

Forschungsberichte zum Themenschwerpunkt

Performers and an Active Audience: Movement in Music Production and Perception

Musizierende und ein aktives Publikum: Der Aspekt der Bewegung beim Ausüben und Wahrnehmen von Musik

Laura Bishop^{*a}, Werner Goebel^b

[a] Austrian Research Institute for Artificial Intelligence, Vienna, Austria. [b] Department of Music Acoustics, University of Music and Performing Arts, Vienna, Austria.

Abstract

Musical communication involves performance and perception processes, both of which engage the sensorimotor system. In much of the performance science literature, however, musical communication is conceptualized as a one-way trajectory from active performer to passive listener, minimizing the contribution of the listener and the collaborative nature of communication. In this paper, we discuss how movement contributes to 1) music performance, through sound production, interperformer coordination, and visual expressivity, and 2) music perception, through the simulation of observed gestures, activation of crossmodal associations, and induction of overt synchronized responses. Embodied music cognition, which treats musical communication as a process of dynamic interaction between individuals, and emphasizes the role of the physical body in mediating between environmental stimuli and subjective experiences, provides a background for our discussion. We conclude the paper with a discussion of how ongoing technological developments are simultaneously enhancing our ability to study musical communication (e.g., via integration of optical motion capture and mobile eye tracking) and, by introducing means of performing music that do not rely on human movement, challenging our understanding of how music and movement relate.

Keywords: music listening, perceptual-motor coupling, visual communication, ensemble coordination, embodied music cognition

Zusammenfassung

Musikalische Kommunikation umfasst Performanz- und Wahrnehmungsprozesse, die beide das sensomotorische System einbeziehen. In weiten Teilen der performanzwissenschaftlichen Literatur wird musikalische Kommunikation jedoch als ein nur in eine Richtung gehender Prozess vom aktiven Musizierenden zum passiven Zuhörenden konzipiert, die den Beitrag des Zuhörenden und die kollaborative Natur der Kommunikation in den Hintergrund stellt. In diesem Beitrag wird diskutiert, wie Bewegung auf der einen Seite beiträgt zu Klangproduktion, Koordination zwischen den Spielenden und visueller Ausdruckskraft, und auf der anderen zu Musikwahrnehmung durch die Simulation von beobachteten Gesten, Aktivierung von intermodalen Assoziationen und Induzierung von deutlichen, synchronisierten Reaktionen. Das Konzept der Embodied Music Cognition, welches musikalische Kommunikation als einen Prozess der dynamischen Interaktion zwischen Individuen behandelt und die Rolle des physischen Körpers bei der Vermittlung zwischen Umweltreizen und subjektiven Erfahrungen betont, bietet einen theoretischen Hintergrund für unsere Diskussion. Wir schließen unseren Beitrag mit einer Diskussion darüber, wie aktuelle technologische Entwicklungen gleichzeitig unsere Fähigkeit, musikalische Kommunikation zu beobachten, verbessern (z.B. durch Integration von optischer Bewegungserfassung und mobilem Eyetracking), aber das Aufkommen von Musikformen, welche nicht mehr auf menschliche Bewegung angewiesen sind (wie z.B. Computermusik), unser Verständnis von Musik und Bewegung auf eine neue Art herausfordert.

Schlüsselwörter: Musikhören, perzeptiv-motorische Kopplung, visuelle Kommunikation, Ensemble-Koordination, verkörperte Musikkognition

Jahrbuch Musikpsychologie, 2018, Vol. 28: Musikpsychologie — Musik und Bewegung, Artikel e19, doi:10.5964/jbdgm.2018v28.19

Eingereicht: 2017-12-19. Akzeptiert: 2018-03-16. Publiziert (VoR): 2018-08-13.

Begutachtet von: Clemens Wöllner; Birgitta Burger.

*Korrespondenzanschrift: Austrian Research Institute for Artificial Intelligence, Freyung 6/6, A-1010 Vienna, Austria. E-mail: laura.bishop@ofai.at



Dieser Open-Access-Artikel steht unter den Bedingungen einer Creative Commons Namensnennung 4.0 International Lizenz, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/deed.de>). Diese erlaubt für beliebige Zwecke (auch kommerzielle) den Artikel zu verbreiten, in jedwedem Medium zu vervielfältigen, Abwandlungen und Bearbeitungen anzufertigen, unter der Voraussetzung, dass der Originalartikel angemessen zitiert wird.

Introduction

Music performance takes many forms in our society, but usually involves trained and practiced musicians playing for an audience. In some musical traditions, conventions dictate silent and motionless listening behaviour from the audience; in other musical traditions, the audience is encouraged to move to or sing along with the music. Occasionally, music performance takes the form of a participatory activity that people do together as a group.

Music performance thus provides a venue for interaction between people – indeed, some scientists hypothesize that social bonding effects, in part, encouraged the widespread evolution of music-making abilities in early humans (Fitch, 2005; Huron, 2001; Tarr, Launay, & Dunbar, 2014). In the performance science literature, however, musical communication has often been conceptualized as a one-way trajectory from (active) performer to (passive) listener. Such a perspective fails to acknowledge that music-making is an inherently social process in which (live or prospective) audiences can influence performers' behaviour.

Music production and perception are active processes that engage overlapping perceptual-motor mechanisms, and movement forms a critical and inseparable part of both forms of musical experience. Our aim in this paper is to highlight the role of the sensorimotor system in music perception and the role of the audience in musical communication. We situate our discussion in the theoretical context of embodiment, which defines cognition as encompassing both internal processing and observable interactions with the world (Leman, 2012; Leman & Maes, 2014).

Following a discussion of the embodied music cognition paradigm, we consider the “musical product” (i.e., the presented performance) in terms of the movements – sound-producing, communicative, and expressive – that go into it. We explore the idea that movement not only underpins music production, but is also a part of the musical product itself (e.g., when it carries expressive and/or communicative information that supplements the audio signal). In the subsequent section, we discuss how movement underlies the audiovisual perception of performed music. We consider how audience members' prior experiences shape their perception of sound-producing movements, and how sounded music activates sensations of motion in listeners and, in some cases, encourages overt movement. We close with a discussion of how recent technological developments have simultaneously given us improved means to study musical communication and raised new questions for researchers to address.

Embodied Music Cognition and Perceptual-Motor Coupling

Embodied music cognition considers musical communication to be a nonlinear process characterized by dynamic interactions between performers, listeners, and their shared environment. The paradigm is in contrast to the “individualist” approach, which treats performers and listeners as separable from each other and from the musical stimulus, potentially understating the extent to which these three components interact (Moran, 2014). Some authors have argued that a focus on the Western art music tradition has enforced the individualist perspective, with its seemingly linear process of a composer writing a score, performers playing their interpretation of the score, and listeners interpreting the sounded music (Maes et al., 2014; Moran, 2014). In other musical traditions, the process of creating music can be more overtly collaborative and emergent (i.e., shaped dynami-

cally in real-time) – for example, group improvisation requires clear interaction between performers, and mother-infant lullaby singing involves clear bidirectional interaction between performer and listener, as the mother’s performance is shaped in real-time by the responses of her infant (see below; [Trevvarthen, 2012](#)).

Within the embodiment paradigm, there are diverging perspectives regarding the possible role of representational cognition in musical communication. Drawing on dynamical systems theory, some argue against the use of mental representations, and instead propose a framework in which musical interaction is dynamic, emergent, and autonomous ([Schiavio & Høffding, 2015](#)). Cognitive processes are said to be distributed across the collaborating musicians and their environment, rather than constrained within individuals, so “hidden” intentions – mental representations that are known only to individual performers – should not play a role. Those who take a less radical approach consider mental representations to be grounded in specific actions and bodily states. In this way, [Leman and Maes \(2014\)](#) describe the human body as a mediator linking people’s subjective experiences with their environment. The body mediates in both directions, helping to encode expressive ideas into sound as well as to decode expression from sound.

The research presented in this chapter is largely in line with this more moderate approach to embodiment. By this perspective, interaction with the environment occurs by way of action-perception loops. As also described in the literature on perceptual-motor coupling, actions are coded in the brain in terms of the consequences they have on the environment ([Hommel et al., 2001](#); [Maes et al., 2014](#); [Prinz, 1990](#)). Activation runs in both directions: executing an overt action activates expectations for the associated effect, while perceiving the effect primes the motor commands needed to produce it. Multiple levels of action-perception loops may run in parallel ([Leman, 2012](#)). Low-level loops, in which sensory input drives motor activity, may allow regulation of some aspects of performance technique (e.g., posture, breathing, finger position), while high-level loops, which draw on a repertoire of “gestures” (i.e., meaning-carrying movements) may facilitate planning of expressive contours.

The embodiment paradigm has been criticized for being overly broad and poorly defined – while its arguments are widely consistent with findings in the literature, it does not establish hypotheses specific enough to be empirically validated by disproving the alternative (disembodied) explanation ([Matyja, 2016](#)). However, it is a useful perspective to take in the context of the current paper, as it encourages us to reconsider the usual conceptual division between perception and performance and examine what the audience contributes to the process of music creation.

Movement as a Part of Music Production

Traditionally, body movement has been necessary for musical sound production, but musicians’ gestures serve other functions too: they enable changes in tone quality, facilitate coordination between ensemble members, convey expressive information to the audience, and support the performer’s own affective experience ([Bishop & Goebel, 2018](#); [Dahl et al., 2010](#)). All categories of gesture form a part of the musical product that is presented to an audience, in the event that the performance can be viewed as well as heard. In the following sections, we discuss three categories of movements: those involved in sound production, those that support ensemble coordination, and those that communicate expression visually.

Sound Production

A central aim of the research on instrumental playing technique is to determine how variations in acoustic parameters are controlled by the performer. At a cognitive level, some musicians report focusing on the image of a desired sound as they play, which they say helps to guide their performance (Trusheim, 1991). The results of empirical studies support this claim. As explained above, perception of a sound is presumed to facilitate activation of the motor commands necessary to produce that sound. Imagining the sound has been shown to have a similar facilitatory effect on body movements (Bishop, Bailes, & Dean, 2013; Keller, Dalla Bella, & Koch, 2010).

The motor commands used to externalize these images and deliver musical output have been studied in increasing depth as technology for capturing fine movements has improved (Goebel, 2017). How skilled musicians maintain rhythmic consistency has been a primary focus in this line of research. In a study of piano performance, Tominaga et al. (2016) used a glove fitted with sensors to show how rotational velocity at finger joints and movement independence between fingers underlie rhythmic consistency. In saxophone performance, maintaining rhythmic consistency involves achieving precise coordination of finger and tongue movements. Using newly-developed sensor-equipped reeds, Hofmann and Goebel (2014) found higher consistency for coupled finger-tongue movements than for tongue-only (isotonal) movements at a slow tempo, but higher consistency for tongue-only movements at a fast tempo.

How skilled musicians manipulate timbre and dynamics has also been a focus of study. In an investigation of piano timbre, Goebel, Bresin, and Fujinaga (2014) used accelerometers to capture the kinematics of piano keys and hammers as they were played with “struck” and “pressed” finger movements. Pianists claim that these movements produce different tone qualities, and, indeed, listeners were able to discriminate between them even when key velocity was held constant. The points at which the striking finger hit the key and key bottom, identified using the accelerometer data, were found to provide acoustic cues that facilitated discrimination between timbres. In violin playing, timbre seems to be largely controlled through manipulations of bow force. This was determined by Schoonderwaldt (2009a), who used a bowing machine to measure the influence of bow force, velocity, and bow-bridge distance on intonation and spectral centroid, an acoustic correlate of perceived timbre. To achieve different dynamic levels, violinists manipulate bow force, bow-bridge distance and bow angle simultaneously (Schoonderwaldt, 2009b).

Interperformer Coordination

Ensemble musicians aim primarily to coordinate their sound. Sometimes they deliberately coordinate their sound-producing gestures as well – for example, orchestral string musicians typically use coordinated bowing patterns. More often, though, the sounds that must be coordinated are the result of different types of sound-producing gestures, which require different attack durations. Some coordination of expressive gestures (e.g., body sway) also occurs (Keller, Dalla Bella, & Koch, 2010; Moran et al., 2015), though it is unclear whether this is intended (or noticed) by performers, or simply a byproduct of performers sharing an interpretation of the music. Also unclear is the extent to which coordination of ancillary movements shapes the audience’s perception of ensemble coherence.

To coordinate their sound, ensemble musicians do not generally need to be able to see each other’s movements. Both trained and novice musicians synchronize with sounded rhythms in the absence of visual cues, even when the sounded rhythm contains irregularities and error-correction is needed to maintain synchroniza-

tion (Hove et al., 2013; Konvalinka et al., 2010; van der Steen et al., 2015). Nevertheless, musicians do look towards each other when coordinating a performance, particularly at moments where their interpretations are likely to diverge, and uncertainty about each other's intentions is high – for instance, at the start of a piece, following a long pause, or at sudden tempo or metrical changes (Bishop & Goebel, 2015; Kawase, 2014).

Some recent studies of ours have considered the visual signals that are exchanged at piece onset (Bishop & Goebel, 2017a, 2017b). In one study, we presented audiovisual point-light recordings of pianists' cueing-in gestures to musician viewers, who were instructed to tap in synchrony with the beat of the music, aligning their first taps with the pianists' first onsets (Bishop & Goebel, 2017b). Viewers synchronized most successfully with gestures that were low in jerk and large in magnitude. They also synchronized more successfully with gestures given by experienced ensemble pianists than with gestures given by pianists lacking ensemble experience.

Conductors' gestures have been subjected to similar analysis. A study by Wöllner et al. (2012) showed that observers synchronize more successfully with gestures that are low in jerk and high in prototypicality than with gestures that are high in jerk and low in prototypicality. In conductors' as well as instrumentalists' gestures, beats seem to correspond to peaks in acceleration, rather than specific points in the spatial trajectory (Bishop & Goebel, 2017a, 2017b; Luck & Sloboda, 2009; Wöllner et al., 2012).

Of course, it is relatively infrequent that musicians make such explicit use of visual signalling as we describe here. It has been proposed that ensemble coordination is largely supported by an exchange of low-level sensory information (e.g., psychoacoustic sound parameters and movement kinematics) that induces entrainment between performers (Pachet, Roy, & Foulon, 2017). Musicians presumably draw more or less on these higher- and lower-level processes at any given moment, depending on the performance conditions (MacRitchie, Varlet, & Keller, 2017).

Visual Expressivity

Musicians' body movements are of substantial communicative value to an observing audience. The question of how visual cues in the form of observed body gestures contribute to our perception of music performance has generated some debate among researchers and, of course, among musicians, whose primary focus is the sound their movements produce, and who are not always pleased to think that the visual modality could have a prevailing impact on their audience's experience.

One line of research in this area tests the hypothesis that auditory and visual contributions to the perception of music expression are integrated, rather than additive. In a study by Vuoskoski et al. (2014), recordings of pianists performing with deadpan, normal, or exaggerated levels of expression were used. Audio and visual tracks were recombined (e.g., normal audio paired with deadpan video) and presented to trained and novice musicians, who provided ratings of auditory expressivity. Visual stimuli affected ratings of auditory expressivity even when matched with incongruous audio, though the magnitude of the effect differed between performing pianists. These results suggest a degree of malleability in auditory and visual cues that affects how much influence one modality has over the other. Similar results were subsequently obtained from an experiment testing emotional impact instead of perceived auditory expressivity (Vuoskoski et al., 2016).

The magnitude of the effect that visual cues can have on even a highly-trained audience's perception of expressivity was demonstrated by Behne and Wöllner (2011). Musician participants were presented with audiovi-

sual recordings of different pianists' performances of the same two pieces and asked to provide expressivity ratings. The recordings that were presented used the same audio track but different video for the different performers, who had actually played in sync with the same pre-recorded audio. Ratings differed between the recordings, despite the audio being identical. High interrater variability was also observed, suggesting that the integration of auditory and visual cues could be an idiosyncratic process.

Tsay (2013) tested professional musicians and untrained listeners' abilities to identify the winners of piano competitions, given audio, visual, or audiovisual presentation of performance excerpts. Both groups of participants made the most accurate judgments when presented with silent video excerpts, suggesting that visual cues might be even more informative than audio cues for discriminating competition winners. On the other hand, the apparent visual advantage could be attributable to a biased selection of stimulus excerpts (Platz et al., 2016). Further study using more systematically selected stimuli would be needed to confirm this surprising finding.

The research we describe here shows that the audience's perception of a musical performance derives from an integration of auditory and visual kinematic cues, and it highlights the audience's role in assigning meaning to the performance. In the next section, we consider how movement can be perceived through musical sound as well as visually.

Movement as a Part of Music Perception

Musical communication is a creative, dynamic process comprising performance and perception components. In the previous section, we considered how body movements underlie performers' contributions to musical communication, and we must acknowledge that the audience's feedback, whether real or imagined, helps to shape those movements. When engaging with a musical performance, audience members draw on their own abilities, constraints, and experiences to construct some meaning from the performers' audiovisual signals. As such, they can be considered active contributors to the creative process of musical communication.

The term "communicative musicality" has been used to describe how coordinated companionship arises from the temporal and expressive attributes of social behaviour (Trevarthen, 2012). Three attributes are defined: pulse, the regular patterning of a performer's output through time that allows the audience to anticipate what might follow; quality, the expressive contours of sound and body gestures; and narrative, the linking of pulse and quality into units that allow the performer and audience to share a sense of passing time. The application of this paradigm to the study of mother-infant musical interaction has shown how infants respond behaviourally to the timing and affective quality of their mothers' singing (e.g., with body rocking, timed verbal utterances, and facial expressions), and how the mothers' behaviour is shaped by evidence of their infants' engagement. The communicative behaviour observed between mothers and infants provide clear examples of how listeners' body movements can both facilitate their own understanding of a musical performance and provide feedback to the performer.

The embodied music cognition paradigm posits that movement underlies music perception just as it underlies performance. In this section, we discuss 1) how audience members' perceptions of sound-producing movement are shaped by their prior experience, 2) how movement is "heard" in sounded music through associations of motion and acoustic parameters, and 3) why music sometimes prompts listeners to move.

Audiovisual Perception of Human Motion: Effects of Experience

Audience members become active participants in a music performance the moment the sounded and/or observed music enters their perceptual systems. If the performance can be seen and heard, a critical part of the perception process is the binding of auditory and visual signals into distinct perceptual events that correspond to sound-producing gestures and their acoustic effects. This process of audiovisual integration draws on action-perception loops that strengthen with exposure to different gestures and their associated effects. The more experience a person has with a repertoire of gesture-sound pairs, the more precise their gesture-sound associations become. Strengthened associations have been observed for pitch (Keller & Koch, 2008), dynamics and articulation (Bishop, Bailes, & Dean, 2013), and timing. In particular, tolerance for audiovisual asynchrony in musical stimuli has been shown to decrease with increasing musical expertise (Petrini et al., 2009).

Tolerance for asynchrony or incongruity in perceived gesture-sound pairs also depends on the type of motion observed and the type of sound produced. Less asynchrony is tolerated for piano playing than for (bowed) violin playing (Bishop & Goebel, 2014). Piano playing involves percussive gestures and yields sounds for which the physical onset (i.e., the point where the sound begins) and the perceptual onset (i.e., the point where the sound becomes perceptible to the listener) are virtually simultaneous. In contrast, violin playing involves continuous gestures and yields sounds for which a small interval between the physical and perceptual onsets can exist, and a range of perceptual onsets might be tolerated (Vos & Rasch, 1981).

Strengthening of action-perception loops occurs with both perceptual and motor experience. For example, among pianists, listening practice (without overt movement) results in better recall (i.e., performance) of simple melodies than does motor practice without sound (Brown & Palmer, 2013). Such an effect shows how the action-perception loops drawn upon during performance are strengthened by listening experience.

On the other hand, the learning benefits of combined perceptual-motor experience has been shown to outweigh the learning benefits of perceptual experience without overt movement. Aglioti et al. (2008) found that skilled basketball players predicted the success of free shots at a basket faster and more accurately than did either coaches or sports journalists, who had extensive visual but limited motor experience. In the music domain, Wöllner and Cañal-Bruland (2010) found that skilled string musicians predict note onset times from violinists' cueing gestures more precisely than do skilled non-string musicians.

These findings suggest that audience members draw on movement in the form of action-perception loops during the early stages of interpreting perceived music, when the binding of audio and visual signals occurs. As discussed above, the way audio and visual signals combine has a potentially strong influence over audience members' perceptions of expression. The learning that occurs with observation of others' performance, even if it occurs to a lesser extent than with overt practice, shows how the perceptual-motor system is tuned to change in a way that facilitates prediction abilities and supports multisensory associations.

Hearing Movement in Sound

Cross-modal correspondences are symmetric associations that people make between parameters in different sensory modalities. Associations between pitch height and spatial height, for example, are widespread and seemingly independent of musical training and linguistic background (Eitan, 2017). Cross-modal correspondences

ces can modulate overt movement, speeding motor responses when the sounded stimulus and target movement are “correctly” matched (Keller & Koch, 2008).

Eitan and Granot (2006) tested for associations between music and motion. Listeners were presented with auditory sequences in which only one parameter was varied (e.g., rising or falling pitch, increasing or decreasing loudness), and imagined a human figure moving with the music. Their descriptions of the imagined figures’ movements suggested a number of correspondences, including an association between increasing/decreasing loudness and decreasing/increasing distance change, and an association between pitch contour and vertical position change.

Some of the features that people associate with acoustic parameters they also associate with emotional constructs. For example, some positively-valenced words (e.g., happy) are associated with a high spatial position, while their antonyms (e.g., sad) are associated with a low spatial position (Eitan & Timmers, 2010; Gozli et al., 2013). A study by Weger et al. (2007) showed how associations between emotional valence, spatial height, and acoustic pitch can interrelate: positively- and negatively-valenced words (e.g., kiss, dead) were found to prime judgements of the pitch height of sounded tones. Eitan (2017) suggests that emotion may mediate some acoustic-visuospatial associations, such as the association of spatial height with pitch.

Taking a different perspective, the “FEELA” (Force-Effort-Energy-Loudness-Arousal) hypothesis relates affective parameters of music (e.g., arousal) to the corresponding acoustic parameters (e.g., acoustic intensity) and parameters of the movement needed to produce the underlying sound (Olsen & Dean, 2016). In a recent study, groups of listeners were presented with passages of classical and electroacoustic music and gave continuous judgements of perceived physical exertion, arousal, and valence. For passages with music that could be readily attributed to human movement (the classical and some of the electroacoustic pieces), perceived exertion was a significant predictor of perceived arousal and valence. For the passages of electroacoustic music that could not be attributed to human movement, exertion judgements did not seem to influence the perception of arousal. Acoustic intensity, in turn, was a significant predictor of perceived exertion. Thus, listeners seem to hear physical exertion in music that they associate with human sound-producing movements, and this supports the perception of arousal.

Studies of cross-modal correspondences suggest that part of the meaning an audience gets from perceived music can come from the associations they make between acoustic and motion parameters. While hearing movement in music is itself potentially meaningful, associations with emotional constructs may additionally contribute. Such findings are in line with the proposed role of the human body as a mediator between subjective experience and the environment that is involved in the construction of musical meaning (Leman & Maes, 2014).

Inducing Movement Through Sounded Music

Some music induces a sense of movement in listeners, inciting in both trained and novice musicians an urge to synchronize with the beat (Janata, Tomic, & Haberman, 2012). This psychological phenomenon is referred to as groove. Not all music creates a sense of groove. Ratings of groove strength are consistently lower for some genres of music (e.g., folk) than others (e.g., soul/R&B; Janata et al., 2012). The perception of groove is thought to relate to microtiming, the small deviations from metronomic timing that occur even in metrical, beat-based music; however, this notion has proven difficult to validate experimentally. Several studies have shown

quantized music to receive high ratings of groove, and that these ratings decline as the magnitude of microtiming deviations increases (Hofmann, Wesolowski, & Goebel, 2017; Senn et al., 2017).

Other parameters have been shown to contribute to the perception of groove for both trained and novice musicians. Witek et al. (2014) showed that a moderate degree of syncopation maximizes both perceived groove and reported enjoyment of groove-based music. Other parameters that encourage perceptions of groove include a tempo near to the average frequency of human locomotion, the use of low frequencies on bass instruments, and spectral flux (a measure of rate of change in the frequency spectrum that relates to perceived activity level) in the low frequencies (Stupacher, Hove, & Janata, 2016). Ratings of groove have also been found to increase when the entrances of different voices are staggered rather than simultaneous, and when the music contains more rather than fewer instruments (Hurley, Martens, & Janata, 2014).

The results of the study by Hurley et al. (2014) show that some of the acoustic features likely to grab listeners' attention (e.g., staggered entrances) also increase their tendency for stimulus-coupled movement. Increased attention facilitates processing of auditory signals, and could potentially increase activity of action-perception loops, encouraging overt movement. Burger et al. (2013) found that some characteristics of musical timing and timbre relate consistently to characteristics of music-induced movement. For example, increased pulse clarity encouraged a wide variety of whole body movements, and low frequency spectral flux correlated positively with head velocity. Rhythmic information in beat-based music is primarily communicated through low frequency voices, so the correlation with head velocity might reflect a tendency to pair head movements with beats. Synchronization with a sounded beat has been shown to improve timing perception (Manning & Schutz, 2013), and as an obvious signal of engagement, it could also have social bonding effects if observed by the performers or other audience members.

Future Directions

In recent years, we have seen the development of sensor and camera systems that measure musicians' movements and audience perceptions of them in great detail. Simultaneously, with the advent of technology-mediated performance, computers have been playing an increasing role in human musical performance, disrupting the traditional relationship between musical sound and movement (Emerson & Egermann, 2018). Today, we can listen to instrumental music that has been digitally enhanced, resulting in sound that is not entirely attributable to observable gestures. We can also see performances on digital musical interfaces for which the gesture-sound mapping is unfamiliar to us or even invisible (e.g., if sound output is generated via laptop controls). In this section, we discuss how these technological developments have the potential to guide our future research endeavours, both by introducing new methods for the study of human interaction and by raising new questions about how music is treated by the perceptual system.

Integrating Capture Techniques to Quantify Visual Interaction

The techniques available for studying music-related movements include sensors capable of making fine-grained measurements of movement parameters that are not readily apparent to an external viewer, such as finger forces in pianists (Kinoshita et al., 2007) or tonguing patterns in saxophonists (Hofmann & Goebel, 2014). Larger-scale movements, from finger trajectories to body sway, can be assessed using inertial or optical motion capture. A number of techniques are also available for assessing perception of musicians' gestures, including

eye tracking, brain imaging, and EMG (for measuring covert muscle activation). In this section, we focus on two categories of techniques – motion capture and eye tracking – that have particular potential for research on the role of movement in musical communication.

Inertial and optical motion capture systems are widely used in the study of performance gestures (Goebel, Dixon, & Schubert, 2014). Inertial sensors, including accelerometers and gyroscopes, are typically affixed to a musician's body or instrument. Accelerometers track 3D acceleration and gyroscopes track orientation and angular velocity at the measured point. Optical motion capture uses cameras to triangulate the position of markers that are affixed to musicians' bodies or instruments, and has been widely used in studies investigating the expressive gestures of performing musicians (Nusseck, Wanderley, & Spahn, 2017; Vuoskoski et al., 2014), communicative gestures in music ensembles (Bishop & Goebel, 2017b), and motor responses to perceived music (Toiviainen, Luck, & Thompson, 2010). Motion capture recordings are also commonly used as visual stimuli in experiments investigating viewers' perceptions of human movement (e.g., Bishop & Goebel, 2017b; Moran et al., 2015; Petrini et al., 2009; Wöllner et al., 2012).

Most eye tracking systems use infrared cameras to detect corneal reflections around the pupil and measure its movements. Eye tracking systems can be remote or mobile: remote systems are typically mounted to a computer screen and suitable for monitoring gaze directed towards a stationary stimulus (e.g., text or images), while mobile systems are mounted on the subject's head and can be used to study gaze behaviour in a 3D environment (e.g., a performance space). Both remote and mobile eye trackers calculate measures such as pupil position, point of regard, and pupil dilation. Mobile eye tracking is a useful technique for studying gaze behaviour in performing musicians, who typically require more freedom of movement than remote eye trackers can cope with. Bigand et al. (2010) used mobile eye tracking to investigate the gaze behaviour of an orchestral conductor, and in our lab, we use mobile eye trackers to examine ensemble musicians' attention towards each other (Bishop, Cancino-Chacón, & Goebel, 2017).

There are ongoing efforts by several research groups, including ours, to integrate mobile eye tracking and optical motion capture (Burger, Puupponen, & Jantunen, 2017). One of the critical steps involved in integrating these systems is establishing a method for synchronizing the different data streams. In our lab, we use a synchronization device (issued by the makers of our motion capture system) to send TTL triggers to the eye tracking software at the start and stop of each motion capture recording. This method ensures precise and reliable synchronization and allows us to check for drift between data streams. However, it does introduce some constraints to the recording set-up, as the glasses have to be connected via (customized) cable to a computer capable of receiving TTL triggers, somewhat restricting performers' freedom of movement.

An integrated motion capture-eye tracking system greatly simplifies the analysis of eye gaze data. With mobile eye tracking, extensive manual coding of video data is typically required, since in contrast to remote eye tracking, the visual scene is constantly changing and areas of interest are not static. If mobile eye trackers are used in combination with motion capture, however, detection of subjects' gaze targets can be automatized by remapping gaze coordinates into the motion capture coordinate system. Moments where the gaze target is the musical score or another performer (i.e., an object or person defined with markers) are then readily identifiable.

Complementing such an integrated system, and further facilitating the study of body movements and gaze behaviour, would be the development of advanced data analysis tools. Currently, MATLAB users can access the Mocap Toolbox developed by Burger and Toiviainen (2013), which provides functions for analyzing and visualiz-

ing movement data. It would be widely beneficial to develop such tools to include functions for the analysis of gaze data captured using an integrated motion capture-eye tracking system. When applied to the study of human interaction, an integrated motion capture-eye tracking system allows precise identification of which body movements people attend to and the extent to which gaze itself acts as a visual cue. With visual attention and body movements of multiple people (e.g., performer and listener or performer and co-performer) captured in parallel, musical communication can be studied from both ends of the loop simultaneously.

Challenging the Inseparability of Movement and Music

Ongoing changes in the way music is created, distributed, and experienced by audiences challenge the understanding we have of embodied music perception. For example, while people do attend live concerts on occasion, most of the music they encounter on a daily basis enters the perceptual system unimodally, as an auditory signal without corresponding visual cues. The pervasiveness of this unimodal auditory presentation raises some questions: when people hear music without knowledge of the movements needed to produce it, is their perception less embodied than it would have been, had they some familiarity with those movements? Earlier, we discussed some perceptual sub-processes that draw on listeners' motor resources. Some of these sub-processes, such as the ability to predict the sounded outcomes of performer gestures and the integration of audio and visual signals, are enhanced by visual experience. On the other hand, cross-modal correspondences are thought to reflect learned associations between commonly co-occurring events and sounds that develop through a combination of general (not music-specific) statistical learning and familiarity with cultural/linguistic conventions (Eitan, 2017). Thus, visual exposure to performance may be unnecessary for listeners to "hear" movement in sounded music.

Music listening today is often a solitary activity, done alone, over headphones, with no possibility of observing the performers. Launay (2015) questioned how we should reconcile this with the hypothesized origins of music-making as a means of social bonding. He suggested that the ability to hear movement in sounded music, which is prompted by the presence of either rhythm or familiar instrumental sounds (associated in memory with certain gestures), allows listeners to infer the existence of a performer, and making music listening a social activity even in minimally interactive conditions (e.g., solo listening to an audio track).

Changes to the way music is created include the development of algorithms capable of performing music and the introduction of digital musical interfaces (DMIs) that human performers can use to create and manipulate digital music in new and creative ways. Some of the sounds produced via these methods, including machine-like or environmental sounds, are not likely to be attributed to human movements by our perceptual systems. Research has already shown that audience members' aesthetic judgments of DMI performance suffer when they are unable to figure out the gesture-sound mappings of the interface (Emerson & Egermann, 2018). Audiences also show an inability to pair audio and visual signals when gesture-sound causality is not perceptible, giving equivalent ratings of performance quality for correctly paired audio-video excerpts and incorrectly paired excerpts that combine audio and video from different parts of the performance.

An outstanding question is whether music comprised of such sounds is, like music comprising sounds that result from human movements, experienced in an embodied way. The results of the study by Olsen and Dean (2016), discussed above, suggest that features of electroacoustic music not attributable to human movement can be associated with movement parameters, although in contrast to music that is attributable to human

movement, the association may not influence listeners' perception of arousal. In contrast to this finding, electronic dance music and hip-hop turntable performances tend to encourage both overt movement and high arousal in their audiences, showing that music comprising digital sounds can engage listeners motorically and emotionally. Presumably, the organization of digital sounds into distinct beat-based rhythms results induces motor activation in listeners, regardless of whether they associate the sounds with causal gestures.

Conclusions

The aim of this paper was to show that musical communication is a dynamic and collaborative process involving performers and an active audience, for whom music perception is a motoric process. Traditionally, overt movement has been critical for performed music, necessary for sound production as well as a part of the musical product presented to audiences. Today, this is no longer entirely the case, since 1) the audience does not always see the performance and 2) sounded music can be generated without sound-producing gestures. Research findings are in line with the prediction that music perception is embodied: motor resources are drawn upon throughout the perception process and while constructing meaning out of the musical signal. At present, however, the literature still lacks strong tests of the embodiment paradigm, as well as an indication of how robust current findings are to a broader definition of "music" that includes genres outside the Western art music tradition, including genres in which the potential for interpersonal interaction is higher (e.g., group improvisation) and genres in which it is lower (e.g., electroacoustic music).

Funding

This research was funded by Austrian Science Fund grant P29427.

Competing Interests

The authors have declared that no competing interests exist.

Acknowledgments

The authors have no support to report.

References

- Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, *11*, 1109-1116. <https://doi.org/10.1038/nn.2182>
- Behne, K. E., & Wöllner, C. (2011). Seeing or hearing the pianists? A synopsis of an early audiovisual perception experiment and a replication. *Musicae Scientiae*, *15*(3), 324-342. <https://doi.org/10.1177/1029864911410955>
- Bigand, E., Lalitte, P., Lerdahl, F., Boucheix, J.-M., Gérard, Y., & Pozzo, T. (2010). Looking into the eyes of a conductor performing Lerdahl's "Time after Time". *Musicae Scientiae*, *14*(2, Suppl.), 275-294. <https://doi.org/10.1177/10298649100140S215>
- Bishop, L., Bailes, F., & Dean, R. T. (2013). Musical imagery and the planning of dynamics and articulation during performance. *Music Perception*, *31*(2), 97-117. <https://doi.org/10.1525/mp.2013.31.2.97>

- Bishop, L., Cancino-Chacón, C. E., & Goebel, W. (2017). *Mapping visual attention during duo performance of temporally-ambiguous music*. Paper presented at the International Symposium on Performance Science, Reykjavik, Iceland.
- Bishop, L., & Goebel, W. (2014). Context-specific effects of musical expertise on audiovisual integration. *Frontiers in Psychology*, 5, Article 1123. <https://doi.org/10.3389/fpsyg.2014.01123>
- Bishop, L., & Goebel, W. (2015). When they listen and when they watch: Pianists' use of nonverbal audio and visual cues during duet performance. *Musicae Scientiae*, 19(1), 84-110. <https://doi.org/10.1177/1029864915570355>
- Bishop, L., & Goebel, W. (2018). Beating time: How ensemble musicians' cueing gestures communicate beat position and tempo. *Psychology of Music*, 46(1), 84-106. <https://doi.org/10.1177/0305735617702971>
- Bishop, L., & Goebel, W. (2017a). Communication for coordination: Gesture kinematics and conventionality affect synchronization success in piano duos. *Psychological Research*. Advance online publication. <https://doi.org/10.1007/s00426-017-0893-3>
- Bishop, L., & Goebel, W. (2017b). Music and movement: Musical instruments and performers. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition*. New York, NY, USA: Routledge.
- Brown, R. M., & Palmer, C. (2013). Auditory and motor imagery modulate learning in music performance. *Frontiers in Human Neuroscience*, 7, Article 320. <https://doi.org/10.3389/fnhum.2013.00320>
- Burger, B., Puupponen, A., & Jantunen, T. (2017). *Synchronizing eye tracking and optical motion capture: How to bring them together*. Paper presented at the Conference on Music and Eye-Tracking, Frankfurt, Germany.
- Burger, B., Thompson, M. R., Luck, G., Saarikallio, S., & Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology*, 4, Article 183. <https://doi.org/10.3389/fpsyg.2013.00183>
- Burger, B., & Toiviainen, P. (2013). *MoCap Toolbox – A Matlab toolbox for computational analysis of movement data*. Paper presented at the Proceedings of the Sound and Music Computing Conference 2013, Stockholm, Sweden.
- Dahl, S., Bevilacqua, F., Bresin, R., Clayton, M., Leante, L., Poggi, I., & Rasamimanana, N. (2010). Gestures in performance. In R. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 36-68). London, United Kingdom: Routledge.
- Eitan, Z. (2017). Musical connections: Cross-modal correspondences. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* (pp. 213-224). New York, NY, USA: Routledge.
- Eitan, Z., & Granot, R. Y. (2006). How music moves: Musical parameters and listeners' images of motion. *Music Perception*, 23(3), 221-247. <https://doi.org/10.1525/mp.2006.23.3.221>
- Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition*, 114, 405-422. <https://doi.org/10.1016/j.cognition.2009.10.013>
- Emerson, G., & Egermann, H. (2018). Gesture-sound causality from the audience's perspective: Investigating the aesthetic experience of performances with digital musical instruments. *Psychology of Aesthetics, Creativity, and the Arts*, 12(1), 96-109. <https://doi.org/10.1037/aca0000114>

- Fitch, W. T. (2005). The evolution of music in comparative perspective. *Annals of the New York Academy of Sciences*, 1060, 29-49. <https://doi.org/10.1196/annals.1360.004>
- Goebel, W. (2017). Movement and touch in piano performance. In B. Müller et al. (Eds.), *Springer handbook of human motion* (pp. 1-18). Berlin, Germany: Springer International.
- Goebel, W., Bresin, R., & Fujinaga, I. (2014). Perception of touch quality in piano tones. *The Journal of the Acoustical Society of America*, 136(5), 2839-2850. <https://doi.org/10.1121/1.4896461>
- Goebel, W., Dixon, S., & Schubert, E. (2014). Quantitative methods: Motion analysis, audio analysis, and continuous response techniques. In D. Fabian, R. Timmers, & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (p. 221-239). Oxford, United Kingdom: Oxford University.
- Gozli, D. G., Chow, A., Chasteen, A. L., & Pratt, J. (2013). Valence and vertical space: Saccade trajectory deviations reveal metaphorical spatial activation. *Visual Cognition*, 21(5), 628-646. <https://doi.org/10.1080/13506285.2013.815680>
- Hofmann, A., & Goebel, W. (2014). Production and perception of legato, portato and staccato articulation in saxophone playing. *Frontiers in Psychology*, 5, Article 690. <https://doi.org/10.3389/fpsyg.2014.00690>
- Hofmann, A., Wesolowski, B. C., & Goebel, W. (2017). The tight-interlocked rhythm section: Production and perception of synchronization in jazz trio performance. *Journal of New Music Research*, 46(4), 329-341. <https://doi.org/10.1080/09298215.2017.1355394>
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24, 849-937. <https://doi.org/10.1017/S0140525X01000103>
- Hove, M. J., Fairhurst, M. T., Kotz, S. A., & Keller, P. (2013). Synchronizing with auditory and visual rhythms: An fMRI assessment of modality differences and modality appropriateness. *NeuroImage*, 67, 313-321. <https://doi.org/10.1016/j.neuroimage.2012.11.032>
- Hurley, B. K., Martens, P. A., & Janata, P. (2014). Spontaneous sensorimotor coupling with multipart music. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1679-1696. <https://doi.org/10.1037/a0037154>
- Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences*, 930, 43-61. <https://doi.org/10.1111/j.1749-6632.2001.tb05724.x>
- Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, 141(1), 54-75. <https://doi.org/10.1037/a0024208>
- Kawase, S. (2014). Assignment of leadership role changes performers' gaze during piano duo performances. *Ecological Psychology*, 26(3), 198-215. <https://doi.org/10.1080/10407413.2014.929477>
- Keller, P. E., Dalla Bella, S., & Koch, I. (2010). Auditory imagery shapes movement timing and kinematics: Evidence from a musical task. *Journal of Experimental Psychology: Human Perception and Performance*, 36(2), 508-513. <https://doi.org/10.1037/a0017604>
- Keller, P. E., & Koch, I. (2008). Action planning in sequential skills: Relations to music performance. *Quarterly Journal of Experimental Psychology*, 61(2), 275-291. <https://doi.org/10.1080/17470210601160864>

- Kinoshita, H., Furuya, S., Aoki, T., & Altenmüller, E. (2007). Loudness control in pianists as exemplified in keystroke force measurements on different touches. *The Journal of the Acoustical Society of America*, 121(5), 2959-2969. <https://doi.org/10.1121/1.2717493>
- Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology*, 63(11), 2220-2230. <https://doi.org/10.1080/17470218.2010.497843>
- Launay, J. (2015). Musical sounds, motor resonance, and detectable agency. *Empirical Musicology Review*, 10(1-2), 30-40. <https://doi.org/10.18061/emr.v10i1-2.4579>
- Leman, M. (2012). Musical gestures and embodied cognition. In *Journées d'Informatique Musicale: Proceedings* (pp. 5-7). Mons, Belgium: University of Mons.
- Leman, M., & Maes, P.-J. (2014). The role of embodiment in the perception of music. *Empirical Musicology Review*, 9(3-4), 236-246. <https://doi.org/10.18061/emr.v9i3-4.4498>
- Luck, G., & Sloboda, J. A. (2009). Spatio-temporal cues for visually mediated synchronization. *Music Perception*, 26(5), 465-473. <https://doi.org/10.1525/mp.2009.26.5.465>
- MacRitchie, J., Varlet, M., & Keller, P. E. (2017). Embodied expression through entrainment and co-representation in musical ensemble performance. In M. Lesaffre, P. J. Maes, & M. Leman (Eds.), *The Routledge companion to embodied music interaction* (pp. 150-159). New York, NY, USA: Routledge.
- Maes, P.-J., Leman, M., Palmer, C., & Wanderley, M. M. (2014). Action-based effects on music perception. *Frontiers in Psychology*, 4, Article 1008. <https://doi.org/10.3389/fpsyg.2013.01008>
- Manning, F., & Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin & Review*, 20, 1133-1139. <https://doi.org/10.3758/s13423-013-0439-7>
- Matyja, J. R. (2016). Embodied music cognition: Trouble ahead, trouble behind. *Frontiers in Psychology*, 7, Article 1891. <https://doi.org/10.3389/fpsyg.2016.01891>
- Moran, N. (2014). Social implications arise in embodied music cognition research which can counter musicological "individualism". *Frontiers in Psychology*, 5, Article 676. <https://doi.org/10.3389/fpsyg.2014.00676>
- Moran, N., Hadley, L. V., Bader, M., & Keller, P. E. (2015). Perception of 'back-channeling' nonverbal feedback in musical duo improvisation. *PLoS One*, 10(6), Article e0130070. <https://doi.org/10.1371/journal.pone.0130070>
- Nusseck, M., Wanderley, M. M., & Spahn, C. (2017). Body movements in music performances: The example of clarinet players. In B. Müller et al. (Eds.), *Springer handbook of human motion* (pp. 1-14). Berlin, Germany: Springer International.
- Olsen, K. N., & Dean, R. T. (2016). Does perceived exertion influence perceived affect in response to music? Investigating the "FEELA" hypothesis. *Psychomusicology: Music, Mind, and Brain*, 26(3), 257-269. <https://doi.org/10.1037/pmu0000140>
- Pachet, F., Roy, P., & Foulon, R. (2017). Do jazz improvisers really interact? In M. Lesaffre, P. J. Maes, & M. Leman (Eds.), *The Routledge companion to embodied music interaction* (pp. 167-176). New York, NY, USA: Routledge.

- Petrini, K., Dahl, S., Rocchesso, D., Waadeland, C. H., Avanzini, F., Puce, A., & Pollick, F. (2009). Multisensory integration of drumming actions: Musical expertise affects perceived audiovisual asynchrony. *Experimental Brain Research*, 198(2-3), 339-352. <https://doi.org/10.1007/s00221-009-1817-2>
- Platz, F., Kopiez, R., Wolf, A., & Thiesen, F. (2016). *Are visual and auditory cues reliable predictors for determining the finalists of a music competition?* Paper presented at the Jahrestagung der Deutschen Gesellschaft für Musikpsychologie, Vienna, Austria.
- Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167-201). Berlin, Germany: Springer.
- Schiavio, A., & Høffding, S. (2015). Playing together without communicating? A prereflective and enactive account of joint musical performance. *Musicae Scientiae*, 19(4), 366-388. <https://doi.org/10.1177/1029864915593333>
- Schoonderwaldt, E. (2009a). The violinist's sound palette: Spectral centroid, pitch attending and anomalous low frequencies. *Acta Acustica united with Acustica*, 95, 901-914. <https://doi.org/10.3813/AAA.918221>
- Schoonderwaldt, E. (2009b). The player and the bowed string: Coordination of bowing parameters in violin and viola performance. *The Journal of the Acoustical Society of America*, 126(5), 2709-2720. <https://doi.org/10.1121/1.3203209>
- Senn, O., Bullerjahn, C., Kilchenmann, L., & von Georgi, R. (2017). Rhythmic density affects listeners' emotional response to microtiming. *Frontiers in Psychology*, 8, Article 1709. <https://doi.org/10.3389/fpsyg.2017.01709>
- Stupacher, J., Hove, M. J., & Janata, P. (2016). Audio features underlying perceived groove and sensorimotor synchronization in music. *Music Perception*, 33(5), 571-589. <https://doi.org/10.1525/mp.2016.33.5.571>
- Tarr, B., Launay, J., & Dunbar, R. I. M. (2014). Music and social bonding: "Self-other" merging and neurohormonal mechanisms. *Frontiers in Psychology*, 5, Article 1096. <https://doi.org/10.3389/fpsyg.2014.01096>
- Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception*, 28(1), 59-70. <https://doi.org/10.1525/mp.2010.28.1.59>
- Tominaga, K., Lee, A., Altenmüller, E., Miyazaki, F., & Furuya, S. (2016). Kinematic origins of motor inconsistency in expert pianists. *PLoS One*, 11(8), Article e0161324. <https://doi.org/10.1371/journal.pone.0161324>
- Trevarthen, C. (2012). Communicative musicality: The human impulse to create and share music. In D. Hargreaves, D. Miell, & R. MacDonald (Eds.), *Musical imaginations: Multidisciplinary perspectives on creativity, performance, and perception*. Oxford, United Kingdom: Oxford University Press.
- Trusheim, W. H. (1991). Audiation and mental imagery: Implications for artistic performance. *The Quarterly Journal of Music Teaching and Learning*, 2, 138-147.
- Tsay, C.-J. (2013). Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences of the United States of America*, 110(36), 14580-14585. <https://doi.org/10.1073/pnas.1221454110>
- van der Steen, M. C., Jacoby, N., Fairhurst, M. T., & Keller, P. E. (2015). Sensorimotor synchronization with tempo-changing auditory sequences: Modeling temporal adaptation and anticipation. *Brain Research*, 1626, 66-87. <https://doi.org/10.1016/j.brainres.2015.01.053>

- Vos, J., & Rasch, R. (1981). The perceptual onset of musical tones. *Perception & Psychophysics*, 29(4), 323-335. <https://doi.org/10.3758/BF03207341>
- Vuoskoski, J. K., Thompson, M. R., Clarke, E. F., & Spence, C. (2014). Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception & Psychophysics*, 76, 591-604. <https://doi.org/10.3758/s13414-013-0582-2>
- Vuoskoski, J. K., Thompson, M. R., Spence, C., & Clarke, E. F. (2016). Interaction of sight and sound in the perception and experience of musical performance. *Music Perception*, 33(4), 457-471. <https://doi.org/10.1525/mp.2016.33.4.457>
- Weger, U. W., Meier, B. P., Robinson, M. D., & Inhoff, A. W. (2007). Things are sounding up: Affective influences on auditory tone perception. *Psychonomic Bulletin & Review*, 14(3), 517-521. <https://doi.org/10.3758/BF03194100>
- Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLOS One*, 9(4), Article e94446. <https://doi.org/10.1371/journal.pone.0094446>
- Wöllner, C., & Cañal-Bruland, R. (2010). Keeping an eye on the violinist: Motor experts show superior timing consistency in a visual perception task. *Psychological Research*, 74, 579-585. <https://doi.org/10.1007/s00426-010-0280-9>
- Wöllner, C., Parkinson, J., Deconinck, F. J. A., Hove, M. J., & Keller, P. (2012). The perception of prototypical motion: Synchronization is enhanced with quantitatively morphed gestures of musical conductors. *Journal of Experimental Psychology: Human Perception and Performance*, 38(6), 1390-1403. <https://doi.org/10.1037/a0028130>