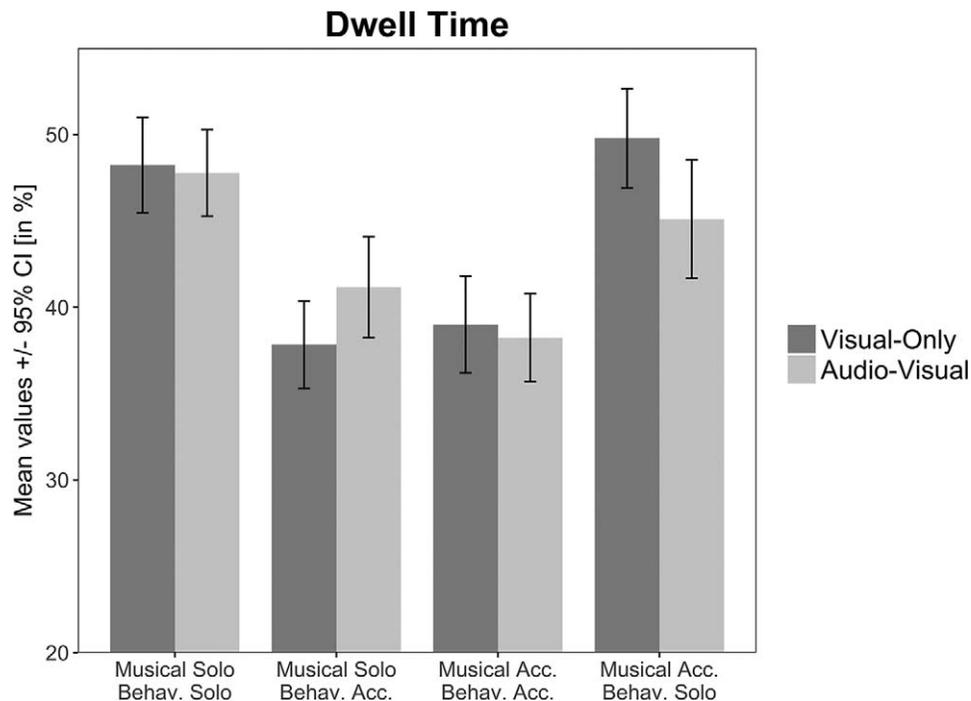


FIGURE 4. Observers' average dwell time on duo musicians performing congruently (musical and behavioral roles aligned) and incongruently (musical and behavioral roles opposed). *Note:* Behav. = Behavioral. Acc. = Accompaniment.

directed towards visual information, and 58.82% reported that musical and visual information attract their attention similarly.

Focus of gaze. Mean dwell time was analyzed in order to investigate the focus of observers' gaze. Given the results of Experiment 1 and our hypothesis that visuo-kinematic cues dominate auditory cues in the competition for visual attention, we first computed a 2×2 ANOVA with the independent variables Musical Role (soloist vs. accompanist) and Behavioral Role (soloist vs. accompanist). Results showed a highly significant effect of Behavioral Role, $F(1, 108) = 95.106$, $p < .001$, $\eta_p^2 = .468$. Observers looked longer at musicians behaving as soloists ($M = 47.725$, $SEM = 0.629$) compared to musicians behaving as accompanists ($M = 39.055$, $SEM = 0.629$). There was no main effect of Musical Role, $F(1, 108) = .655$, $p = .420$, $\eta_p^2 = .006$. Next, to investigate how the mode of presentation affected each congruent/incongruent role, a 4×2 ANOVA with the independent variable Role (musical solo-behavioral solo, musical solo-behavioral accompaniment, musical accompaniment-behavioral solo, and musical accompaniment-behavioral accompaniment) and Mode (audio-visual vs. visual-only) was performed. The results showed a significant main effect of Role, $F(3, 104) = 34.130$, $p < .001$, $\eta_p^2 = .496$ and

a significant interaction between Role and Mode, $F(3, 104) = 3.625$, $p = .016$, $\eta_p^2 = .095$.

Benjamini-Hochberg-corrected post hoc comparisons revealed longer dwell times in the musical solo-behavioral solo condition ($M = 48.007$, $SEM = 0.860$) compared to the musical solo-behavioral accompaniment ($M = 39.493$, $SEM = 0.860$), $p < .001$, and to the musical accompaniment-behavioral accompaniment conditions ($M = 38.618$, $SEM = 0.860$), $p < .001$. Significantly shorter dwell times were observed in the musical accompaniment-behavioral accompaniment condition compared to the musical accompaniment-behavioral solo condition ($M = 47.443$, $SEM = 0.860$), $p < .001$.

The significant interaction was caused by differences in dwell times on musicians performing incongruently (see Figure 4). There was a trend towards increased dwell times for performers playing the solo part but behaving as accompanists in the audio-visual compared to the visual-only mode, although this comparison was not significant, $t(26) = 1.991$, $p = .057$, $r = .483$. On the other hand, there were decreased dwell times for performers playing the accompaniment but behaving as soloists in the audio-visual compared to the visual-only mode, $t(26) = 2.421$, $p = .023$, $r = .558$.

TABLE 3 Correlations Between Mean Dwell Time and Movement Features

Movement Feature	Correlation with Mean Dwell Time Split By Mode and Condition			
	AUDIO-VISUAL - CONGRUENT	AUDIO-VISUAL - INCONGRUENT	VIDEO-ONLY - CONGRUENT	VIDEO-ONLY - INCONGRUENT
Quantity of Motion (H)	.263	-.012	.374	.542**
Quantity of Motion (S)	.179	.076	.405	.694***
Fluidity (H)	.405	.099	.206	.141
Fluidity (S)	.256	.172	.230	.376
Bounding Rectangle (H)	.389	.160	.423	.542***
Bounding Rectangle (S)	.411	.198	.429	.614***
Shoulder Speed (X)	.210	.196	.380	.658***
Shoulder Speed (Y)	.276	.239	.449	.685***
Shoulder Speed (Z)	.315	.242	.432	.697***

Note: We used the Benjamini-Hochberg procedure to correct for multiple testing. $N = 28$, ** $p < .01$, *** $p < .001$.
H = head; S = shoulders; X = x-axis; Y = y-axis; Z = z-axis.

Subsequently, we correlated the nine movement features with the mean dwell times, separately for the congruent and incongruent performances in the audio-visual and visual-only modes, respectively. To correct for multiple comparisons, we used again the Benjamini-Hochberg adjustment. There were no significant correlations in the congruent and incongruent performances of the audio-visual mode. However, there were significant correlations in the visual-only mode when musicians performed incongruently (see Table 3) for quantity of motion (head and shoulders), bounding rectangle (head and shoulders), and shoulder speed (all three axes). None of the correlations was significant in the visual-only mode when musicians behaved congruently.

DISCUSSION

The second experiment revealed that certain visuo-kinematic cues attract observers' visual attention, independent of the musical cues (i.e., solo or accompaniment) associated with a duo performer. Our results further suggest that the absence of auditory information may increase the effect of visual attention allocation. When both duo performers behaved according to their musical roles (i.e., in the congruent condition), observers looked longer at the soloists, regardless of whether audio was present or not. When musicians performed incongruently, however, the mode of presentation (audio-visual vs. visual-only) had an impact on dwell times. Observers looked longer at accompanists behaving as soloists in the visual-only compared to the audio-visual mode. For soloists performing incongruently, dwell times did not differ between the two modes of presentation. These findings were substantiated by correlation analyses that revealed that

significant positive correlations between movement features and average dwell times only occurred in the visual-only mode when musicians behaved incongruently. Specifically, significant correlations were obtained for quantity of motion (head and shoulders), bounding rectangle (head and shoulders), and shoulder speed (all three axes).

Before these results can be discussed in light of theory and previous research, it is crucial to consider that visual perception and visual attention are two different concepts (Remington, 1980; Rensink, O'Regan, & Clark, 1997; Shulman, Remington, & McLean, 1979). For instance, covert visual attention allows us to gaze at one object while attending to another in the periphery. In some circumstances, attention may even be directed inwardly, such as when we are distracted by intrusive thoughts or switch to states of mind-wandering (Seli et al., 2018). While a shift of gaze from performer A to performer B does not necessarily mean that an observer is also attending to performer B, it is more likely that they are now focusing on performer B rather than performer A. Since observers in our study could move their gaze freely over the screen (i.e., there was no fixation cross), it is unlikely that an observer would try to peripherally attend to performer A while looking at performer B. In other words, the location of our observers' gaze in our experimental setup most likely coincided with their visual attention, although additional objective (e.g., EEG to detect default mode network) and subjective (continuous self-reports, c.f. Broughton, Schubert, Harvey, & Stevens, 2019) measurements may be needed for more conclusive evidence.

The pattern of our results suggests that bottom-up perceptual processes were primarily driving visual attention. Visuo-kinematic cues of performers moving

more expressively likewise captured more of observers' attention, without observers necessarily being aware of or able to control this behavior. This suggestion supports Tsay's (2013) notion of a natural, automatic, and unintentional dependence on visual cues during musical performances. More importantly, our results help contextualize previous studies reporting that observers' visual attention depends on the musical part (Kawase & Obata, 2016) and that melodic parts of a musical performance attract listeners' attention (Gregory, 1990), especially if the melody or the solo part is in a higher register than the accompaniment (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2005; Trainor, Marie, Bruce, & Bidelman, 2014). In a musical performance with unexpected visuo-kinematic cues such as an accompanist behaving as a soloist, saliency of the auditory information (i.e., the musical solo part) does not always seem to be sufficient to attract visual attention. Further studies are needed that manipulate both musical features (e.g., pitch range, dynamics, rhythm) and bodily expressive gestures (e.g., normal vs. exaggerated) in solo and accompaniment parts to investigate the effects of audio-visuo-kinematic processing on visual (and auditory) attention.

Our results also show that the mode of presentation significantly affects observers' eye movements such that the absence of auditory information increases the dwell time on accompanists behaving as soloists. It is possible that the inappropriateness and unexpectedness of accompanists' exaggerated head movements caused by musicians' need to exert unusual auditory-motor couplings (see Discussion of Experiment 1) attracted even more visual attention in the absence of competing auditory cues. The issue therefore seems to be not only one of saliency, but also one of audio-visual integration and context-dependent, selective auditory and visual attention. While it has been shown that audio-visual integration can occur very early in sensory perception (Giard & Peronnet, 1999), more recent evidence suggests that multimodal integration and attention can occur at different levels of information processing (Koelewijn, Bronkhorst, & Theeuwes, 2010). Our findings suggest that when audio and visuals are present in music performance, vision still dominates sound for attention, but the presence of auditory information can modulate visual attention—possibly similarly to the way the ventriloquist effect can be modulated by the quality of visuo-spatial information (Alais & Burr, 2004).

Given these results, observers' ability to attend to musical cues in the presence of salient visuo-kinematic cues should not be underestimated, which is in line with evidence demonstrating that auditory stimuli affect

visual attention, whereas visual stimuli do not influence auditory attention (Driver & Spence, 1998). For example, even if observers' visual attention is directed towards performers' expressive movements in most circumstances, observers may still be able to focus their auditory attention elsewhere. This process may require some cognitive effort—for instance, effort similar to that needed by judges in musical competitions who must assess the musical quality of a performance. Accordingly, hiring committees and audition panels have embraced “blind” screenings (Goldin & Rouse, 2000) not only out of the pursuit of fairness, but also in response to critics who disparage those who prioritize visually stimulating choreography over musicians' sonic enactment of a composer's work (e.g., Tommasini, 2003).

General Discussion

In this study, our two primary goals were as follows: first, to investigate how soloists' and accompanists' bodily gestures change according to their role in musical duos and second, how bodily gestures affect observers' visual attention. We hypothesized that musicians behaving as soloists would display more expressive bodily gestures and that visuo-kinematic cues associated with these gestures would attract visual attention, independent of the coinciding auditory cues (i.e., solo vs. accompaniment musical parts). While our findings suggest that musicians' embodied intention changes their bodily gestures, this may result in unnatural, exaggerated movements. In a typical musical performance where visual and auditory information are present, a musician behaving as soloist attracts most visual attention—an effect that is present even when audio is absent. These results further various research areas including bodily gestures in musical duo performances (Goebel & Palmer, 2009; Keller & Appel, 2010; King & Ginsborg, 2011) and audience gaze during concerts (Kawase & Obata, 2016). Nevertheless, it is important to consider several limitations in further pursuing these research topics.

The pattern of our results in both experiments may be explained by expectations and knowledge structures formed through previous experiences. For instance, it would not be uncommon to act as or observe an understated soloist, but it would be strange to act as or observe an overstated accompanist. Thus, our expectations for acting out or observing usual patterns of behavior in musical performances may have affected low-level audio-motor couplings (Baumann et al., 2007; Palomar-García, Zatorre, Ventura-Campos, Bueichekú, & Ávila, 2017; Stephan, Lega, & Penhune,

2018) and responses to visuo-kinematic cues (Tsay, 2013, 2014; Vuoskoski et al., 2014), respectively, possibly through increased cognitive load.

In our experiment, asking musicians to behave as soloists while playing the accompaniment part, and vice versa, may have altered certain aspects of the musical sound. Though comprehensive categorization of performers' gestures (Delalande, 1988; Jensenius et al., 2010) has greatly improved the coherence of subsequent studies, these gesture types are nevertheless interdependent and blend into one another. In addition, changing communicative gestures (e.g., to signal the status of a soloist while playing the accompaniment) represents an extra effort required to coordinate the auditory-motor couplings and may be reflected not only in the visual-kinematic cues but also in the auditory output (i.e., the musical sound). In the incongruent condition, the musical accompaniment may have sounded more expressive and the musical solos more stagnant, respectively. Moreover, the embodied intentions of performer A might have affected the auditory-motor couplings of performer B and vice versa.

It should also be taken into account that the musical stage itself invites musicians to enact a certain role. In some performance settings, musicians know in advance whether they are to perform as the soloist or accompanist and can prepare their performance accordingly. In other performance settings, the roles of soloist and accompanist can be (come) more blurred. Future research on duo performers and other ensemble musicians could vary the degree to which the musical roles are separable in order to investigate the effects on bodily gestures.

Familiarity and expertise of duo musicians are further important factors that can influence musicians' enactment of bodily gestures and observers' perception of musical performances. King and Ginsborg (2011) have demonstrated that duo musicians tend to use bodily gestures to a larger extent and more variedly when performing with familiar partners of the same level of expertise compared to unfamiliar partners with a different level of musical expertise. Given that our participants had at least two years of experience playing together in duos, they were highly familiar with their duo partners and may have used a wider range of bodily gestures than they would have with unfamiliar partners. In line with King and Ginsborg's argument that familiar duos have an advantage of cognitive processing as they can anticipate and adapt to their partner's intentions and actions, it can be expected that the execution of unusual auditory-motor couplings such as the ones required in Experiment 1 would have been perceived as even more difficult for unfamiliar duos.

We tested different instrumentalists playing different kinds of musical genres and styles. Although we provide strong evidence that musicians' head and shoulder movements change similarly when asked to act as a soloist as compared with an accompanist (i.e., more extensive, smoother, and faster), we acknowledge that there are potential discrepancies among different musical genres in terms of how these embodied intentions are enacted. Thus, further studies are needed to differentiate the impact of different genres, instruments, and styles on bodily gestures in music ensemble performance.

Coordination and cooperation within musical ensembles—whether in duos or larger orchestras—is facilitated by an interconnected system of agents who are receptive to deviations from the expected output; further, these deviations are not limited to a single modality, but may be changed, for example, from within the sonic, haptic, or visual realms. Given these interdependencies, there is a trade-off between studying musicians' behavior in ensembles for optimized ecological validity and independence of data points for optimized statistical analysis. Advanced techniques to model behavior in group settings (e.g., Granger-coupling, see Chang et al., 2019) or to develop computational models of social interaction (c.f. Volpe, D'Ausilio, Badino, Camurri, & Fadiga, 2016) may provide a better understanding of the broad repertoire of body movements utilized to cooperate and coordinate with co-performers and to communicate expressiveness to audience members during musical performances. These bodily gestures need also to be taken into account when determining an observer's appreciation or liking of a musical performance (Griffiths & Reay, 2018; Thompson, Graham, & Russo, 2005). Understanding the process of judging ensemble performances (Tsay, 2014), or specific musicians within an ensemble, will require a careful investigation of visual and auditory cues as well as consideration of observers' visual and auditory expectations in performance settings. These are influenced not only by performance traditions within a particular musical culture but also by the way young professionals are educated, and thus will require cooperation with educational psychologists and music teachers.

Our findings have practical implications for rehearsal and performance preparation. Williamon and Davidson (2002) provided evidence that duo performers adapt their bodily gestures during rehearsals. We propose that additionally considering the impact of embodied intention on body movements and observers' perception of performers' body movements can inform rehearsing strategies as well as approaches to teaching presentation

and stage skills at conservatoires. As put by Davidson and Broughton (2016, p. 3), “the body is crucial to several processes in the task of solo performance: thought, feeling, production, and communication, regardless of the particular intricacies of the instrument or voice.” Thus, our findings emphasize the importance of monitoring expressive gestures and visuo-kinematic cues depending on a musician’s role in an ensemble where bodily gestures fulfil multiple functions.

To conclude, we have shown that performers adapt their bodily gestures depending on their role within a musical duo. Soloists, especially accompanists acting as soloists, display more extensive head and shoulder movements and attract more visual attention than musicians acting as accompanists. Furthermore, extensive bodily gestures that do not fit the musical role attract more visual attention when auditory information is absent. Auditory and visuo-kinematic cues are thus both central aspects of musical performances and their interaction greatly informs our perception of performed music.

Author Note

Mats B. Küssner and Edith Van Dyck share first authorship of this paper.

We would like to thank the editors, three anonymous reviewers, and Ilana Harris for their helpful comments and suggestions. A special thanks goes to Olivia Geibel for her assistance with the literature review.

The study was approved by the Ethics Committee of the Faculty of Arts and Philosophy of Ghent University, Belgium, and all procedures followed were in accordance with the statements of the Declaration of Helsinki.

The authors received no financial support for the research, authorship, and/or publication of this article.

Correspondence concerning this article should be addressed to Mats B. Küssner, Institut für Musikwissenschaft und Medienwissenschaft, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. E-mail: mats.kussner@gmail.com

References

- ALAIS, D., & BURR, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14*, 257–262. <https://doi.org/10.1016/j.cub.2004.01.029>
- BAUMANN, S., KOENEKE, S., SCHMIDT, C. F., MEYER, M., LUTZ, K., & JANCKE, L. (2007). A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Research, 1161*, 65–78. <https://doi.org/10.1016/j.brainres.2007.05.045>
- BENJAMINI, Y., & HOCHBERG, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*, 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- BERTELSON, P., & ASCHERSLEBEN, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin and Review, 5*, 482–489. <https://doi.org/10.3758/BF03208826>
- BINDEMANN, M., BURTON, A. M., LANGTON, S. R. H., SCHWEINBERGER, S. R., & DOHERTY, M. J. (2007). The control of attention to faces. *Journal of Vision, 7*(10), 15. <https://doi.org/10.1167/7.10.15>
- BISHOP, L., & GOEBL, W. (2015). When they listen and when they watch: Pianists’ use of nonverbal audio and visual cues during duet performance. *Musicae Scientiae, 19*, 84–110. <https://doi.org/10.1177/1029864915570355>
- BROUGHTON, M. C., SCHUBERT, E., HARVEY, D. G., & STEVENS, C. J. (2019). Continuous self-report engagement responses to the live performance of an atonal, post-serialist solo marimba work. *Psychology of Music, 47*, 109–131. <https://doi.org/10.1177/0305735617736378>
- BROUGHTON, M. C., & STEVENS, C. (2009). Music, movement and marimba: An investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music, 37*, 137–153. <https://doi.org/10.1177/0305735608094511>
- BURGER, B., SAARIKALLIO, S., LUCK, G., THOMPSON, M. R., & TOIVIAINEN, P. (2013). Relationships between perceived emotions in music and music-induced movement. *Music Perception, 30*, 517–533. <https://doi.org/10.1525/mp.2013.30.5.517>
- BURGER, B., & TOIVIAINEN, P. (2013). MoCap Toolbox—A Matlab toolbox for computational analysis of movement data. In R. Bresin (Ed.), *Proceedings of the Sound and Music Computing Conference 2013*, (pp. 172–178). Berlin: Logos Verlag Berlin.
- BURGER, B., & TOIVIAINEN, P. (2020). Embodiment in electronic dance music: Effects of musical content and structure on body movement. *Musicae Scientiae, 24*, 186–205. <https://doi.org/10.1177/1029864918792594>
- CAMURRI, A., LAGERLÖF, I., & VOLPE, G. (2003). Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies, 59*, 213–225. [https://doi.org/10.1016/S1071-5819\(03\)00050-8](https://doi.org/10.1016/S1071-5819(03)00050-8)
- CASTELLANO, G., MORTILLARO, M., CAMURRI, A., VOLPE, G., & SCHERER, K. (2008). Automated analysis of body movement in emotionally expressive piano performances. *Music Perception, 26*, 103–119. <https://doi.org/10.1525/mp.2008.26.2.103>

- CHANG, A., KRAGNESS, H. E., LIVINGSTONE, S. R., BOSNYAK, D. J., & TRAINOR, L. J. (2019). Body sway reflects joint emotional expression in music ensemble performance. *Scientific Reports*, 9, 205. <https://doi.org/10.1038/s41598-018-36358-4>
- CHANG, A., LIVINGSTONE, S. R., BOSNYAK, D. J., & TRAINOR, L. J. (2017). Body sway reflects leadership in joint music performance. *Proceedings of the National Academy of Sciences*, 114, E4134–E4141. <https://doi.org/10.1073/pnas.1617657114>
- CHAPADOS, C., & LEVITIN, D. J. (2008). Cross-modal interactions in the experience of musical performances: Physiological correlates. *Cognition*, 108, 639–651. <https://doi.org/10.1016/j.cognition.2008.05.008>
- CONNOR, C. E., EGETH, H. E., & YANTIS, S. (2004). Visual attention: Bottom-up versus top-down. *Current Biology*, 14, R850–R852. <https://doi.org/10.1016/j.cub.2004.09.041>
- DAHL, S., & FRIBERG, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, 24, 433–454. <https://doi.org/10.1525/mp.2007.24.5.433>
- DAVIDSON, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, 21, 103–113.
- DAVIDSON, J. W. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? *Journal of Human Movement Studies*, 26, 279–301.
- DAVIDSON, J. W. (2002). Understanding the expressive movements of a solo pianist. In K.-E. Behne, G. Kleinen, & H. de la Motte-Haber (Eds.), *Musikpsychologie: Jahrbuch der Deutschen Gesellschaft für Musikpsychologie: Band 16. Wirkungen und kognitive Verarbeitung in der Musik* (pp. 7–29). Göttingen: Hogrefe.
- DAVIDSON, J. W. (2005). Bodily communication in musical performance. In D. Miell, D. J. Hargreaves, & R. MacDonald (Eds.), *Musical Communication* (pp. 215–238). Oxford, UK: Oxford University Press.
- DAVIDSON, J. W. (2007). Qualitative insights into the use of expressive body movement in solo piano performance: A case study approach. *Psychology of Music*, 35, 381–401. <https://doi.org/10.1177/0305735607072652>
- DAVIDSON, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies. *Psychology of Music*, 40, 595–633. <https://doi.org/10.1177/0305735612449896>
- DAVIDSON, J. W., & BROUGHTON, M. C. (2016). Bodily mediated coordination, collaboration, and communication in music performance. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed.). Oxford, UK: Oxford University Press.
- DAVIDSON, J. W., & COULAM, A. (2006). Exploring jazz and classical solo singing performance behaviours: A preliminary step towards understanding performer creativity. In I. Deliège & G. A. Wiggins (Eds.), *Musical creativity: Multidisciplinary research in theory and practice* (pp. 181–199). London, UK: Psychology Press.
- DELALANDE, F. (1988). La gestique de Gould: Éléments pour une sémiologie du geste musical. In G. Guertin (Ed.), *Glenn Gould pluriel* (pp. 85–111). Saint-Zénon, Quebec: Louise Courteau.
- DRIVER, J., & SPENCE, C. (1998). Attention and the crossmodal construction of space. *Trends in Cognitive Sciences*, 2, 254–262. [https://doi.org/10.1016/S1364-6613\(98\)01188-7](https://doi.org/10.1016/S1364-6613(98)01188-7)
- EEROLA, T., JAKUBOWSKI, K., MORAN, N., KELLER, P. E., & CLAYTON, M. (2018). Shared periodic performer movements coordinate interactions in duo improvisations. *Royal Society Open Science*, 5(2), 171520. <https://doi.org/10.1098/rsos.171520>
- FUJIOKA, T., TRAINOR, L. J., ROSS, B., KAKIGI, R., & PANTEV, C. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *Journal of Cognitive Neuroscience*, 17, 1578–1592. <https://doi.org/10.1162/089892905774597263>
- GIARD, M. H., & PERONNET, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11, 473–490. <https://doi.org/10.1162/089892999563544>
- GLOWINSKI, D., MANCINI, M., COWIE, R., CAMURRI, A., CHIORRI, C., & DOHERTY, C. (2013). The movements made by performers in a skilled quartet: A distinctive pattern, and the function that it serves. *Frontiers in Psychology*, 4, 841. <https://doi.org/10.3389/fpsyg.2013.00841>
- GOEBL, W., & PALMER, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception*, 26, 427–438. <https://doi.org/10.1525/mp.2009.26.5.427>
- GOLDIN, C., & ROUSE, C. (2000). Orchestrating impartiality: The impact of 'blind' auditions on female musicians. *American Economic Review*, 90, 715–741. <https://doi.org/10.1257/aer.90.4.715>
- GREGORY, A. H. (1990). Listening to polyphonic music. *Psychology of Music*, 18, 163–170. <https://doi.org/10.1177/0305735690182005>
- GRIFFITHS, N. K. (2008). The effects of concert dress and physical appearance on perceptions of female solo performers. *Musicae Scientiae*, 12, 273–290. <https://doi.org/10.1177/102986490801200205>
- GRIFFITHS, N. K. (2010). 'Posh music should equal posh dress': An investigation into the concert dress and physical appearance of female soloists. *Psychology of Music*, 38, 159–177.
- GRIFFITHS, N. K., & REAY, J. L. (2018). The relative importance of aural and visual information in the evaluation of western canon music performance by musicians and nonmusicians. *Music Perception*, 35, 364–375. <https://doi.org/10.1525/mp.2018.35.3.364>

- HAYHOE, M., & BALLARD, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9, 188–194. <https://doi.org/10.1016/j.tics.2005.02.009>
- HUANG, J., & KRUMHANSL, C. L. (2011). What does seeing the performer add? It depends on musical style, amount of stage behavior, and audience expertise. *Musicae Scientiae*, 15, 343–364. <https://doi.org/10.1177/1029864911414172>
- JENSENIUS, A. R., WANDERLEY, M. M., GODØY, R. I., & LEMAN, M. (2010). Musical gestures: Concepts and methods in research. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 12–35). Abingdon-on-Thames, UK: Routledge.
- JENSENIUS, A. R., ZELECHOWSKA, A., & GONZALEZ SANCHEZ, V. E. (2017). The musical influence on people's micromotion when standing still in groups. In T. Lokki, J. Pätynen, & V. Välimäki (Eds.), *Proceedings of the 14th Sound and Music Computing Conference 2017*, (pp. 195–200). Aalto University.
- JUCHNIEWICZ, J. (2008). The influence of physical movement on the perception of musical performance. *Psychology of Music*, 36, 417–427. <https://doi.org/10.1177/0305735607086046>
- KAWASE, S., & OBATA, S. (2016). Audience gaze while appreciating a multipart musical performance. *Consciousness and Cognition*, 46, 15–26. <https://doi.org/10.1016/j.concog.2016.09.015>
- KAYSER, C., PETKOV, C. I., LIPPERT, M., & LOGOTHETIS, N. K. (2005). Mechanisms for allocating auditory attention: An auditory saliency map. *Current Biology*, 15, 1943–1947. <https://doi.org/10.1016/j.cub.2005.09.040>
- KELLER, P. E., & APPEL, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception*, 28, 27–46. <https://doi.org/10.1525/mp.2010.28.1.27>
- KING, E., & GINSBORG, J. (2011). Gestures and glances: Interactions in ensemble rehearsal. In A. Gritten & E. King (Eds.), *New perspectives on music and gesture* (pp. 177–201). Farnham, UK: Ashgate. <https://doi.org/10.4324/9781315598048-17>
- KOCH, C., & ULLMAN, S. (1987). Shifts in selective visual attention: Towards the underlying neural circuitry. In L. M. Vaina (Ed.), *Matters of Intelligence: Conceptual Structures in Cognitive Neuroscience* (pp. 115–141). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-3833-5_5
- KOELEWIJN, T., BRONKHORST, A., & THEEUWES, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica*, 134, 372–384. <https://doi.org/10.1016/j.actpsy.2010.03.010>
- LIU, T., ABRAMS, J., & CARRASCO, M. (2009). Voluntary attention enhances contrast appearance. *Psychological Science*, 20, 354–362. <https://doi.org/10.1111/j.1467-9280.2009.02300.x>
- LOFTUS, G. R., & MACKWORTH, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 565–572. <https://doi.org/10.1037/0096-1523.4.4.565>
- MACRITCHIE, J., BUCK, B., & BAILEY, N. J. (2013). Inferring musical structure through bodily gestures. *Musicae Scientiae*, 17, 86–108. <https://doi.org/10.1177/1029864912467632>
- MALM, W. P. (2000). *Traditional Japanese music and musical instruments*. Tokyo, Japan: Kodansha International.
- MCGURK, H., & MACDONALD, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748. <https://doi.org/10.1038/264746a0>
- MEHR, S. A., SCANNELL, D. A., & WINNER, E. (2018). Sight-over-sound judgments of music performances are replicable effects with limited interpretability. *PLOS ONE*, 13(9), e0202075. <https://doi.org/10.1371/journal.pone.0202075>
- PALMER, C., & DEUTSCH, D. (2012). Music performance: Movement and coordination. In D. Deutsch (Ed.), *The perception of music* (pp. 405–422). London, UK: Elsevier.
- PALOMAR-GARCÍA, M.-Á., ZATORRE, R. J., VENTURA-CAMPOS, N., BUEICHEKÚ, E., & ÁVILA, C. (2017). Modulation of functional connectivity in auditory-motor networks in musicians compared with nonmusicians. *Cerebral Cortex*, 27, 2768–2778. <https://doi.org/10.1093/cercor/bhw120>
- PLATZ, F., & KOPIEZ, R. (2012). When the eye listens: A meta-analysis of how audio-visual presentation enhances the appreciation of music performance. *Music Perception*, 30, 71–83. <https://doi.org/10.1525/mp.2012.30.1.71>
- REMINGTON, R. W. (1980). Attention and saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 726–744. <https://doi.org/10.1037/0096-1523.6.4.726>
- RENSINK, R. A., O'REGAN, J. K., & CLARK, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8, 368–373. <https://doi.org/10.1111/j.1467-9280.1997.tb00427.x>
- ROSENHOLTZ, R. (1999). A simple saliency model predicts a number of motion popout phenomena. *Vision Research*, 39, 3157–3163. [https://doi.org/10.1016/S0042-6989\(99\)00077-2](https://doi.org/10.1016/S0042-6989(99)00077-2)
- RUNESON, S., & FRYKHOLM, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, 112, 585–615. <https://doi.org/10.1037/0096-3445.112.4.585>
- SCHUTZ, M., & LIPSCOMB, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, 36, 888–897.
- SELL, P., KANE, M. J., SMALLWOOD, J., SCHACTER, D. L., MAILLET, D., SCHOOLER, J. W., & SMILEK, D. (2018). Mind-wandering as a natural kind: A family-resemblances view. *Trends in Cognitive Sciences*, 22, 479–490. <https://doi.org/10.1016/j.tics.2018.03.010>

- SHULMAN, G. L., REMINGTON, R. W., & MCLEAN, J. P. (1979). Moving attention through visual space. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 522–526. <https://doi.org/10.1037/0096-1523.5.3.522>
- STEPHAN, M. A., LEGA, C., & PENHUNE, V. B. (2018). Auditory prediction cues motor preparation in the absence of movements. *NeuroImage*, 174, 288–296. <https://doi.org/10.1016/j.neuroimage.2018.03.044>
- THOMPSON, M. R., & LUCK, G. (2012). Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae*, 16, 19–40. <https://doi.org/10.1177/1029864911423457>
- THOMPSON, W. F., GRAHAM, P., & RUSSO, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, 2005(156), 203–227. <https://doi.org/10.1515/semi.2005.2005.156.203>
- THOMPSON, W. F., RUSSO, F. A., & LIVINGSTONE, S. R. (2010). Facial expressions of singers influence perceived pitch relations. *Psychonomic Bulletin and Review*, 17, 317–322. <https://doi.org/10.3758/PBR.17.3.317>
- TIMMERS, R., MAROLT, M., CAMURRI, A., & VOLPE, G. (2006). Listeners' emotional engagement with performances of a Scriabin étude: An explorative case study. *Psychology of Music*, 34, 481–510. <https://doi.org/10.1177/0305735606067165>
- TODOROV, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, 7, 907–915. <https://doi.org/10.1038/nn1309>
- TODOROV, E., & JORDAN, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5, 1226–1235. <https://doi.org/10.1038/nn963>
- TOMMASINI, A. (2003, November 10). A showman revs up the classical genre. *The New York Times*. <https://www.nytimes.com/2003/11/10/arts/music-review-a-showman-revs-up-the-classical-genre.html>
- TRAINOR, L. J., MARIE, C., BRUCE, I. C., & BIDELEMAN, G. M. (2014). Explaining the high voice superiority effect in polyphonic music: Evidence from cortical evoked potentials and peripheral auditory models. *Hearing Research*, 308, 60–70. <https://doi.org/10.1016/j.heares.2013.07.014>
- TSAY, C.-J. (2013). Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences*, 110, 14580–14585. <https://doi.org/10.1073/pnas.1221454110>
- TSAY, C.-J. (2014). The vision heuristic: Judging music ensembles by sight alone. *Organizational Behavior and Human Decision Processes*, 124, 24–33. <https://doi.org/10.1016/j.obhdp.2013.10.003>
- VINES, B. W., KRUMHANSL, C. L., WANDERLEY, M. M., DALCA, I. M., & LEVITIN, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, 118, 157–170. <https://doi.org/10.1016/j.cognition.2010.11.010>
- VINES, B. W., KRUMHANSL, C. L., WANDERLEY, M. M., & LEVITIN, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, 101, 80–113. <https://doi.org/10.1016/j.cognition.2005.09.003>
- VOLPE, G., D'AUSILIO, A., BADINO, L., CAMURRI, A., & FADIGA, L. (2016). Measuring social interaction in music ensembles. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693), 20150377. <https://doi.org/10.1098/rstb.2015.0377>
- VUOSKOSKI, J. K., THOMPSON, M. R., CLARKE, E. F., & SPENCE, C. (2014). Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception and Psychophysics*, 76, 591–604. <https://doi.org/10.3758/s13414-013-0582-2>
- WADDELL, G., & WILLIAMON, A. (2017). Eye of the beholder: Stage entrance behavior and facial expression affect continuous quality ratings in music performance. *Frontiers in Psychology*, 8, 513. <https://doi.org/10.3389/fpsyg.2017.00513>
- WANDERLEY, M. M., VINES, B. W., MIDDLETON, N., MCKAY, C., & HATCH, W. (2005). The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, 34, 97–113. <https://doi.org/10.1080/09298210500124208>
- WAPNICK, J., MAZZA, J. K., & DARROW, A. A. (2000). Effects of performer attractiveness, stage behavior, and dress on evaluation of children's piano performances. *Journal of Research in Music Education*, 48, 323–335. <https://doi.org/10.2307/3345367>
- WAPNICK, J., MAZZA, J. K., & DARROW, A.-A. (1998). Effects of performer attractiveness, stage behavior, and dress on violin performance evaluation. *Journal of Research in Music Education*, 46, 510–521. <https://doi.org/10.2307/3345347>
- WILLIAMON, A., & DAVIDSON, J. W. (2002). Exploring co-performer communication. *Musicae Scientiae*, 6, 53–72. <https://doi.org/10.1177/102986490200600103>
- ZATORRE, R. J., CHEN, J. L., & PENHUNE, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8, 547–558. <https://doi.org/10.1038/nrn2152>