Action-based effects on music perception

Pieter-Jan Maes[1]*, Marc Leman[2], Caroline Palmer[3] and Marcelo M. Wanderley[1]

[1] Department of Music Research, McGill University, Montreal, QC, Canada
[2] Department of Musicology, Ghent University, Ghent, Belgium
[3] Department of Psychology, McGill University, Montreal, QC, Canada

The classical, disembodied approach to music cognition conceptualizes action and perception as separate, peripheral processes. In contrast, embodied accounts of music cognition emphasize the central role of the close coupling of action and perception. It is a commonly established fact that perception spurs action tendencies. We present a theoretical framework that captures the ways in which the human motor system and its actions can reciprocally influence the perception of music. The cornerstone of this framework is the common coding theory, postulating a representational overlap in the brain between the planning, the execution, and the perception of movement. The integration of action and perception in so-called internal models is explained as a result of associative learning processes. Characteristic of internal models is that they allow intended or perceived sensory states to be transferred into corresponding motor commands (inverse modeling), and vice versa, to predict the sensory outcomes of planned actions (forward modeling). Embodied accounts typically refer to inverse modeling to explain action effects on music perception (Leman, 2007). We extend this account by pinpointing forward modeling as an alternative mechanism by which action can modulate perception. We provide an extensive overview of recent empirical evidence in support of this idea. Additionally, we demonstrate that motor dysfunctions can cause perceptual disabilities, supporting the main idea of the paper that the human motor system plays a functional role in auditory perception. The finding that music perception is shaped by the human motor system and its actions suggests that the musical mind is highly embodied. However, we advocate for a more radical approach to embodied (music) cognition in the sense that it needs to be considered as a dynamical process, in which aspects of action, perception, introspection, and social interaction are of crucial importance.

Keywords: embodied music cognition, common coding theory, sensory-motor association learning, dynamical systems, internal model

1. INTRODUCTION

Music is known to be a powerful medium that evokes body movements in listeners, ranging from tapping the feet, shaking the head, swaying the arms and hips, to more sophisticated forms of free or stylized dance. Research has shown that these body movements often reflect the performer's movements from which the music originated (Leman et al., 2009; Godøy and Leman, 2010), certain aspects of the melody, harmony, rhythm and timbre (Maes et al., 2010; Naveda and Leman, 2010; Toiviainen et al., 2010; Burger et al., 2013; Leman et al., 2013), or even the listeners' mood (Van Dyck et al., 2013). These and similar studies importantly indicate that the listeners' musical mind (attention, intention, mood, feelings, etc.) can be accessed through body movement, without the need for symbolic representations like language or musical scores. However, despite the explicit focus on the human body and body movements, these and similar studies do not consider the musical mind as being fundamentally embodied. The findings do not exclude the possibility that movement responses to music are mere peripheral epiphenomena resulting from central cognitive processes. Only recently have studies started to

emerge, demonstrating how the musical mind can be shaped by the human motor system and the movements it produces (Phillips-Silver and Trainor, 2005, 2007; Repp and Knoblich, 2009; Sedlmeier et al., 2011; Iordanescu et al., 2013; Loehr, 2013; Manning and Schutz, 2013; Maes and Leman, 2013; Timm et al., 2013). This line of research reflects an important paradigm shift within cognitive science. The classical view, inspired by the developments of computer science and artificial intelligence in the 1950s–1960s, pertains to an "information processing" approach that considers a strictly unidirectional information flow from perception (input) to cognition (central processing unit) to action (output) (Neisser, 1967; Laske, 1974; Fodor, 1975; Pylyshyn and Demopoulos, 1986; Massaro, 1990). Accordingly, sensory information received from the external world is perceived, translated into a syntactic code of meaningful symbols, and processed according to a systematic set of rules. Then, body movements and other sorts of behavior are considered as mere outcomes of these higher-level, formal symbol manipulations. Hence, in this classical view of cognition, perception and action are completely separated from each other, and are outside central cognition

[what Hurley (2001) describes as the "sandwich model of cognition"]. This classical model is obsolete, as research shows that perception and action are strongly intertwined and can mutually exert influence on each other. In what became the *embodied* cognition theory, the human body - with its perceptual and motor systems - and its interaction with the outside world, became central to human cognition (Varela et al., 1991; Leman, 2007; Chemero, 2009; Krueger, 2009; Glenberg, 2010; Shapiro, 2010). Within this framework of embodied cognition, the common coding theory (Prinz, 1990, 1997; Hommel et al., 2001) has been an influential theory postulating a close coupling between perception and action. Although the theory is not readily falsifiable, it provides a general framework for developing more detailed and testable explanatory models (cf. Hommel et al., 2001). In essence, the theory states that the planning or execution of an action, and the mere perception of the (multi-)sensory consequences of that action, are similarly represented (coded) in the brain, thereby recruiting both sensory and motor brain areas. Important in this theory is that the integration of motor and sensory representations leads to internal models of the relationship between the body and the external environment, which can contain inverse and forward components (Wolpert et al., 1995). Inverse models represent an information flow from perception to action, in the sense that they allow the system to estimate from incoming sensory information the corresponding motor commands required to generate that specific sensory state [cf. Rizzolatti et al. (2001): direct-matching hypothesis]. In contrast, forward models represent an information flow from action to perception, in the sense that they allow to predict the likely sensory outcome of a planned or executed action (Davidson and Wolpert, 2005; Bubic et al., 2010; Waszak et al., 2012). Currently, the idea is gaining consensus that the combination of inverse and forward modeling processes guides people's interaction with the external world, including motor control and sensory processing.

In the present paper, we set the common coding theory, and the related theory of internal models, as a theoretical framework for understanding action-based effects on music perception. We conjecture that a focus on both inverse and forward modeling processes can provide a comprehensive view of how the human motor system and its actions influence music perception. In the domain of embodied music cognition, one typically refers to inverse modeling processes to explain action-based effects on music perception. Music spurs body movements that amount to expressive qualities, intentions, inner feelings, etc. Many of the musical elements that contribute to expressivity (e.g., dynamics, articulation, touch, phrasing, vibrato, rubato, etc.) directly relate to physical aspects of movement and space. Inverse modeling processes enable us to render (or decode) perceived patterns of musical expressivity into corresponding body movements. This corporeal mirroring process is responsible for listeners' tendency to ascribe intentions, inner feelings, etc. to music (Godøy, 2003; Leman, 2007; Cox, 2011). We want to extend this "traditional" embodied perspective to the role of the human body in music cognition with a focus on forward modeling processes. From this perspective, it is not about how the body resonates with the music, but rather about how predicted sensory outcomes of planned or performed actions can be projected onto the perceived

music. Recently, there has been a proliferation of studies addressing the role of forward models in action-based effects on visual, auditory, and somatosensory perception. In the domain of visual perception, several papers review action-based effects on visual perception (Schütz-Bosbach and Prinz, 2007; Shin et al., 2010; Witt, 2011; Halász and Cunnington, 2012). Currently, such a review of studies investigating action-based effects on auditory perception does not exist. An important goal of the present paper is to provide such a review of studies in support of the proposed theories and principles.

The paper is structured as follows. In section 2, we argue that sensory-motor association learning can be considered a central mechanism underlying the development of internal models. Accordingly, we claim that the ability to predict the auditory consequences of one's actions, which is one of the core mechanisms of action-based effects on perception, depends on previous acquired sensory-motor associations. Further in that section, we define the concepts of temporal contiguity and probabilistic contingency as two main principles underlying associative learning processes. Additionally, we discuss musical instrument playing as a special but highly illustrative case of sensory-motor association learning. In section 3, we provide extensive empirical evidence for the claim that the principle of motor resonance, inherent in inverse models (section 3.1), together with auditory predictions generated by forward models (section 3.2), can modulate auditory perception. Also, we demonstrate that deficits in the motor system may have impaired auditory perception as a consequence (section 3.3). To conclude, an extensive discussion is presented in which we advocate a radical approach to embodied music cognition based on dynamical systems. Moreover, we pinpoint music as an ideal study object to extend this approach based on dynamical systems to embodied cognition, as it incorporates expressivity, introspection (affect, motivation, intentions, metacognition, etc.), and social interaction as crucial components.

2. ASSOCIATIVE LEARNING

Above, we outlined the common coding of action and perception as a core mechanism underlying people's engagement with music (motor control and sensory processing). However, this account does not address the question of how action and perception become integrated. We advocate that this integration is established, in large part, through associative learning processes. The study of these processes can be traced back to the philosophy of Aristotle who stated that things that occur near each other in time and/or space are readily associated (i.e., law of contiguity). During the Enlightenment, these ideas were further developed by the Associationist School (e.g., David Hume, John Locke, John Stuart Mill, etc.). In the nineteenth century William James stated, as an elementary law of association, that "when two elementary brain-processes have been active together or in immediate succession, one of them, on reoccurring, tends to propagate its excitement into the other" (James, 1890, p.566). In the late 1940s, this principle was paraphrased in Hebb's law "neurons that fire together wire together." A more recent account is the theory of associative sequence learning (ASL) introduced by (Heyes and Ray, 2000). The ASL theory suggests that imitation is mediated by associative processes that establish links between sensory

and motor representations. This theory has been applied to the human mirror neuron system (MNS) in an attempt to reconsider its origin and function. The classical view on the MNS—as originated in the work of Gallese et al. (1996); Rizzolatti et al. (2001); Kohler et al. (2002)—is that it is an innate system, only marginally influenced by sensory-motor experience, and inherently codes the meaning of actions (e.g., goals, intentions, etc.). This view was soon adopted to explain various important psychological and social functions, such as action understanding, learning by imitation, empathy, and social interaction. However, critical voices have been raised in opposition to this classical view, in particular to the idea that mirror neurons are adapted by evolution to directly and consistently encode action goals (Hickok, 2009; Heyes, 2010; Catmur, 2012). The alternative view—what Heyes (2010) termed the *associative hypothesis*—states that the development of the MNS is promoted by sensory-motor associative learning. Empirical evidence is provided in the context of music and dance. Haslinger et al. (2005) compared expert pianists with musically naive controls with fMRI while observing piano-playing and non-piano-playing finger movements. The results showed that the expert pianists exhibited stronger activation in brain areas associated with the MNS (inferior fronto-parieto-temporal region) compared to the control participants. Similarly, in the context of dance, Calvo-Merino et al. (2005) showed that activation in brain areas related with the MNS in expert dancers (classical ballet and capoeira) was higher when they observed a familiar dance style. In conclusion, the associative hypothesis states that, through systematically repeated experiences, sensory events are associated with particular motor acts and excitatory links between both are created, resulting in the development of "internal models." Accordingly, when a sensory representation is activated, the corresponding motor representation is automatically co-activated (inverse modeling), and vice versa: when an action is merely planned or executed, the corresponding sensory representation is automatically co-activated (forward modeling). As will be explained further in Section 3, both inverse and forward modeling processes can contribute to action-based effects on auditory perception.

An important challenge of future research is to further identify the neural substrates underlying associative learning processes. Studies pinpoint the cerebellum (Imamizu and Kawato, 2009; Timmann et al., 2010), the striatum—an input nucleus of the basal ganglia—(Pasupathy and Miller, 2005; Williams and Eskandar, 2006; Lalazar and Vaadia, 2008; Melcher et al., 2012), prefrontal areas (Deiber et al., 1997; Bangert and Altenmüller, 2003; Pasupathy and Miller, 2005), the supplementary motor area (Pasupathy and Miller, 2005), and the premotor cortex (Deiber et al., 1997; Schubotz, 2007; Chen et al., 2009; Imamizu and Kawato, 2009) as important neural structures underlying association learning leading to the development of internal models and predictive mechanisms. In the field of music research, evidence suggests that the striatum is involved in prediction and anticipation. Grahn and Rowe (2013) assessed the role of the putamen—one of the two nuclei that make up the striatum—in beat prediction. Their findings show that the putamen becomes active only after having established a predictable sense of the beat. Accordingly, they conclude that putamen activity reflects the

process of internally generating a model of the stimulus rhythm. In a study of Leaver et al. (2009), anticipatory/predictive imagery of musical melodies was shown to be associated with activation in a variety of cortical (frontal and parietal) and subcortical (basal ganglia and cerebellum) structures. Interestingly, different neural substrates underlay different stages of development of learned conditional associations between melodies ("moderately learned" vs. "well-learned"). Findings show that the supplementary motor area and the basal ganglia (putamen) are particularly important in early and moderate stages of learning, while the frontal cortex seems to dominate end stages (cf. Pasupathy and Miller, 2005). These dynamics in neural activation involved in sensorimotor association learning characterizes motor skill learning in general. Studies have demonstrated that the recruitment of distributed brain regions in the process of acquiring motor skills depends on the type of motor task (motor sequence learning vs. motor adaptation) and on the stage of learning (fast learning, slow learning, consolidation, automatization, retention) (Ungerleider et al., 2002; Luft and Buitrago, 2005; Doyon et al., 2009).

2.1. CONTINUITY AND CONTINGENCY

Auditory-motor association learning—i.e., the acquisition of knowledge of sound-movement relationships—is modulated by both temporal "contiguity" and probabilistic "contingency" (Cooper et al., 2012). "Contiguity" refers to the proximity of two events (e.g., movement and sound) in time and space. The concept originates in Aristotle's law of contiguity, stating that things that occur near each other in time and/or space are readily associated. It is not, however, the case that association learning occurs every time two events are linked together in time or space. Instead, it is necessary that the relationship between the events is predictable. "Contingency" refers to this degree of probability or the likelihood that two or more events belong together. In statistical terms, contingency is related to covariance, being a measure of how much two random variables change together.

Elsner and Hommel (2004) present two experiments in which the role of contiguity and contingency were investigated in the development of sensory-motor associations. Each experiment consisted of a training phase followed by a test phase. In the training phase, participants learned action-effect associations by repeatedly pressing keys (action) triggering corresponding tones (effect). In the subsequent test phase, tones were presented and participants were asked to make speeded responses to these stimuli by pressing keys either in a consistent fashion (i.e., action-effect mapping as in the training phase) or inconsistent fashion (i.e., other action-effect mapping as in the training phase). If an action-effect association was established in the training phase, then participants were expected to respond faster in an acquisition-consistent fashion than in an acquisition-inconsistent fashion. In the training phase of Experiment 1, the contiguity between action and effect was systematically manipulated by adding an increasing delay between the two (50, 1000, and 2000 ms). In the test phase, participants responded faster in acquisition-consistent test blocks compared to acquisition-inconsistent test blocks when action-effects training delays were 50 or 1000 ms. Accordingly, association learning seemed to be successful only with action-effect delays of up to

1000 ms, signaling an effect of contiguity in association learning. In the training phase of Experiment 2, the contingency between action and effect was systematically manipulated by varying the relative frequencies of the presence or absence of tones with corresponding keypresses. Again, it was shown that the acquisition-consistency effect in the test phase was affected by the contingency of action and effect in the training phase. Together, these findings show that both the contiguity and contingency between actions (here, keypresses) and auditory events (here, sinusoidal tones, MIDI marimba/flute tones) are important in the process of acquiring sensory-motor associations.

An interesting experimental paradigm in which contiguity and contingency could be further investigated is the counter-mirror sensory-motor training paradigm (Cook et al., 2010). In this paradigm, previously established associations between motor and sensory events are manipulated by repeatedly pairing the observation of an action with the execution of another action. One typically finds (e.g., by measuring neural responses, or reaction times) that the original sensory-motor association gets weakened, depending on the principles of contingency and contiguity. This paradigm has been applied to visual-motor learning processes, but not yet to auditory-motor learning processes. However, the paradigm offers unique possibilities to study for instance how counter-mirror training can alter auditory-motor links established in musical instrument playing.

2.2. MUSICAL INSTRUMENT LEARNING

Learning to play an instrument can be considered a special, highly illustrative case of sensory-motor association learning in which action and perception become intricately interwoven. The act of playing an instrument can be considered as a goal-directed, intentional act (Dalla Bella and Palmer, 2011). Ultimately, the goal of playing a musical instrument is to produce a certain sound. However, in order to reach that goal, one first needs to obtain knowledge about the relationship between the actions afforded by the musical instrument, and the auditory consequences of these actions. This knowledge is gradually acquired by exploring and manipulating the possibilities afforded by the instrument using (at first) arbitrary actions that lead to (at first) unexpected auditory events (Hommel, 2003). In that process of exploration and interaction, one systematically and repeatedly associates performed actions with heard sounds, and internal models are developed as a result, capturing the relationship between actions and sound. For example, in the case of the piano, one starts to understand that the key-to-pitch mapping is functionally organized (left-right motion corresponds to low-high pitch), or that depressing the sustain pedal creates a legato effect. At that point, playing a musical instrument may become a goal-directed act, in the sense that performers have the ability to intentionally produce certain sounds by performing certain actions. Additionally, it must be noted that the process of exploration in which action and perception mutually interact, is a continuous process throughout the life of a music performer. It incorporates aspects of creativity, intuition and surprise, and can in itself be a "raison d'être" of playing an instrument (cf. Sudnow, 1978).

A large body of empirical studies exist that support these ideas. For example, it has been shown that when people are trained to play a musical instrument, auditory-motor linkages are developed as a result of that training (Pascal-Leone, 2001; Bangert and Altenmüller, 2003; Lotze et al., 2003; Lahav et al., 2005; D'Ausilio et al., 2006; Lahav et al., 2007; Hyde et al., 2009; Herholz and Zatorre, 2012). Also studies have shown that during passive music listening, trained musicians exhibit stronger auditory-motor couplings compared to non-musicians (Haueisen and Knösche, 2001; Gaser and Schlaug, 2003; Baumann et al., 2007). This supports the idea that auditory-motor linkages are established by intensive training which involves long-term skill acquisition and the repetitive rehearsal of the same skills (Brown and Palmer, 2012, 2013).

It is evident that sensory-motor association processes are important for voluntary action control, as in musical instrument performance (Hommel, 1997, 2003; Elsner and Hommel, 2001). However, more important in the light of the present paper is the idea that sensory-motor relationships, and the integration of these relationships into internal models, may influence perceptual processes and accordingly shape the musical mind. In the following sections, we will discuss empirical evidence demonstrating that sensory-motor association learning, with musical instrument training as a special case, may lead to action-based effects on auditory perception.

3. EMPIRICAL EVIDENCE: A REVIEW
3.1. INVERSE MODEL: PERCEPTION → ACTION

Inverse models enable us to predict the motor commands that are required to achieve a desired sensory state. It is obvious that this is of utmost importance when playing a musical instrument. But inverse models hold an important role in music perception as well, as they allow to predict and simulate the physical aspects of motion and space implied in the music. There is ample evidence that merely listening to sounds or music automatically triggers motor responses, as a function of their previously established associations [motor resonance (Schütz-Bosbach and Prinz, 2007), perceiving action (Hurley, 2008), etc.]. This has been shown in neurophysiological studies (Haueisen and Knösche, 2001; Bangert and Altenmüller, 2003; Gaser and Schlaug, 2003; Lahav et al., 2005, 2007; D'Ausilio et al., 2006; Baumann et al., 2007; Chen et al., 2008). Additionally, results from behavioral studies show that motor responses to sounds are typically faster when the specific sounds and actions have been repeatedly and consistently paired on previous occasions (Elsner and Hommel, 2001; Rusconi et al., 2006; Lidji et al., 2007; Trimarchi and Luzzatti, 2011; Stewart et al., 2013a,b). These findings provide support for the idea that an action becomes automatically activated (or, primed) as a result of the mere perception of the auditory consequences normally associated with that action.[1] Other studies have focused on overt body movements that people make in response to music for music presented in visual form, or via motion imagery (Eitan and Granot, 2006; Leman et al., 2009; Caramiaux et al., 2010; Godøy, 2010; Kozak et al., 2012; Bernardi et al., 2013; Küssner, 2013; Lotze, 2013). These studies show that people can consistently translate acoustic properties of sound and music into body movements, although Küssner (2013) reports

---

[1] see Cox and Hasselman (2013) for some critical remarks on typical effect-priming studies.

that musicians are more consistent (i.e., less varying) in visualizing sound and music by means of drawings. More important in the scope of the present article is the idea that the power of music to induce body movements in listeners implies that merely listening to music becomes a kinaesthetic experience. Musical groove is a relevant example of a musical quality that induces body movements in listeners (Janata et al., 2012; Stupacher et al., 2013). The notion of music-induced body movement may be related to two ideas showing how inverse models, and the related concept of motor resonance (or, motor simulation), can shape people's engagement with music and by extension the "musical mind."

First, the recruitment of the body into the process of music listening causes a connection to be made between the music and the expressive qualities inherent to the movements that the music induces. The human body acts thereby as a mediator between physical phenomena (sensory and motor processes) and subjective, mental states (Leman, 2007). An interesting model to capture the subtle qualities of movement expressivity is the Effort/Shape model that originated in the Laban Movement Analysis (LMA) method (Laban, 1947; Laban and Ullmann, 1966). This model is particularly appropriate, as it provides an integrated conceptual system connecting a set of physical movement properties with expressive qualities (e.g., weight, flow, space, time, etc.). The model has been used in research to show how music-induced body movements correlate with verbal descriptors used by people to describe their perception of the music (Maes et al., 2014).

Second, it is interesting to note that music-induced body movements may instigate a sense of *imagined participation* with the production of the sound. This idea of imagined participation is addressed in a broad range of musicological studies with different terminology, such as *imagined activity* (Maus, 1988), *kinaesthetic empathy* (Mead, 1999), *imaginary agency* (Levinson, 2006), *simulated control* (Leman, 2007), and *active perception* (Krueger, 2009). What these accounts have in common is their reference to a direct, sensory-motor engagement with music, to how music literally "moves" people, and to how people feel immersed in, and resonate with, the physical sound energy. In that sense, motor resonance may create the illusion of taking part in the actual skillful production of the music, which would be impossible in real life. Musical motion, however, is not limited to purely physical movements of the human body. Schubotz (2007) provides an answer to the question of how people can simulate or anticipate events that could not be readily reproduced by their own motor system (e.g., rhythm of ocean waves, the flight of a mosquito, or an unfolding sequence of abstract stimuli on a computer screen). Schubotz demonstrates and explains that even abstract events—including auditory events—recruit our motor system (in particular the premotor cortex and its parietal projection areas) in order to support simulation and prediction processes (see also Southgate, 2013). Accordingly, the micro and macro dynamics and subtleties inherent in the musical textures and structures, as for instance in the "Clocks and Clouds" (1973) of György Ligeti or in electronic music productions (e.g., Infected Mushroom, Aphex Twin, etc.), can evoke a fascinating continuum of spatial imagery and motion, with which the listener may float along. Accordingly, motor resonance may generate an experience of flow, being a state of heightened focus and immersion, typically accompanied with intense feelings of enjoyment and creativity (Csikszentmihalyi, 1988). This aspect of motor resonance is an essential component of musical aesthetic experiences and is fundamental for shaping the "musical mind." Additionally, it may be a factor that explains the ability of music to alter people's experience of space and time (Schäfer et al., 2013), and to contribute to people's general well-being (Croom, 2011).

3.2. FORWARD MODEL: ACTION → PERCEPTION

As explained, forward internal models represent an information flow from action to perception, in the sense that they allow the prediction of the likely sensory outcome of a planned or executed action [cf. "perceptual resonance" (Schütz-Bosbach and Prinz, 2007), "active perception" (Hurley, 2008), etc.]. Research has pinpointed the cerebellum as a crucial locus for internal forward models (Wolpert et al., 1998; Blakemore et al., 2001; Knolle et al., 2012a; Ebner, 2013), presumably in interaction with other brain structures [e.g., prefrontal areas (Lappe et al., 2013)]. In this context, it is important to note that different predictive mechanisms exist which are supported by different brain systems. O'Reilly et al. (2013) for example differentiate between statistical and dynamic predictive models. Statistical models capture the stochastic probability that two or more events are associated— for example an action event and a reward or sensory event—and are developed over a history of discrete events. Alternatively, in dynamic forward models, the relation between two events is deterministic and predictions are computed via explicit reference to pre-learned environmental dynamics.

Studies have shown that predictive models are important for motor control (Wolpert et al., 1995; Hommel, 1997), as well as for the processing of sensory information coming from the external environment (Halász and Cunnington, 2012). In the present study, we focus on the latter in the context of auditory perception. We will discuss how sensory predictions generated by forward models may influence the perception of sound and music. It will be shown that sensory predictions can either attenuate, facilitate, or disambiguate auditory perception (cf. Halász and Cunnington, 2012).

3.2.1. Attenuation

Performing an action for which one can predict the sensory consequences attenuates the perception of the actual sensory outcome, as reflected in self-reports and neuronal responses. In the domain of auditory perception, this phenomenon was first studied in speech production (Houde et al., 2002; Heinks-Maldonado et al., 2006). Later on, studies appeared in which the phenomenon of motor-induced suppression (MIS) was studied with tones generated by keypresses. Despite the fact that the tones and the actions that produce them (i.e., action-to-pitch mapping) are highly simple, a parallel can be drawn with musical instrument playing, like playing the piano, trumpet, etc.

A study conducted by Aliu et al. (2009) demonstrates that the auditory response to tones generated by self-produced keypresses is attenuated relative to the response following passive listening to the same tones. However, because self- and externally-generated tones were presented in separate blocked conditions, it could not be ruled out that the observed attenuation effect

was modulated by differences in contextual task demands (e.g., allocation of attention, arousal, etc.). To clarify this matter, Baess et al. (2011) mixed self- and externally-generated tones within blocks. The results of this study yielded an even larger attenuation effect for self-generated tones than that observed in the blocked condition. Also, Timm et al. (2013) conducted a study to further investigate the relationship between attention and the effects of motor prediction in perceiving auditory stimuli. The study adapted the mixed paradigm of Baess et al. (2011) and additionally incorporated different conditions in which attention was allocated to either the sound, the motor act, or to visual stimuli. Findings of this study demonstrated that an attenuation effect for self-generated sounds was independent from the allocation of attention. Other studies investigated whether the attenuation of auditory action effects occurs when actions are merely observed, instead of being self-generated. Sato (2008) hypothesized that, if there is a human mirror neuron system that codes a bidirectional association between action execution and action perception, then the mere observation of (well-learned) actions leading to a certain auditory event should bring about a similar auditory attenuation effect as when the action is self-generated. The results of this study confirmed this hypothesis, as similar auditory attenuation was observed for self-generated and merely observed sound-producing actions. However, this finding was later contradicted in a study by Weiss and Schütz-Bosbach (2012), using a comparable experimental protocol as in Sato (2008). They compared auditory action effects for self-generated actions, observed unanticipated actions, and observed anticipated actions. The results showed that the attenuation of a sound is significantly higher when the sound-producing action is self-generated compared to merely observed. Moreover, this effect was shown to be independent of whether the observed action could be anticipated or not. This finding raises questions about the role that forward internal models play in the prediction mechanisms underlying action effects on auditory attenuation (cf. Sato, 2008). More research is needed to clarify this point. In a last study we address here, Knolle et al. (2012b) examined whether auditory attenuation is a function of the degree of predictability of the self-generated sound. The result of this study indicated a lowering of the attenuation effect when self-generated sounds deviate from the expected outcome.

Together, these and similar studies (Baess et al., 2008; Hughes et al., 2013a,b; Jones et al., 2013; Loehr, 2013; Sanmiguel et al., 2013) provide strong evidence in support of the existence of an internal, motor-based prediction mechanism that can modulate auditory perception. Planning or executing an action causes a copy of the motor command to be made (i.e., "efferent copy," or "corollary discharge"), which enables a prediction of the auditory outcome of that motor command. A comparison between the prediction and the actual auditory input ("reafference input") leads to a small prediction error, and subsequently to a minimal response in the auditory cortex reflecting an attenuated perception (Aliu et al., 2009). This mechanism enables to discriminate between auditory inputs that are a consequence of our own actions and those that reach us from the external world. It is important to consider that this mechanism requires (learned) internal models about the relationship between sensory and motor representations. Only recently, studies have started

to unveil the neural substrates of motor-based sensory prediction (Nelson et al., 2013; Roussel et al., 2013). However, more research is needed in order to obtain a full picture of the neural mechanisms underlying the action effect of auditory attenuation.

3.2.2. Facilitation

Manning and Schutz (2013) examined to what extent "moving to the beat" objectively improves timing perception. They presented participants with sequences of 16 isochronous tones divided into groups of four followed by a probe tone. In the last group, the second, third, and fourth "tones" were silent (i.e., timekeeping segment). The probe tone was "on-time" (i.e., sounding after the same inter-onset interval), slightly early, or slightly late. The task of the participant was to judge whether the final probe tone sounded "on-time." In one condition, participants were asked to tap along with the beat, while they remained still in the other condition. The results show that late offsets were better detected when participants could move during the timekeeping segment. Additionally, it was found that "better" tappers (i.e., less variability) performed better on the detection task overall. In general, these findings confirm that movement may improve time perception. Iordanescu et al. (2013) obtained similar results using a standard temporal-bisection paradigm. Participants were presented with sequences of three brief clicks with the location of the second click randomly varied. Participants had to judge whether the second click was temporally closer to the first or the third click. In the "active" condition, participants initiated each trial themselves by pressing the space bar, while trials were externally generated in the "passive" condition. Again, in line with the results of Manning and Schutz (2013), people in the active condition demonstrated a higher auditory sensitivity to temporal intervals. Moreover, it was shown that this effect was not attributable to the tactile sensation from a keypress. It is interesting to note that the finding that body movement can enhance time perception has been picked up by research in the domain of human-computer interaction (HCI) design. Maes et al. (2012, 2013) present a dance application and a music conducting application aiming to enhance users' understanding of temporal musical structures by teaching them how to articulate these temporal structures into corresponding body movements (dancing, conducting).

In another study, Brown and Palmer (2012) investigated how motor and auditory learning contribute to auditory memory for music. Pianists were asked to learn melodies on a Musical Instrument Digital Interface (MIDI) piano keyboard in each of four conditions (auditory only, motor only, strongly coupled auditory-motor [i.e., normal performance], or weakly coupled auditory-motor [i.e., performing along with auditory recordings (acoustically similar or varying) without hearing their own feedback]). After learning, participants heard melodies (half target, half foils) in a subsequent recognition test and were instructed to indicate which melodies they had encountered in the learning conditions. It was found that motor learning (combined with strongly coupled auditory learning) enhanced auditory recognition beyond auditory learning alone. Results were explained by the ability of sensory-motor associations formed during learning to provide additional retrieval cues and to shape auditory perception through mental simulation of action plans.

3.2.3. Disambiguation

Music may have a certain degree of ambiguity in terms of perceptual and/or affective content. As discussed below, studies indicate that it is possible for a listener to disambiguate this content by planning or executing body movements during listening. Forward models provide an appropriate explanation for this disambiguation effect (Halász and Cunnington, 2012). The planning or execution of body movements enables one to automatically predict the sensory consequences of these actions. Consequently, these predicted sensory states can be projected onto the auditory or musical material, which may guide (i.e., disambiguate) the corresponding perception. Some additional remarks need to be made, however. First, planning body movements does not only generate predictions of sensory states, but equally of subjective mental states related to affect and expressivity (e.g., valence, arousal, etc.). In that sense, it is equally possible that subjective states are attributed to the music (Thompson et al., 2005; Juchniewicz, 2008; Sedlmeier et al., 2011; Maes and Leman, 2013). Second, auditory or musical material doesn't necessarily need to be ambiguous in order for body movements to guide our perception in a specific direction. Music presents the performer and listener with a flood of different auditory cues and accents. Body movements can help selectively direct attention to certain cues, and accordingly to impose a certain structure onto the music. According to Urista (2003); Pierce (2007), body movements can help to isolate and explore musical elements as melody, beat, and structural levels. Hence, cue selection (and, cue identification) facilitated by body movement may refine music listening in general, and shape our perception and understanding of the music. Third, studies show that merely observing body movements, instead of actually planning or executing them, may equally influence perceptual and aesthetic judgments of the produced music (Thompson et al., 2005; Schutz and Lipscomb, 2007; Juchniewicz, 2008). Fourth, it is possible that executed or observed body movements modulate auditory perception instantaneously, i.e., at the moment one listens to the music (Thompson et al., 2005; Schutz and Lipscomb, 2007; Juchniewicz, 2008; Repp and Knoblich, 2009; Sedlmeier et al., 2011). Additionally, it is also possible that when one repeatedly pairs body movements to music, the resulting action-based effects on music perception may endure for a longer period of time, in the sense that the specific way of perceiving music may retain when merely listening to the music without the need to intentionally plan or execute the corresponding body movements (Phillips-Silver and Trainor, 2005, 2007; Maes and Leman, 2013). So by sensory-motor associative learning processes, music may become integrated with actions and more importantly, with the sensory and affective states inherent to these actions. It is a form of "evaluative conditioning" leading to effects of disambiguation and cue selection (Juslin and Västfjäll, 2008; Maes and Leman, 2013). Moreover, depending on the nature of the learning process (e.g., duration, continuity, contingency, etc.), these effects can be retained for different amounts of time.

In the following section, we discuss several studies that illustrate these effects of disambiguation and cue selection. Phillips-Silver and Trainor (2005, 2007) addressed the interaction between body movement and the perception of musical rhythm. The procedures of the experiments conducted in these studies contained a training phase and a subsequent test phase. In the training phase, infants were passively bounced (Phillips-Silver and Trainor, 2005), or adults bounced actively by bending their knees (Phillips-Silver and Trainor, 2007) on every second (duple) vs. every third (triple) beat of an ambiguous musical rhythm pattern. In the subsequent test phase, infants' listening preferences were tested for two auditory versions of the rhythm pattern (duple and triple form) (Phillips-Silver and Trainor, 2005). In Phillips-Silver and Trainor (2007), the adults were asked to listen to two auditory rhythms (duple and triple rhythm) and to select the one they thought matched what they heard during the training phase. The results showed that the preferences and interpretations were oriented toward the auditory stimulus that matched the metrical form of their movement training.

In a study by Naveda and Leman (2009) it was shown that Samba music has a polymetric ambiguity, whereas Samba dance patterns typically have binary tendencies. Accordingly, the authors suggest that "perception of samba may be movement-based in the sense that through self-movement (of the dancer in response to music) musical patterns get rhythmically disambiguated."

In a study by Sedlmeier et al. (2011), it was shown that real or imagined body movements during music listening may codetermine music preferences. The experimenters activated or inhibited specific muscles of the participants whose innervations have been shown to be associated with positive and negative emotions. This was realized by instructing the participants to perform three specific kinds of body movements or actions (activating/inhibiting "smiling muscles," vertical/horizontal head movements, and flexion/extension of the arms). Activation of the positively associated muscle groups during listening to music led to higher preference ratings for that music than activation of the negatively associated ones. This suggests that body movements, both real and imagined, may play an important role in the development of music preferences.

Su and Pöppel (2012) tested the hypothesis that the use of body movement is not merely a reaction to hearing rhythmic input, but could actively assist the processing of temporal structures in auditory events. They suggest that a self-initiated movement frequency, which is not tuned-in at first, could be attracted to one of the underlying periodicities of the presented sequence. Doing so guides the listener to start "hearing" the pulse at that level, forming a positive audio-motor feedback loop. The authors show that in the absence of overt movement, by contrast, this tuning process must then rely on the internal motor entrainment and/or the ability to analyze the sequence. Unlike musicians, non-musicians seemed to lack an effective internal motor simulation that entrained to the pulse when it was not regularly present at the rhythmic surface, nor did they possess additional musical knowledge as a compensatory strategy.

A study by Iversen et al. (2009) investigated how the perception of a simple rhythmically ambiguous phrase (i.e., a repeating series of two tones followed by a rest) depends upon its intrinsic metrical interpretation. Participants were asked to mentally place the downbeat on either the first or the second tone of the rhythmical phrase. Using magnetoencephalography (MEG) it was shown that different metrical interpretations evoked different

neural responses, specifically in the upper beta range (20–30 Hz). This led the authors—given the suggested role of beta in motor processing—to the hypothesis that the motor system influences metrical interpretation of sound, even in the absence of overt movement. In another study, Maes and Leman (2013) addressed the question of whether expressive body movements can condition children's perception of musical expressiveness. They trained children with a happy or a sad choreography in response to music that had an expressively ambiguous character. Afterwards, the children's perception of musical expressiveness in terms of valence and arousal was assessed. The results suggested that the expressive qualities of the movements they learned to associate with the music had a significant impact on how children perceived musical expressiveness.

In a study by Repp and Knoblich (2009), participants were asked to play pairs of octave-ambiguous (Shepard) tones which were a tritone apart. Although each tone of a pair is characterized by a specific pitch class (e.g., C - F#), they are ambiguous in pitch height. Participants were asked to play the pairs of tones by pressing corresponding piano keys or computer keyboard keys, either in left-to-right or right-to-left direction. Consecutively, they had to judge whether each pitch interval was rising or falling. Results showed that the participants gave significantly more "rising" responses when the order of keypresses was left-to-right than when it was right-to-left. Moreover, this effect was larger for pianists compared to non-pianist musicians, most likely because the specific pitch-to-sound mapping is stronger in pianists (Experiment 1). Additionally, the same effect was found when pianists merely observed another person pressing keys on a piano keyboard (Experiment 2).

Other studies have shown that merely observing musician's body movements can alter perceptual and aesthetic judgments of the produced music. Schutz and Lipscomb (2007) examined to what extent visual information of a marimba player's gestures can influence the perception of the duration of the produced tone. For the experiment, video recordings were made of a marimba player performing a series of tones using two stroke types ("long" gesture and "short" gesture). The tones that were produced by both stroke types were acoustically indistinguishable. The visual and auditory components were separated from each other and fully crossed in order to create realistic musical stimuli. Then, participants were asked to indicate perceived tone duration by means of a 101-point slider. In an audio-only condition, no significant differences occurred between the ratings. However, in the audio-visual condition, participants rated the tones produced with "long" gestures as significantly longer than the tones produced with "short" gestures. In another study Thompson et al. (2005) showed that facial expressivity and expressive hand gestures of music performers (i.e., vocal and guitar performance) can influence listeners' auditory perception of musical dissonance, melodic interval size, and affective valence. Similar findings are provided by Juchniewicz (2008), showing that the type of physical movement exhibited by a piano player while performing a musical excerpt (i.e., "no movement," "head and facial movement," and "full body movement") alters listeners' judgments of the piano performance in terms of phrasing, dynamics, rubato and overall musical performance.

3.3. MOTOR DISORDERS

The previously discussed action-based effects on auditory perception were rooted in learned auditory-motor associations. Apart from that, another category of action-based effects can be distinguished. Several studies have shown that motor dysfunction leads to considerable changes in individuals' perception and recognition of auditory and musical features. Pazzaglia et al. (2008) claimed a causative link between auditory recognition and action execution. Working with apraxia patients (limb apraxia, buccofacial apraxia, or both), they showed that deficits in performing gestures are causally linked to the patients' inability to recognize these gestures by their mere sounds. In the study, apraxia patients were asked to listen to a sound and then choose from among four pictures the one corresponding to the heard sound. Limb and buccofacial apraxia patients were impaired in recognizing sounds linked to respectively limb and buccofacial human actions. The authors advocated that lesions in frontal and parietal brain areas, which are actively associated with deficits in execution tasks, were responsible for the observed gesture-comprehension deficits. Also, studies indicate that the perception of musical features is impaired by motor dysfunctions. Beste et al. (2011) demonstrated effects of movement deterioration on rhythm processing in Huntington's disease patients. While listening to music, patients exhibited weaker activations overall in brain areas involved in the assessment of musical rhythms (cerebellar structures). Also, a study of Parkinson's disease patients by Grahn and Brett (2009) found that basal ganglia dysfunction results in an impairment of the processing of rhythms that have a beat. However, as the authors discuss, it cannot be excluded that pathological factors other than movement deterioration may contribute to impaired rhythm processing. For instance dopamine depletion, typical for Parkinson's disease, has been shown to affect emotional processing (Lotze et al., 2009), which may further modulate the processing of rhythms. In another study by Lucas et al. (2013), impaired temporal information processing in Parkinson's disease patients has been ascribed to a deficit in the process of sensorimotor integration. These and other studies (see e.g., Grahn, 2012 for a review) demonstrate that rhythm perception involves a close link between auditory and motor processes. The existence of such links has been exploited for motor rehabilitation purposes in the domain of Parkinson's disease, Huntington's disease, and stroke. In this context, musical activities involving movement (control) and rhythm (perception) have been shown to improve general motor performance in Parkinson's disease patients (Nombela et al., 2013a,b) and stroke patients (Altenmüller et al., 2009). It would be interesting to investigate further to what extent improvements in motor skill benefit performance on perceptual tasks.

4. DISCUSSION

Traditionally, body movements—whether performed by a music performer or by a listener—were considered as the mere output of internal cognitive processes that involved a system of symbolic representations. Only recently, empirical evidence has begun to appear indicating that the human motor system and its actions may actually modulate people's experience, perception, and understanding of sound and music. The present article

was intended to provide a theoretical framework in which action-based effects on auditory perception may be understood. Additionally, the article serves as a review in which we investigate how the theory applies to recent empirical findings. The presented theoretical framework is centered around the common coding theory (Prinz, 1990; Hommel et al., 2001). The basic assertion of this theory is that the planning or execution of an action recruits the same sensory-motor brain areas as the mere perception of the sensory consequences of that action. We have argued that associative learning, in which actions and sensory states are repeatedly experienced together, are of crucial importance in order for action and perception to become integrated, and to form so-called internal models. These internal models contain inverse and forward components. Inverse models allow incoming sensory information to activate the motor codes associated with the production of that sensory state (cf. direct-matching hypothesis Rizzolatti et al., 2001). In contrast, forward models allow the sensory outcomes to be predicted from planned actions (Waszak et al., 2012). The combination of inverse and forward models regulate goal-directed motor control (Wolpert et al., 1995; Hommel, 1997), as well as the processing of sensory information coming from the external environment (Halász and Cunnington, 2012). We explained that both inverse and forward models contribute to action-based effects on auditory perception. Inverse models allow that mere listening to music results in the activation of motor codes, which is often manifested in overt movement responses (cf. motor simulation, motor resonance, action mirroring, etc.). These body movements are experienced and understood as intentionally, expressively, and semantically meaningful, and cause the music to be experienced and understood accordingly. Forward models have an impact on music perception in a different way. They allow us to make predictions about the auditory outcomes of planned or executed actions, which guide and shape the perception of sound and music. Predictions may either attenuate, facilitate, or disambiguate the perception of sound and music. Together, these findings show that the human motor system and its actions have an impact on music perception and cognition. It is tempting to conclude based on this evidence that the "musical mind" is fundamentally embodied. However, according to Wilson and Golonka (2013), the assertion that (music) cognition is embodied has more radical and far-ranging implications. They claim that "embodiment is not simply another factor acting on otherwise disembodied cognitive processes." This would retain the traditional Cartesian view that the brain is in control and, in the case of people's engagement in musical activities, literally "runs the show." Instead, "radical embodiment" encompasses a perspective on the body, the mind, and the environment as substantial elements of a dynamical system (Chemero, 2009). In essence, the term "dynamical system" points to a system of elements which are coupled, mutually interactive, and evolve over time (Thelen and Smith, 1998). An important feature of dynamical systems is the ability to self-organize. Order and coherence appear out of the mutual interactions of the elements of the system without the use of explicit instructions, representations, or symbols. The dynamical system approach can be applied to motor control and development (Turvey, 1990; Kelso, 1995; Thelen and Smith, 1998; Warren, 2006), as well as to

cognition (Port and Van Gelder, 1995; Van Gelder, 1998; Beer, 2000; Chemero, 2009; McClelland et al., 2010; Shapiro, 2013). Music seems especially relevant as many musical activities— e.g., music production, dance, music listening—provide an ecological setting in which the intrinsic dynamics of action and perception can be studied (Bader, 2013a,b). Moreover, it is interesting to note that people's engagement with music involves not only sensory and motor components but also other components, such as "introspection"—referring to internal states that include affect, motivation, intentions, metacognition, etc. (Barsalou, 2009)—and "social interaction." Currently, research on internal models focuses almost exclusively on sensory and motor processes. However, to explain people's interaction with music, and by extension with the world in general, it is necessary to include aspects of introspection and social interaction into theories on internal models. The integration of these aspects into the present theoretical framework can deepen our understanding of music, and of the musical mind as fundamentally embodied. In the following paragraphs, we briefly discuss these two components.

4.1. MUSICAL EXPRESSIVITY

An important aspect of people's engagement with music— whether in listening to music or the actual production of music— is musical expressivity. The musical elements that are said to constitute musical expressivity are manifold: dynamics, articulation, touch, phrasing, vibrato, etc. In the case of music production, musical expressivity is often—but not exclusively—related to the contents of the composition, and the main task of the musician is to render the composition into sound. Of course there is always a certain degree of interpretation and expressivity from the performer's side. Music performance however does not necessarily rely on a pre-composed score, as in the case of improvisation or jam sessions, where music may be created for the sake of exploring different sounds, rhythms, dynamics, etc. Apart from whether music is the result of playing a composition or improvisation, what is conspicuous about many of the various elements contributing to musical expressivity is that they directly relate to their physical origin, namely the body movements that produced the music (Repp, 1993; Shove and Repp, 1995; Johnson, 1997; Godøy, 2003; Leman, 2007; Cox, 2011). Accordingly, musical expressivity can be said to appeal to, at least to some extent, *kinaesthetic sensations* related to the effort and shape of body movements (Laban, 1947; Laban and Ullmann, 1966). Further, this kinaesthetic sensitivity may be associated with subjective phenomena like feeling, emotion, intentionality, etc. (Leman, 2007; Cochrane, 2010; Sievers et al., 2013). In that sense, the human body has been considered as a mediator between sensory and motor processes and mental states (Leman, 2007). A similar role has been attributed to the body in the context of music listening. A listener is assumed to be able to decode—i.e., identify, imagine, or even physically render—the elements of musical expressivity that relate to physical motion and space, based on their own action repertoire and notion of space. This kinaesthetic sensitivity may be related to subjective mental aspects of feeling, emotion, intentionality, etc. In the same way as planning or executing an action enables people to make predictions of the sensory consequences of that action, it

is possible to make predictions of the consequences on a mental level (e.g., feeling, emotion, intentionality). Accordingly, it is reasonable to assume that the predictions of mental states modulate the perception of musical expressivity. It is only recently that empirical support for this idea emerged (Sedlmeier et al., 2011; Maes and Leman, 2013). Also, it has been shown that the visual observation of performers' body movements influences people's perception of musical expressivity (Davidson, 1993; Thompson et al., 2005; Juchniewicz, 2008). These findings provide support for including expressivity in theories of forward modeling applied to music perception and cognition. According to current theories of internal models, we have reason to believe that the relationship between mental states and action works in the opposite direction as well (cf. inverse models). In that sense, a subjective state coupled to music is assumed to modulate motor responses to music. Support for this idea is given in a study of Van Dyck et al. (2013). According to the current view, internal models guide goal-directed behavior as well as sensory processing. In that sense, internal models are the basic constituents of people's interaction with the outside world. We advocate that this view should be broadened by integrating other aspects of introspection (affect, motivation, intentions, metacognition, etc.). Musical behaviors provide opportunities to study interactions between sensory, motor and introspective processes, and the way these components become associated with each other. The current view of embodied music cognition considers introspection as a result of motor simulation processes (Leman, 2007). In other words, music induces body movements, which consequently trigger subjective aspects of feeling, emotion, intentionality, etc. We advocate that the relationship between body and mind may be bidirectional, as aspects of introspection may also influence motor responses to music.

4.2. SOCIAL INTERACTION

In daily life, much of what we do and experience happens in a social context. A paramount example is people's engagement with music, as in music ensemble playing (Bastien and Hostager, 1988; Seddon, 2005), or when people dance together in a club or festival. These activities can be considered as forms of joint action involving coordinated actions, shared intention, shared attention, shared representations, etc. (Keller, 2008; Goebl and Palmer, 2009; Loehr and Palmer, 2011; Obhi and Sebanz, 2011; Pacherie, 2012; Phillips-Silver and Keller, 2012). Joint action in the context of music playing and dance has been shown to promote social behavior (Kirschner and Tomasello, 2010) and to establish a heightened sense of agency and a sense of we-ness (Pacherie, 2012). Also, studies show that the social context may modulate people's experience and perception of music (Egermann et al., 2011; Liljeström et al., 2012). Currently, a major line of research is devoted to the study of joint action in order to unveil the underlying mechanisms. Accumulating evidence suggests that these mechanisms are similar to the ones involved in individual voluntary motor control and information processing. Accordingly, internal models containing an inverse and forward component may explain how people manage to dynamically adapt to changes in each other's behavior. Inverse models are important for rendering desired joint action sensory outcomes into particular action plans. Supplementary forward models facilitate to anticipate (predict) the sensory consequences of one's own and other's actions.

Our discussion of the components of "introspection" and "social interaction" indicates that musical activities involve a high-dimensional dynamical system in which the body, the mind, and the external environment are continually and mutually interacting. In the case of musical instrument playing, music can be considered as the result of a dynamical interaction between the musicians' motor and sensory system, the constraints and opportunities of the pre-composed musical notation, the musical instruments and the social environment, and the musicians' intentions, personality, mental states, etc. The system in which these components interact is an open system, in the sense that no individual component has causal priority in generating the music (Thelen and Smith, 1998). It is possible, however, that the weight of the individual components on the produced sound varies depending on the specific musical activity (e.g., musical improvisation, historical informed music performance, jam session with an emphasis on social interaction, etc.). Similarly, music listening can be considered as a dynamical process, in which the experience, the perception, and the understanding of music is guided and shaped by the intrinsic dynamics of the body, the mind, and the external environment. In conclusion, adopting a fundamental embodied approach to music cognition requires us to consider music performance—involving motor coordination, control, and development—and music cognition as dynamical processes. The integration of theories on internal models and theories on dynamical systems can thereby enhance our understanding of how our body, mind, and the external environment interact in our engagement with the act of music.

REFERENCES

Aliu, S. O., Houde, J. F., and Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *J. Cogn. Neurosci.* 21, 791–802. doi: 10.1162/jocn.2009.21055

Altenmüller, E., Marco-Pallares, J., Münte, T., and Schneider, S. (2009). Neural reorganization underlies improvement in stroke-induced motor dysfunction by music-supported therapy. *Ann. N.Y. Acad. Sci.* 1169, 395–405. doi: 10.1111/j.1749-6632.2009.04580.x

Bader, R. (2013a). *Nonlinearities and Synchronization in Musical Acoustics and Music Psychology, Vol. 2 of Current Research in Systematic Musicology.* Heidelberg: Springer.

Bader, R. (2013b). "Synchronization and self-organization as basis of musical performance, sound production, and perception," in *Sound-perception-performance, Vol. 1 of Current Research in Systematic Musicology*, ed R. Bader (Heidelberg: Springer), 3–42.

Baess, P., Horváth, J., Jacobsen, T., and Schröger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: an ERP study. *Psychophysiology* 48, 1276–1283. doi: 10.1111/j.1469-8986.2011.01196.x

Baess, P., Jacobsen, T., and Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: evidence for internal forward models with dynamic stimulation. *Int. J. Psychophysiol.* 70, 137–143. doi: 10.1016/j.ijpsycho.2008.06.005

Bangert, M., and Altenmüller, E. (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci.* 4:26. doi: 10.1186/1471-2202-4-26

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1281–1289. doi: 10.1098/rstb.2008.0319

Bastien, D. T., and Hostager, T. J. (1988). Jazz as a process of organizational innovation. *Commun. Res.* 15, 582–602. doi: 10.1177/009365088015005005

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., and Jancke, L. (2007). A network for audio–motor coordination in skilled pianists and non-musicians. *Brain Res.* 1161, 65–78. doi: 10.1016/j.brainres.2007.05.045

Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends Cogn. Sci.* 4, 91–99. doi: 10.1016/S1364-6613(99)01440-0

Bernardi, N. F., De Buglio, M., Trimarchi, P. D., Chielli, A., and Bricolo, E. (2013). Mental practice promotes motor anticipation: evidence from skilled music performance. *Front. Hum. Neurosci.* 7:451. doi: 10.3389/fnhum.2013.00451

Beste, C., Schüttke, A., Pfleiderer, B., and Saft, C. (2011). Music perception and movement deterioration in Huntington's disease. *PLoS Curr.* 3:RRN1252. doi: 10.1371/currents.RRN1252

Blakemore, S.-J., Frith, C. D., and Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport* 12, 1879–1884. doi: 10.1097/00001756-200107030-00023

Brown, R. M., and Palmer, C. (2012). Auditory-motor learning influences auditory memory for music. *Mem. Cogn.* 40, 567–578. doi: 10.3758/s13421-011-0177-x

Brown, R. M., and Palmer, C. (2013). Auditory and motor imagery modulate learning in music performance. *Front. Hum. Neurosci.* 7:320. doi: 10.3389/fnhum.2013.00320

Bubic, A., Von Cramon, D. Y., and Schubotz, R. I. (2010). Prediction, cognition and the brain. *Front. Hum. Neurosci.* 4:25. doi: 10.3389/fnhum.2010.00025

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S., and Toiviainen, P. (2013). Influences of rhythm-and timbre-related musical features on characteristics of music-induced movement. *Front. Psychol.* 4:183. doi: 10.3389/fpsyg.2013.00183

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., and Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb. Cortex* 15, 1243–1249. doi: 10.1093/cercor/bhi007

Caramiaux, B., Bevilacqua, F., and Schnell, N. (2010). "Towards a gesture-sound cross-modal analysis," in *Gesture in embodied communication and human-computer interaction (LNCS)*, Vol. 5934, eds S. Kopp and I. Wachsmuth (Berlin: Springer-Verlag), 158–170.

Catmur, C. (2012). Sensorimotor learning and the ontogeny of the mirror neuron system. *Neurosci. Lett.* 540, 21–27. doi: 10.1016/j.neulet.2012.10.001

Chemero, A. (2009). *Radical Embodied Cognitive Science.* Cambridge, MA: MIT press.

Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* 18, 2844–2854. doi: 10.1093/cercor/bhn042

Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2009). The role of auditory and premotor cortex in sensorimotor transformations. *Ann. N.Y. Acad. Sci.* 1169, 15–34. doi: 10.1111/j.1749-6632.2009.04556.x

Cochrane, T. (2010). A simulation theory of musical expressivity. *Austr. J. Philos.* 88, 191–207. doi: 10.1080/00048400902941257

Cook, R., Press, C., Dickinson, A., and Heyes, C. (2010). Acquisition of automatic imitation is sensitive to sensorimotor contingency. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 840–852. doi: 10.1037/a0019256

Cooper, R. P., Cook, R., Dickinson, A., and Heyes, C. M. (2012). Associative (not Hebbian) learning and the mirror neuron system. *Neurosci. Lett.* 540, 28–36. doi: 10.1016/j.neulet.2012.10.002

Cox, A. (2011). Embodying music: principles of the mimetic hypothesis. *Music Theory Online* 17, 1–24.

Cox, R. F., and Hasselman, F. (2013). The case of Watson vs. James: effect-priming studies do not support ideomotor theory. *PLoS ONE* 8:e54094. doi: 10.1371/journal.pone.0054094

Croom, A. M. (2011). Music, neuroscience, and the psychology of well-being: a précis. *Front. Psychol.* 2:393. doi: 10.3389/fpsyg.2011.00393

Csikszentmihalyi, M. (1988). "The flow experience and its significance for human psychology," in *Optimal Experience: Psychological Studies of Flow in Consciousness*, eds M. Csikszentmihalyi and S. Csikszentmihalyi (Cambridge, UK: Cambridge University Press), 15–35.

Dalla Bella, S., and Palmer, C. (2011). Rate effects on timing, key velocity, and finger kinematics in piano performance. *PLoS ONE* 6:e20518. doi: 10.1371/journal.pone.0020518

D'Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., and Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *Eur. J. Neurosci.* 24, 955–958. doi: 10.1111/j.1460-9568.2006.04960.x

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychol. Music* 21, 103–113. doi: 10.1177/030573569302100201

Davidson, P. R., and Wolpert, D. M. (2005). Widespread access to predictive models in the motor system: a short review. *J. Neural Eng.* 2, S313–S319. doi: 10.1088/1741-2560/2/3/S11

Deiber, M.-P., Wise, S., Honda, M., Catalan, M., Grafman, J., and Hallett, M. (1997). Frontal and parietal networks for conditional motor learning: a positron emission tomography study. *J. Neurophysiol.* 78, 977–991.

Doyon, J., Bellec, P., Amsel, R., Penhune, V., Monchi, O., Carrier, J., Lehéricy, S., and Benali, H. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behav. Brain Res.* 199, 61–75. doi: 10.1016/j.bbr.2008.11.012

Ebner, T. J. (2013). "Cerebellum and internal models," in *Handbook of the Cerebellum and Cerebellar Disorders*, eds M. Manto, J. D. Schmahmann, F. Rossi, D. L. Gruol, and N. Koibuchi (Heidelberg: Springer), 1279–1295. doi: 10.1007/978-94-007-1333-8_56

Egermann, H., Sutherland, M. E., Grewe, O., Nagel, F., Kopiez, R., and Altenmüller, E. (2011). Does music listening in a social context alter experience? A physiological and psychological perspective on emotion. *Musicae Scientiae* 15, 307–323. doi: 10.1177/1029864911399497

Eitan, Z., and Granot, R. Y. (2006). How music moves: musical parameters and listeners images of motion. *Music Percept.* 23, 221–248. doi: 10.1525/mp.2006.23.3.221

Elsner, B., and Hommel, B. (2001). Effect anticipation and action control. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 229–240. doi: 10.1037/0096-1523.27.1.229

Elsner, B., and Hommel, B. (2004). Contiguity and contingency in action-effect learning. *Psychol. Res.* 68, 138–154. doi: 10.1007/s00426-003-0151-8

Fodor, J. A. (1975). *The Language of Thought.* Cambridge, MA: Harvard University Press.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609. doi: 10.1093/brain/119.2.593

Gaser, C., and Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *J. Neurosci.* 23, 9240–9245. doi: 10.1016/S1053-8119(01)92488-7

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdiscipl. Rev. Cogn. Sci.* 1, 586–596. doi: 10.1002/wcs.55

Godøy, R. I. (2003). Motor-mimetic music cognition. *Leonardo* 36, 317–319. doi: 10.1162/002409403322258781

Godøy, R. I. (2010). "Gestural affordances of musical sound," in *Musical Gestures: Sound, Movement, and Meaning*, eds R. I. Godøy and M. Leman (New York, NY: Routledge), 103–125.

Godøy, R. I., and Leman, M. (2010). *Musical Gestures: Sound, Movement, and Meaning.* New York, NY: Routledge.

Goebl, W., and Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Percept.* 26, 427–438. doi: 10.1525/mp.2009.26.5.427

Grahn, J. A. (2012). Neural mechanisms of rhythm perception: current findings and future perspectives. *Top. Cogn. Sci.* 4, 585–606. doi: 10.1111/j.1756-8765.2012.01213.x

Grahn, J. A., and Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex* 45, 54–61. doi: 10.1016/j.cortex.2008.01.005

Grahn, J. A., and Rowe, J. B. (2013). Finding and feeling the musical beat: striatal dissociations between detection and prediction of regularity. *Cereb. Cortex* 23, 913–921. doi: 10.1093/cercor/bhs083

Halász, V., and Cunnington, R. (2012). Unconscious effects of action on perception. *Brain Sci.* 2, 130–146. doi: 10.3390/brainsci2020130

Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., and Ceballos-Baumann, A. O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *J. Cogn. Neurosci.* 17, 282–293. doi: 10.1162/0898929053124893

Haueisen, J., and Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by Music Perception. *J. Cogn. Neurosci.* 13, 786–792. doi: 10.1162/08989290152541449

Heinks-Maldonado, T. H., Nagarajan, S. S., and Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport* 17, 1375–1379. doi: 10.1097/01.wnr.0000233102.43526.e9

Herholz, S. C., and Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76, 486–502. doi: 10.1016/j.neuron.2012.10.011

Heyes, C. M. (2010). Where do mirror neurons come from? *Neurosci. Biobehav. Rev* 34, 575–583. doi: 10.1016/j.neubiorev.2009.11.007

Heyes, C. M., and Ray, E. D. (2000). What is the significance of imitation in animals? *Adv. Study Behav.* 29, 215–245. doi: 10.1016/S0065-3454(08)60106-0

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21, 1229–1243. doi: 10.1162/jocn.2009.21189

Hommel, B. (1997). Toward an action-concept model of stimulus-response compatibility. *Adv. Psychol.* 118, 281–320. doi: 10.1016/S0166-4115(97)80041-6

Hommel, B. (2003). "Acquisition and control of voluntary action," in *Voluntary Action: Brains, Minds, and Sociality*, eds S. Maasen, W. Prinz, and G. Roth (Oxford, UK: Oxford University Press), 34–48.

Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–878. doi: 10.1017/S0140525X01000103

Houde, J. F., Nagarajan, S. S., Sekihara, K., and Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. *J. Cogn. Neurosci.* 14, 1125–1138. doi: 10.1162/089892902760807140

Hughes, G., Desantis, A., and Waszak, F. (2013a). Attenuation of auditory N1 results from identity-specific action-effect prediction. *Eur. J. Neurosci.* 37, 1152–1158. doi: 10.1111/ejn.12120

Hughes, G., Desantis, A., and Waszak, F. (2013b). Mechanisms of intentional binding and sensory attenuation: the role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychol. Bull.* 139, 133–151. doi: 10.1037/a0028566

Hurley, S. (2001). Perception and action: alternative views. *Synthese* 129, 3–40. doi: 10.1023/A:1012643006930

Hurley, S. (2008). The shared circuits model (SCM): how control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behav. Brain Sci.* 31, 1–21. doi: 10.1017/S0140525X07003123

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., and Schlaug, G. (2009). The effects of musical training on structural brain development. *Ann. N.Y. Acad. Sci.* 1169, 182–186. doi: 10.1111/j.1749-6632.2009.04852.x

Imamizu, H., and Kawato, M. (2009). Brain mechanisms for predictive control by switching internal models: implications for higher-order cognitive functions. *Psychol. Res.* 73, 527–544. doi: 10.1007/s00426-009-0235-1

Iordanescu, L., Grabowecky, M., and Suzuki, S. (2013). Action enhances auditory but not visual temporal sensitivity. *Psychon. Bull. Rev.* 20, 108–114. doi: 10.3758/s13423-012-0330-y

Iversen, J. R., Repp, B. H., and Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Ann. N.Y. Acad. Sci.* 1169, 58–73. doi: 10.1111/j.1749-6632.2009.04579.x

James, W. (1890). *The Principles of Psychology*. Vol. 1. New York, NY: Henry Holt. doi: 10.1037/11059-000

Janata, P., Tomic, S. T., and Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *J. Exp. Psychol. Gen.* 141, 54–75. doi: 10.1037/a0024208

Johnson, M. L. (1997). Embodied musical meaning. *Theory Pract.* 22, 95–102.

Jones, A., Hughes, G., and Waszak, F. (2013). The interaction between attention and motor prediction. An ERP study. *Neuroimage* 83, 533–541. doi: 10.1016/j.neuroimage.2013.07.004

Juchniewicz, J. (2008). The influence of physical movement on the perception of musical performance. *Psychol. Music* 36, 417–427. doi: 10.1177/0305735607086046

Juslin, P. N., and Västfjäll, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *Behav. Brain Sci.* 31, 559–621. doi: 10.1017/S0140525X08005293

Keller, P. E. (2008). "Joint action in music performance," in *Enacting Intersubjectivity: a Cognitive and Social Perspective on the Study of Interactions*, eds F. Morganti, A. Carassa, and G. Riva (Amsterdam: IOS press), 205–221.

Kelso, J. A. (1995). *Dynamic Patterns: The Self Organization of Brain and Behaviour*. Cambridge, MA: MIT Press.

Kirschner, S., and Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evol. Hum. Behav.* 31, 354–364. doi: 10.1016/j.evolhumbehav.2010.04.004

Knolle, F., Schröger, E., Baess, P., and Kotz, S. A. (2012a). The cerebellum generates motor-to-auditory predictions: ERP lesion evidence. *J. Cogn. Neurosci.* 24, 698–706. doi: 10.1162/jocn_a_00167

Knolle, F., Schröger, E., and Kotz, S. A. (2012b). Prediction errors in self-and externally-generated deviants. *Biol. Psychol.* 92, 410–416. doi: 10.1016/j.biopsycho.2012.11.017

Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297, 846–848. doi: 10.1126/science.1070311

Kozak, M., Nymoen, K., and Godøy, R. I. (2012). "Effects of spectral features of sound on gesture type and timing," in *Gesture and Sign Language in Human-Computer Interaction and Embodied Communication (LNCS)*, Vol. 7206, eds E. Efthimiou, G. Kouroupetroglou, and S.-E. Fotinea (Berlin: Springer-Verlag), 69–80. doi: 10.1007/978-3-642-34182-3_7

Krueger, J. (2009). Enacting musical experience. *J. Conscious. Stud.* 16, 2–3. doi: 10.1093/acprof:oso/9780199654888.003.0014

Küssner, M. B. (2013). Music and shape. *Lit. Linguist. Comput.* 28, 472–479. doi: 10.1093/llc/fqs071

Laban, R. (1947). *Effort*. London: MacDonald & Evans.

Laban, R., and Ullmann, L. (1966). *Choreutics*. Hampshire: Dance Books Ltd. (Edition published in 2011).

Lahav, A., Boulanger, A., Schlaug, G., and Saltzman, E. (2005). The power of listening: auditory-motor interactions in musical training. *Ann. N.Y. Acad. Sci.* 1060, 189–194. doi: 10.1196/annals.1360.042

Lahav, A., Saltzman, E., and Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J. Neurosci.* 27, 308–314. doi: 10.1523/JNEUROSCI.4822-06.2007

Lalazar, H., and Vaadia, E. (2008). Neural basis of sensorimotor learning: modifying internal models. *Curr. Opin. Neurobiol.* 18, 573–581. doi: 10.1016/j.conb.2008.11.003

Lappe, C., Steinsträter, O., and Pantev, C. (2013). Rhythmic and melodic deviations in musical sequences recruit different cortical areas for mismatch detection. *Front. Hum. Neurosci.* 7:260. doi: 10.3389/fnhum.2013.00260

Laske, O. E. (1974). The information-processing approach to musical cognition. *J. New Music Res.* 3, 109–136.

Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., and Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *J. Neurosci.* 29, 2477–2485. doi: 10.1523/JNEUROSCI.4921-08.2009

Leman, M. (2007). *Embodied music Cognition and Mediation Technology*. Cambridge, MA: MIT Press.

Leman, M., Desmet, F., Styns, F., Van Noorden, L., and Moelants, D. (2009). Sharing musical expression through embodied listening: a case study based on Chinese guqin music. *Music Percept.* 26, 263–278. doi: 10.1525/mp.2009.26.3.263

Leman, M., Moelants, D., Varewyck, M., Styns, F., van Noorden, L., and Martens, J.-P. (2013). Activating and relaxing music entrains the speed of beat synchronized walking. *PLoS ONE* 8:e67932. doi: 10.1371/journal.pone.0067932

Levinson, J. (2006). *Contemplating Art: Essays in Aesthetics*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780199206179.001.0001

Lidji, P., Kolinsky, R., Lochy, A., and Morais, J. (2007). Spatial associations for musical stimuli: a piano in the head? *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1189–1207. doi: 10.1037/0096-1523.33.5.1189

Liljeström, S., Juslin, P. N., and Västfjäll, D. (2012). Experimental evidence of the roles of music choice, social context, and listener personality in emotional reactions to music. *Psychol. Music* 41, 577–597. doi: 10.1177/0305735612440615

Loehr, J. D. (2013). Sensory attenuation for jointly produced action effects. *Front. Psychol.* 4:172. doi: 10.3389/fpsyg.2013.00172

Loehr, J. D., and Palmer, C. (2011). Temporal coordination between performing musicians. *Q. J. Exp. Psychol.* 64, 2153–2167. doi: 10.1080/17470218.2011.603427

Lotze, M. (2013). Kinesthetic imagery of musical performance. *Front. Hum. Neurosci.* 7:280. doi: 10.3389/fnhum.2013.00280

Lotze, M., Reimold, M., Heymans, U., Laihinen, A., Patt, M., and Halsband, U. (2009). Reduced ventrolateral fmri response during observation of emotional gestures related to the degree of dopaminergic impairment in parkinson disease. *J. Cogn. Neurosci.* 21, 1321–1331. doi: 10.1162/jocn.2009.21087

Lotze, M., Scheler, G., Tan, H.-R., Braun, C., and Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage* 20, 1817–1829. doi: 10.1016/j.neuroimage.2003.07.018

Lucas, M., Chaves, F., Teixeira, S., Carvalho, D., Peressutti, C., Bittencourt, J., et al. (2013). Time perception impairs sensory-motor integration in Parkinson's disease. *Int. Arch. Med.* 6, 1–7. doi: 10.1186/1755-7682-6-39

Luft, A. R., and Buitrago, M. M. (2005). Stages of motor skill learning. *Mol. Neurobiol.* 32, 205–216. doi: 10.1385/MN:32:3:205

Maes, P.-J., Amelynck, D., and Leman, M. (2012). Dance-the-Music: an educational platform for the modeling, recognition and audiovisual monitoring of dance steps using spatiotemporal motion templates. *EURASIP J. Adv. Signal Process.* 2012, 1–16. doi: 10.1186/1687-6180-2012-35

Maes, P.-J., Amelynck, D., Lesaffre, M., Leman, M., and Arvind, D. (2013). The "Conducting Master": an interactive, real-time gesture monitoring system based on spatiotemporal motion templates. *Int. J. Hum. Comput. Interact.* 29, 471–487. doi: 10.1080/10447318.2012.720197

Maes, P.-J., and Leman, M. (2013). The influence of body movements on children's perception of music with an ambiguous expressive character. *PLoS ONE* 8:e54682. doi: 10.1371/journal.pone.0054682

Maes, P.-J., Leman, M., Lesaffre, M., Demey, M., and Moelants, D. (2010). From expressive gesture to sound: the development of an embodied mapping trajectory inside a musical interface. *J. Multimodal User Interfaces* 3, 67–78. doi: 10.1007/s12193-009-0027-3

Maes, P.-J., Van Dyck, E., Lesaffre, M., Leman, M., and M, K. P. (2014). The coupling of action and perception in musical meaning formation. *Music Percept.* (forthcoming).

Manning, F., and Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychon. Bull. Rev.* 20, 1133–1139. doi: 10.3758/s13423-013-0439-7

Massaro, D. W. (1990). "An information-processing analysis of perception and action," in *Relationships between Perception and Action: Current Approaches*, eds O. Neumann and W. Prinz (Berlin: Springer-Verlag), 133–166. doi: 10.1007/978-3-642-75348-0_6

Maus, F. (1988). Music as drama. *Music Theory Spectr.* 10, 56–73. doi: 10.2307/745792

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., et al. (2010). Letting structure emerge: connectionist and dynamical systems approaches to cognition. *Trends Cogn. Sci.* 14, 348–356. doi: 10.1016/j.tics.2010.06.002

Mead, A. (1999). Bodily hearing: physiological metaphors and musical understanding. *J. Music Theory* 43, 1–19. doi: 10.2307/3090688

Melcher, T., Winter, D., Hommel, B., Pfister, R., Dechent, P., and Gruber, O. (2012). The neural substrate of the ideomotor principle revisited: evidence for asymmetries in action-effect learning. *Neuroscience* 231, 13–27. doi: 10.1016/j.neuroscience.2012.11.035

Naveda, L., and Leman, M. (2009). A cross-modal heuristic for periodic pattern analysis of samba music and dance. *J. New Music Res.* 38, 255–283. doi: 10.1080/09298210903105432

Naveda, L., and Leman, M. (2010). The spatiotemporal representation of dance and music gestures using topological gesture analysis (TGA). *Music Percept.* 28, 93–111. doi: 10.1525/mp.2010.28.1.93

Neisser, U. (1967). *Cognitive Psychology.* New York, NY: Appleton-Century-Crofts.

Nelson, A., Schneider, D. M., Takatoh, J., Sakurai, K., Wang, F., and Mooney, R. (2013). A circuit for motor cortical modulation of auditory cortical activity. *J. Neurosci.* 33, 14342–14353. doi: 10.1523/JNEUROSCI.2275-13.2013

Nombela, C., Hughes, L. E., Owen, A. M., and Grahn, J. A. (2013a). Into the groove: can rhythm influence parkinson's disease? *Neurosci. Biobehav. Rev.* 37, 2564–2570. doi: 10.1016/j.neubiorev.2013.08.003

Nombela, C., Rae, C., Grahn, J., Barker, R., Owen, A., Rowe, J., et al. (2013b). How often does music and rhythm improve patients' perception of motor symptoms in Parkinson's disease? *J. Neurol.* 260, 1404–1405. doi: 10.1007/s00415-013-6860-z

Obhi, S. S., and Sebanz, N. (2011). Moving together: toward understanding the mechanisms of joint action. *Exp. Brain Res.* 211, 329–336. doi: 10.1007/s00221-011-2721-0

O'Reilly, J. X., Jbabdi, S., Rushworth, M. F., and Behrens, T. E. (2013). Brain systems for probabilistic and dynamic prediction: computational specificity and integration. *PLoS Biol.* 11:e1001662. doi: 10.1371/journal.pbio.1001662

Pacherie, E. (2012). "The phenomenology of joint action: self-agency vs. joint-agency," in *Joint Attention: New Developments*, ed A. Seemann (Cambridge, MA: MIT Press), 343–389.

Pascal-Leone, A. (2001). The brain that plays music and is changed by it. *Ann. N.Y. Acad. Sci.* 930, 315–329. doi: 10.1111/j.1749-6632.2001.tb05741.x

Pasupathy, A., and Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433, 873–876. doi: 10.1038/nature03287

Pazzaglia, M., Pizzamiglio, L., Pes, E., and Aglioti, S. M. (2008). The sound of actions in apraxia. *Curr. Biol.* 18, 1766–1772. doi: 10.1016/j.cub.2008.09.061

Phillips-Silver, J., and Keller, P. E. (2012). Searching for roots of entrainment and joint action in early musical interactions. *Front. Hum. Neurosci.* 6:26. doi: 10.3389/fnhum.2012.00026

Phillips-Silver, J., and Trainor, L. J. (2005). Feeling the beat: movement influences infant rhythm perception. *Science* 308, 1430. doi: 10.1126/science.1110922

Phillips-Silver, J., and Trainor, L. J. (2007). Hearing what the body feels: auditory encoding and rhythmic movement. *Cognition* 105, 533–546. doi: 10.1016/j.cognition.2006.11.006

Pierce, A. (2007). *Deepening Musical Performance through Movement: the Theory and Practice of Embodied Interpretation.* Bloomington, IN: Indiana University Press.

Port, R. F., and Van Gelder, T., editors (1995). *Mind as Motion: Explorations in the Dynamics of Cognition.* Cambridge, MA: MIT Press.

Prinz, W. (1990). "A common coding approach to perception and action," in *Relationships Between Perception and Action: Current Approaches*, eds O. Neumann and W. Prinz (Heidelberg: Springer), 167–201. doi: 10.1007/978-3-642-75348-0_7

Prinz, W. (1997). Perception and action planning. *Eur. J. Cogn. Psychol.* 9, 129–154. doi: 10.1080/713752551

Pylyshyn, Z. W., and Demopoulos, W., editors (1986). *Meaning and Cognitive Structure: Issues in the Computational Theory of Mind.* Norwood, NJ: ABLEX Publishing Company.

Repp, B. H. (1993). "Musical motion: some historical and contemporary perspectives," in *Proceedings of the Stockholm Music Acoustics Conference (SMAC)*, eds A. Friberg, J. Iwarsson, E. Jansson, and J. Sundberg (Stockholm: Kgl. Musikaliska Akademin), 128–135.

Repp, B. H., and Knoblich, G. (2009). Performed or observed keyboard actions affect pianists' judgements of relative pitch. *Q. J. Exp. Psychol.* 62, 2156–2170. doi: 10.1080/17470210902745009

Rizzolatti, G., Fogassi, L., Gallese, V., et al. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–669. doi: 10.1038/35090060

Roussel, C., Hughes, G., and Waszak, F. (2013). A preactivation account of sensory attenuation. *Neuropsychologia* 51, 922–929. doi: 10.1016/j.neuropsychologia.2013.02.005

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

Sanmiguel, I., Todd, J., and Schröger, E. (2013). Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. *Psychophysiology* 14, 1–11. doi: 10.1111/psyp.12024

Sato, A. (2008). Action observation modulates auditory perception of the consequence of others' actions. *Conscious. Cogn.* 17, 1219–1227. doi: 10.1016/j.concog.2008.01.003

Schäfer, T., Fachner, J., and Smukalla, M. (2013). Changes in the representation of space and time while listening to music. *Front. Psychol.* 4:508. doi: 10.3389/fpsyg.2013.00508

Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.* 11, 211–218. doi: 10.1016/j.tics.2007.02.006

Schutz, M., and Lipscomb, S. (2007). Hearing gestures, seeing music: vision influences perceived tone duration. *Perception* 36, 888–897. doi: 10.1068/p5635

Schütz-Bosbach, S., and Prinz, W. (2007). Perceptual resonance: action-induced modulation of perception. *Trends Cogn. Sci.* 11, 349–355. doi: 10.1016/j.tics.2007.06.005

Seddon, F. A. (2005). Modes of communication during jazz improvisation. *Br. J. Music Educ.* 22, 47–61. doi: 10.1017/S0265051704005984

Sedlmeier, P., Weigelt, O., and Walther, E. (2011). Music is in the muscle: how embodied cognition may influence music preferences. *Music Percept.* 28, 297–306. doi: 10.1525/mp.2011.28.3.297

Shapiro, L. (2010). *Embodied Cognition*. Boca Raton, FL: Taylor & Francis.

Shapiro, L. (2013). Dynamics and cognition. *Minds Mach.* 23, 353–375. doi: 10.1007/s11023-012-9290-2

Shin, Y. K., Proctor, R. W., and Capaldi, E. (2010). A review of contemporary ideomotor theory. *Psychol. Bull.* 136, 943–974. doi: 10.1037/a0020541

Shove, P., and Repp, B. H. (1995). "Musical motion and performance: theoretical and empirical perspectives," in *The Practice of Performance*, ed J. Rink (Cambridge, UK: Cambridge University Press), 55–83.

Sievers, B., Polansky, L., Casey, M., and Wheatley, T. (2013). Music and movement share a dynamic structure that supports universal expressions of emotion. *Proc. Natl. Acad. Sci. U.S.A.* 110, 70–75. doi: 10.1073/pnas.1209023110

Southgate, V. (2013). Do infants provide evidence that the mirror system is involved in action understanding? *Conscious. Cogn.* 22, 1114–1121. doi: 10.1016/j.concog.2013.04.008

Stewart, L., Verdonschot, R. G., Kajihara, T., and Sparks, J. (2013a). Action-perception coupling in violinists. *Front. Hum. Neurosci.* 7:349. doi: 10.3389/fnhum.2013.00349

Stewart, L., Verdonschot, R. G., Nasralla, P., and Lanipekun, J. (2013b). Action–perception coupling in pianists: learned mappings or spatial musical association of response codes (SMARC) effect? *Q. J. Exp. Psychol.* 66, 37–50. doi: 10.1080/17470218.2012.687385

Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., and Keller, P. E. (2013). Musical groove modulates motor cortex excitability: a TMS investigation. *Brain Cogn.* 82, 127–136. doi: 10.1016/j.bandc.2013.03.003

Su, Y.-H., and Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychol. Res.* 76, 373–382. doi: 10.1007/s00426-011-0346-3

Sudnow, D. (1978). *Ways of the Hand*. Cambridge, MA: MIT Press.

Thelen, E., and Smith, L. B. (1998). "Dynamic systems theories," in *Handbook of Child Psychology, Theoretical Models of Human Development, 6th Edn.*, eds W. Damon and R. M. Lerner (Chichester: Wiley Online Library), 258–312.

Thompson, W. F., Graham, P., and Russo, F. A. (2005). Seeing music performance: visual influences on perception and experience. *Semiotica* 156, 203–227. doi: 10.1515/semi.2005.2005.156.203

Timm, J., SanMiguel, I., Saupe, K., and Schröger, E. (2013). The N1-suppression effect for self-initiated sounds is independent of attention. *BMC Neurosci.* 14:2. doi: 10.1186/1471-2202-14-2

Timmann, D., Drepper, J., Frings, M., Maschke, M., Richter, S., Gerwig, M., and Kolb, F. (2010). The human cerebellum contributes to motor, emotional and cognitive associative learning: a review. *Cortex* 46, 845–857. doi: 10.1016/j.cortex.2009.06.009

Toiviainen, P., Luck, G., and Thompson, M. R. (2010). Embodied meter: hierarchical eigenmodes in music-induced movement. *Music Percept.* 28, 59–70. doi: 10.1525/mp.2010.28.1.59

Trimarchi, P. D., and Luzzatti, C. (2011). Implicit chord processing and motor representation in pianists. *Psychol. Res.* 75, 122–128. doi: 10.1007/s00426-010-0292-5

Turvey, M. T. (1990). Coordination. *Am. Psychol.* 45, 938–953. doi: 10.1037/0003-066X.45.8.938

Ungerleider, L. G., Doyon, J., and Karni, A. (2002). Imaging brain plasticity during motor skill learning. *Neurobiol. Learn. Mem.* 78, 553–564. doi: 10.1006/nlme.2002.4091

Urista, D. J. (2003). Beyond words: the moving body as a tool for musical understanding. *Music Theory Online* 9.

Van Dyck, E., Maes, P.-J., Hargreaves, J., Lesaffre, M., and Leman, M. (2013). Expressing induced emotions through free dance movement. *J. Nonverbal Behav.* 37, 175–190. doi: 10.1007/s10919-013-0153-1

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behav. Brain Sci.* 21, 615–628.

Varela, F. J., Thompson, E. T., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Warren, W. H. (2006). The dynamics of perception and action. *Psychol. Rev.* 113, 358–389. doi: 10.1037/0033-295X.113.2.358

Waszak, F., Cardoso-Leite, P., and Hughes, G. (2012). Action effect anticipation: neurophysiological basis and functional consequences. *Neurosci. Biobehav. Rev.* 36, 943–959. doi: 10.1016/j.neubiorev.2011.11.004

Weiss, C., and Schütz-Bosbach, S. (2012). Vicarious action preparation does not result in sensory attenuation of auditory action effects. *Conscious. Cogn.* 21, 1654–1661. doi: 10.1016/j.concog.2012.08.010

Williams, Z. M., and Eskandar, E. N. (2006). Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat. Neurosci.* 9, 562–568. doi: 10.1038/nn1662

Wilson, A. D., and Golonka, S. (2013). Embodied cognition is not what you think it is. *Front. Psychol.* 4:58. doi: 10.3389/fpsyg.2013.00058

Witt, J. K. (2011). Action's effect on perception. *Curr. Dir. Psychol. Sci.* 20, 201–206. doi: 10.1177/0963721411408770

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882. doi: 10.1126/science.7569931

Wolpert, D. M., Miall, R. C., and Kawato, M. (1998). Internal models in the cerebellum. *Trends Cogn. Sci.* 2, 338–347. doi: 10.1162/jocn.2009.21055

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.