Sensorimotor Grounding of Musical Embodiment and the Role of Prediction: A Review

Pieter-Jan Maes *

Department of Art, Music, and Theatre Sciences, IPEM, Ghent University, Belgium

Pieter-Jan Maes
holds a Ph.D. in Musicology from Ghent University, with a thesis on "An empirical study of embodied music listening, and its applications in mediation technology." He worked as a postdoctoral fellow at McGill University Montreal, Canada (2012–2013), performing research on timing in music performance. He is currently working as a postdoctoral researcher at IPEM studying the role of expressive body movement, auditory feedback, and prediction mechanisms in musical timing (music performance and auditory-motor synchronization).
pieterjan.maes@UGent.be

In a previous article, we reviewed empirical evidence demonstrating action-based effects on music perception to substantiate the musical embodiment thesis (Maes et al., 2014). Evidence was largely based on studies demonstrating that music perception automatically engages motor processes, or that body states/movements influence music perception. Here, we argue that more rigorous evidence is needed before any decisive conclusion in favor of a "radical" musical embodiment thesis can be posited. In the current article, we provide a focused review of recent research to collect further evidence for the "radical" embodiment thesis that music perception is a dynamic process firmly rooted in the natural disposition of sounds and the human auditory and motor system. Though, we emphasize that, on top of these natural dispositions, long-term processes operate, rooted in repeated sensorimotor experiences and leading to learning, prediction, and error minimization. This approach sheds new light on the development of musical repertoires, and may refine our understanding of action-based effects on music perception as discussed in our previous article (Maes et al., 2014). Additionally, we discuss two of our recent empirical studies demonstrating that music performance relies on similar principles of sensorimotor dynamics and predictive processing.

Keywords: embodied music cognition, music perception, music performance, dynamical systems, predictive coding

1. INTRODUCTION

Under the impetus of ecological and embodied approaches to music, it is now commonly agreed that bodily states and processes and perceptual-motor interactions are inherent to music cognition (Shove and Repp, 1995; Godøy, 2003; Reybrouck, 2005; Leman, 2007; Maes et al., 2014; Schiavio et al., 2015). It is noteworthy that lately, the embodied approach becomes progressively extended from mere perception and cognition, to related areas of musical affect and emotion (Cochrane, 2010; Krueger, 2013), reward and motivation (Chanda and Levitin, 2013; Zatorre and Salimpoor, 2013), and social interaction (Moran, 2014; D'Ausilio et al., 2015). However, despite serious advances in the field, improvements could be made in order to further develop a "radical" embodiment thesis to music cognition and interaction, and maybe more important, to substantiate this thesis with empirical evidence (Mahon and Caramazza, 2008; Chemero, 2011; Wilson and Golonka, 2013; Kiverstein and Miller, 2015). The crux of the radical embodied cognition thesis is the view that music cognition emerges from the real-time interaction of modality-specific processes serving perceptual or sensorimotor functions in the brain, bodily states and dynamics, and environment. In our previous review article, we collected studies demonstrating

that music perception automatically engages multi-sensory and motor processes, as well as studies demonstrating that planned or executed body movement may influence music perception (Maes et al., 2014). This evidence is typically cited in favor of the embodiment thesis given above (Maes et al., 2014; Matyja, 2015; Schiavio et al., 2015). However, as reliably argued by Mahon and Caramazza (2008), this conclusion cannot be fully drawn based on this evidence. The aim of this article is to collect further empirical evidence and computational models to substantiate the role of interaction in the embodiment thesis in the domain of music perception and performance. In a first part, we argue that (psycho)acoustic properties of sound, combined with processes within the human auditory and motor system fundamentally shape perception. On this basis, we propose an explanation of commonalities in musical repertoires across cultures, specifically related to tonality and tempo. Importantly, we will emphasize the necessity to consider the role of long-term knowledge formation through learning, and of related prediction and error-correction processes. More specifically, we consider acoustic properties, the human body, and the auditory system as natural dispositions on top of which learning and predictive processes may operate, leading to a differentiation in perception and musical repertoire. In a second part of this article, a similar approach was applied to music performance, in particular focusing on musical timing. We highlight two of our recent studies demonstrating the role of body dynamics and sensorimotor processes for timing in music performance. Again, we pinpoint how these processes may serve as a basis for learning processes facilitating prediction mechanisms.

In our approach to music perception and performance, we argue that insights from **dynamical system** theories may prove particularly useful. Previously, dynamical systems theories have been applied to the study of motor control (Turvey, 1990; Kelso, 1995; Thelen and Smith, 1998; Warren, 2006), and cognition (Port and Van Gelder, 1995; Van Gelder, 1998; Beer, 2000; McClelland et al., 2010; Buhrmann et al., 2013; Shapiro, 2013). This approach considers the various functions (sensory, motor, affective, cognitive, and social) engaged in people's interaction with music as intrinsically (non-linearly) interwoven and reciprocally deterministic. It involves a process-based approach focusing on the processes and constraints that propagate order and organization within all of the complexities, variability, and change inherent to humans' interaction with music (Ashby, 1962; Fischer and Bidell, 2007; Deacon, 2012).

KEY CONCEPT 1 | Dynamical system

A dynamical system can be defined by some key features:

[1] Assembly: A dynamical system is composed of multiple homogeneous and/or heterogeneous components.

[2] Self-organization: A dynamical system is a self-organizing system, in which order emerges out of the interactions of its components, without being explicitly controlled from within or without.

[3] Stability: Self-organization of a dynamical system is attracted toward stable relationships between its components (cf. error minimization).

[4] Constraint- and processes-based: Interactions, self-organization, and stability within a dynamical system are tied to the disposition of its components and the processes in which they are enmeshed, in interaction with the constraints and opportunities that the environment imposes/offers.

For that purpose, special emphasis will be given to (associative) learning, prediction and error-correction processes that underly music perception and performance, and link to core concepts of music research such as motivation, reward, affect, and agency.

2. MUSIC PERCEPTION

2.1. Dynamics and Natural Dispositions

2.1.1. Pitch Processing, Tonal Perception, and Music Syntactic Processing

Tonality refers to a musical system of tones (i.e., pitch classes) standing in a hierarchical relationship (Lerdahl, 2001). Tonality is characteristic of musical cultures worldwide (Western, Indian, Chinese, and Arabic), throughout history (Gill and Purves, 2009). The perceived stability of particular tones and chords (i.e., their tonal function) depends on the overarching musical context in which they occur. This context dependency has commonly been interpreted as evidence for the existence of abstract cognitive music-syntactic processing, rooted in learning processes, and thus musical enculturation (Krumhansl, 1990; Tillmann et al., 2000; Janata et al., 2002). Long-term exposure to the "statistics" of tonal relationships in tonal music lead to the development of cognitive schemata that capture tonal hierarchies, leading to specific expectancies of harmonic and melodic progress. It is plausible that the statistics of occurring tonal relationships in the musical repertoire drives the development of predictive models of tonal perception. However, the predictive model does not answer the crucial question *why* the co-occurrences got into the repertoire. In the following, some key studies are presented indicating that the perception of tonal hierarchy may be grounded in dynamical, sensory-based mechanisms that draw upon inherent physical and biological constraints (i.e., acoustics, and the human auditory system).

In articles by Bigand et al. (2014) and Collins et al. (2014), an important corpus of empirical studies on tonal perception was readdressed. These studies involve a series of behavioral studies on tonal perception applying the typical harmonic priming paradigm, instructing participants to judge whether a target chord was in-tune or out-of-tune with respect to a single prime chord or a longer harmonic context. Based on the reaction time of people's responses, tone profiles reflecting tonal hierarchies could be constructed. In addition, Bigand et al. (2014) readdressed a series of event-related potential (ERP) studies showing that violations of harmonic regularities are reflected in late positive components and early negative components. Bigand et al. (2014) proposed a sensory-based model to explain the observed data. The computational model used was the auditory short-term memory (ASTM) model introduced by Leman (2000). This ASTM model was used to simulate the afore-mentioned series of empirical studies in the field of tonal music perception. The obtained simulated ratings indicated that most of the behavioral and neurophysiological responses observed in these studies could be accounted for by the ASTM model. Considering the "nature" of this auditory short-time memory, neuroimaging studies have pinpointed the role of a dynamic interplay between attentional

and sensory systems (Kaiser, 2015), as well as sensorimotor processes in support of the temporary maintenance of auditory information in working memory (Schulze and Koelsch, 2012). In another study, by Collins et al. (2014), it was further argued that low-level sensory properties of musical sounds shape tonal hierarchy and expectation. Though, they concluded that musical syntax requires multiple representational stages along a sensory-to-cognition continuum. These results suggest that music-syntactic processing is, to a certain extent, a stimulus-driven process, rooted in the (psycho)acoustic properties of sound. Although different tonal systems and tuning systems are employed in different musical traditions, it is highly peculiar however that the most widely used scales typically consist of five or seven pitch classes. Additionally, research has shown that these scales have specific spectral characteristics, in the sense that the component intervals of these scales have greatest spectral similarity to a harmonic series (Gill and Purves, 2009). In the following, some studies are discussed that provide a further explanation for why these scales are preferred above others.

In a series of works, Sethares (1993, 2005) explained how the use of tuning systems and tone scales stems from the specific timbre of musical instruments being used. The timbre of tones produced by Western musical instruments is predominantly characterized by a fundamental frequency with a harmonic series of partials. Drawing on previous work by Plomp and Levelt (1965), Sethares calculated the combined dissonance among all partials of a tone, leading to a dissonance curve in which minima (i.e., local consonance) correspond with many Just Intonation scale steps. Although, not entirely undisputed (cf. Cousineau et al., 2012), sensory dissonance seems to be a property of the human auditory system related to the critical bandwidth of the basilar membrane. Hence, the use of certain scales across different cultures seems to be directly related to both the timbre of the musical instruments and the working of the human auditory system.

In another series of studies, Large and colleagues provide a neurophysiological explanation of why certain tone scales are preferred above others (Large, 2011a,b; Large and Almonte, 2012; Lerud et al., 2014). For that purpose, they introduced a computer model (i.e., gradient-frequency neural networks; GFNNs) of auditory pitch processing based on a non-linear, dynamical systems approach. Characteristic for the model is that the response frequencies of neural oscillators at different stages along the auditory pathway do not simply reflect the frequencies contained in the input stimulus; oscillations may occur at harmonics (e.g., 2:1 and 3:1), subharmonics (e.g., 1:2, 1:3), and more complex integer ratios (e.g., 3:2) of the stimulus frequency, and in the case of multi-frequency stimulation, at summation and difference frequencies (cf. Lee et al., 2009). According to Large and colleagues, this non-linear resonance behavior predicts the generalized preference for simple integer ratios, forming a biological basis of the most widely used tonal systems and tuning methods. Interestingly, it is further argued that internal connectivity between naturally coupled neural oscillators may be strengthened by Hebbian learning processes, providing an explanation why different tonal systems exist

among different musical traditions (cf. Chandrasekaran et al., 2013). In their work, Bidelman et al. (Bidelman, 2013; Bidelman and Grall, 2014), give further evidence for a neurobiological predisposition for consonant pitch relationships, basic to tonal hierarchies. Their research indicated that perceptual correlates of pitch hierarchy are automatically and pre-attentively mapped onto activation patterns within the subcortical auditory nervous system. Their work provides further evidence that the structural foundations of musical pitch perception inherently relate to innate processing mechanisms within the human auditory system.

This corpus of research suggests that the dynamics and (short-time) processes within the auditory apparatus play a fundamental role in music perception. In addition, it is shown that the acoustic environment (e.g., timbre of musical instruments) provides the fundament of most well-known tuning systems and tone scales. This focus on auditory and environmental conditions is valuable as it may explain well *why* specific tonal relationships got into the musical repertoire. This supports our central claim that the auditory apparatus and acoustic environment provide the genuine predispositions on top of which long-term learning processes may operate to differentiate further the musical repertoire. In that regard, we agree with Koelsch stating that "long-term knowledge about the relations of tones and chords is not a necessary condition for the establishment of a tonal hierarchy, but a sufficient condition for the modulation of such establishment (that is, the establishment of such a hierarchy is shaped by cultural experience)" (Koelsch, 2012, p.105).

2.1.2. Processing of Time and the Resonance Model of Tempo Perception

Next to tonality, another musical feature is characteristic of the repertoire of Western music, namely tempo. van Noorden and Moelants (1999) and Moelants (2002) analyzed the perceived tempo distribution of a huge sample of Western music. The results showed that the perceived tempi in Western music group around the 500 ms period (120 bpm or 2 Hz). Interestingly, studies in the domain of motor and auditory-motor control report related findings. For instance, when people are asked to make (unconstrained) cyclical movements with a finger or wrist at a comfortable regular tempo, it is shown that their preferred tempi peak slightly below 120 bpm (Kay et al., 1987; Collyer et al., 1994). Also, research indicated that the preferred step frequency in free-cadence walking during short and extended periods is about 2 Hz (120 steps per min; Murray et al., 1964; MacDougall and Moore, 2005). Additionally, it has been shown that auditory-motor synchronization is most accurate when the tempo of the external auditory stimulus is around 120 bpm (Styns et al., 2007; Leman et al., 2013). The observed correspondence between periods of cyclic (auditory-)motor behavior and preferred musical tempi suggests a link between the musical repertoire, music perception, and the human motor system. This idea is at the center of van Noorden's and Moelants' resonance model of time perception, which ties time perception in music to natural resonance frequencies in the human body. The model views the human body, as it is capable to perform periodic movements, as a

harmonic oscillator with a distinct natural resonance frequency. Resonance is the phenomenon whereby an external periodic force drives the amplitude of oscillations to a relative maximum. Resonance occurs when the frequency of the driving force is close to the natural frequency of the oscillator. In the model of van Noorden and Moelants, the external periodic force is the musical beat that drives corresponding periodic movement responses. The model suggests that the distribution of preferred tempi in the Western repertoire is matched to the natural predisposition of the human motor system to move comfortably and spontaneously at a pulse rate around 2 Hz. Again, similar to our discussion of tonality, this work strongly suggests that the "statistics" of the musical repertoire, and the preference and perception of tempo (i.e., whether a musical piece is slow or fast) are rooted in physical and biological dispositions inherent to the sensory and motor systems.

2.2. Predictive Processing

Musical input is often noisy or ambiguous making it difficult to reliably discern basic auditory features such as pitch, duration, and dynamics. In cognitive science, it becomes common to refer to Bayesian statistical inference to explain perception under such "uncertain" conditions. The basic idea is that people constantly make predictions of ensuing sensory events, based on current sensory input and learned sensorimotor regularities in our environment. Based on Bayesian principles, these predictions may lead to more accurate (optimal) perceptions under noisy and ambiguous conditions.

This predictive framework may further refine our understanding of the action-based effects that were reviewed in our previous article, in particular disambiguation effects (Maes et al., 2014). We have explained how repeated experience of auditory-motor regularities lead to internal models of our interaction with the world. As explained, the forward component of these models allows predicting the auditory states or phenomena that go along with performed body movements. Hence, we hypothesize here that action-induced expectancies may promote a selective response to incoming auditory information during music listening causing people to "sample" the auditory features in an "optimal" way so that expectations are confirmed, and prediction errors avoided (Clark, 2015). In other words, auditory and musical features that confirm "prior beliefs" are given priority and enter into perception, leading to the observed disambiguation effects. This explanation considers perception as a process of active engagement, in which the interaction between auditory stimulation, sensorimotor predictions, and attention play a central role (Schröger et al., 2015).

In addition to action-based disambiguation effects on music perception, predictive processing based on Bayesian inference has been applied to general perception phenomena (Temperley, 2007), and more specifically to auditory scene analysis (Winkler and Schröger, 2015; note also early inferentialist approaches by Bregman, 2008 and Levitin, 2006, rhythm, Sadakata et al., 2006; Lange, 2013; Vuust and Witek, 2014, and duration Cicchini et al., 2012; Aagten-Murphy et al., 2014).

Although the concept of the "**Bayesian brain**" is promising, we should be aware that many issues are left to be resolved (Eberhardt and Danks, 2011; Jones and Love, 2011; Bowers and Davis, 2012; Marcus and Davis, 2013; Orlandi, 2014b). Two of the most prominent questions relate to how the "hypothesis space" is restricted, and to whether the Bayesian brain deploys explicit internal representations of the rules that underlie prediction. At the present, typical Bayesian inferential accounts do not offer a solution to these questions. Recent accounts may offer further clarification by approaching Bayesian perception from an embedded perspective rooted in natural scene statistics and ecology (Orlandi, 2014a,b; Judge, 2015). The debate between inferentialism and ecology is outside the scope if this article. What is important however is to acknowledge that human prediction in both accounts is closely linked to our sensitivity to statistical regularities in our environment. Then the question remains why certain regularities and rules got into the musical repertoire at the expense of others? Therefore, we reviewed empirical evidence in the previous section showing that the dynamics and natural dispositions of our acoustic environment, human body, and auditory system constrains specific regularities. Then, on top of that, cultural- and history-specific habits and preferences may further differentiate the musical repertoires.

KEY CONCEPT 2 | The Bayesian brain

In its essence, the theory of the Bayesian brain posits that the human brain is compelled to infer the probable causes of its sensations, and to predict future states of the world. This ability relies on the existence of learned (generative/internal) models of the world, which include prior knowledge about sensorimotor regularities (cf. the concept of internal models discussed in Maes et al., 2014). This idea can be traced back to von Helmholtz' understanding of perception as inference (von Helmholtz, 1962). Nowadays, this idea is reflected in the theory of *predictive coding*, and the related *Bayesian coding hypothesis*, which stand out as most dominant accounts in psychology and neuroscience to explain the brain's functions, ranging from sensory perception to high-level cognition (Rao and Ballard, 1999; Knill and Pouget, 2004; Friston and Kiebel, 2009; Clark, 2013; Pouget et al., 2013; Summerfield and de Lange, 2014). In that regard, perception is a process of probabilistic inference, whereby perception is viewed as the "compromise" between sensations and prior knowledge, made in an attempt to minimize the discrepancy between both (cf. *free energy principle*, Friston, 2010). This can be realized in various ways: by optimizing the reliability and precision of sensory input through attention processes (Rao, 2005; Feldman and Friston, 2010; Brown et al., 2011; Kaya and Elhilali, 2014), by introducing perceptual bias (Geisler and Kersten, 2002), or by updating one's prior beliefs.

---

3. MUSIC PERFORMANCE

In this section, we present two recent empirical studies in support of the central idea of this article, namely that—similar to music perception—music performance may heavily rely on the inherent dynamics of the human motor system, in combination with predictive processing allowing online adaptation to changing environments. Thereby, we focus on temporal control of performers' body movements. For musicians, it is important to temporally coordinate muscle activity in order to control their musical instrument, or vocal chords in the case of singers. Traditional accounts posit the existence of a cognitively controlled internal clock mechanism to keep track of time,

indicating when muscles need to be activated (Church, 1984; Allman et al., 2014). However, as human cognitive resources are fairly limited, this account is probably incomplete to fully explain temporal behavior. Accumulating evidence is gathered demonstrating that the human sensorimotor system may support time perception and production, in addition to cognitively-based mechanisms (Hopson, 2003; Mauk and Buonomano, 2004; Ross and Balasubramaniam, 2014). In a series of experiments, we defined and tested two hypotheses to provide more insights into the factors and mechanisms regulating sensorimotor timing strategies in music performance; limited to solo performance, and regular tone production tasks. First, based on theories on **emergent timing**, it was hypothesized that the dynamic control of body movements might lead to temporal regularities with a minimum of explicit cognitive control. Second, it was hypothesized that actions might be aligned to perceived temporal patterns in self-generated auditory feedback, leading to corresponding regular motor patterns.

KEY CONCEPT 3 | Event timing vs. emergent timing

Research on motor control and coordination suggests that the timing of rhythmic body movements is a hybrid phenomenon. Typically a distinction is made between discrete and continuous (quasi-periodic) movements (Robertson et al., 1999; Zelaznik et al., 2002; Delignières et al., 2004; Larue, 2005; Zelaznik et al., 2008; Torre and Balasubramaniam, 2009; Studenka et al., 2012; Janzen et al., 2014). Whereas discrete rhythmic movements are characterized by salient events separated by pauses in bodily movement, continuous rhythmic movements are smooth without interspersed pauses. Importantly, research suggests that these movement types rely on different control mechanisms. Discrete movements are regulated by an "event-based timing system." Here, the basic idea is that timing is explicitly controlled by a dedicated clock capable of keeping track of time. One of the most influential accounts of event-based timing is the pacemaker-accumulator model (Gibbon, 1977). In contrast, continuous movements are regulated by an "emergent timing system," which pertains to a dynamical systems perspective on motor control. According to this perspective, coordinated (regular) body movements are to a high extent the result of the motor system's dynamics with a minimum of explicit, central control (Turvey, 1977; Thelen, 1991; Kelso, 1995; Warren, 2006).
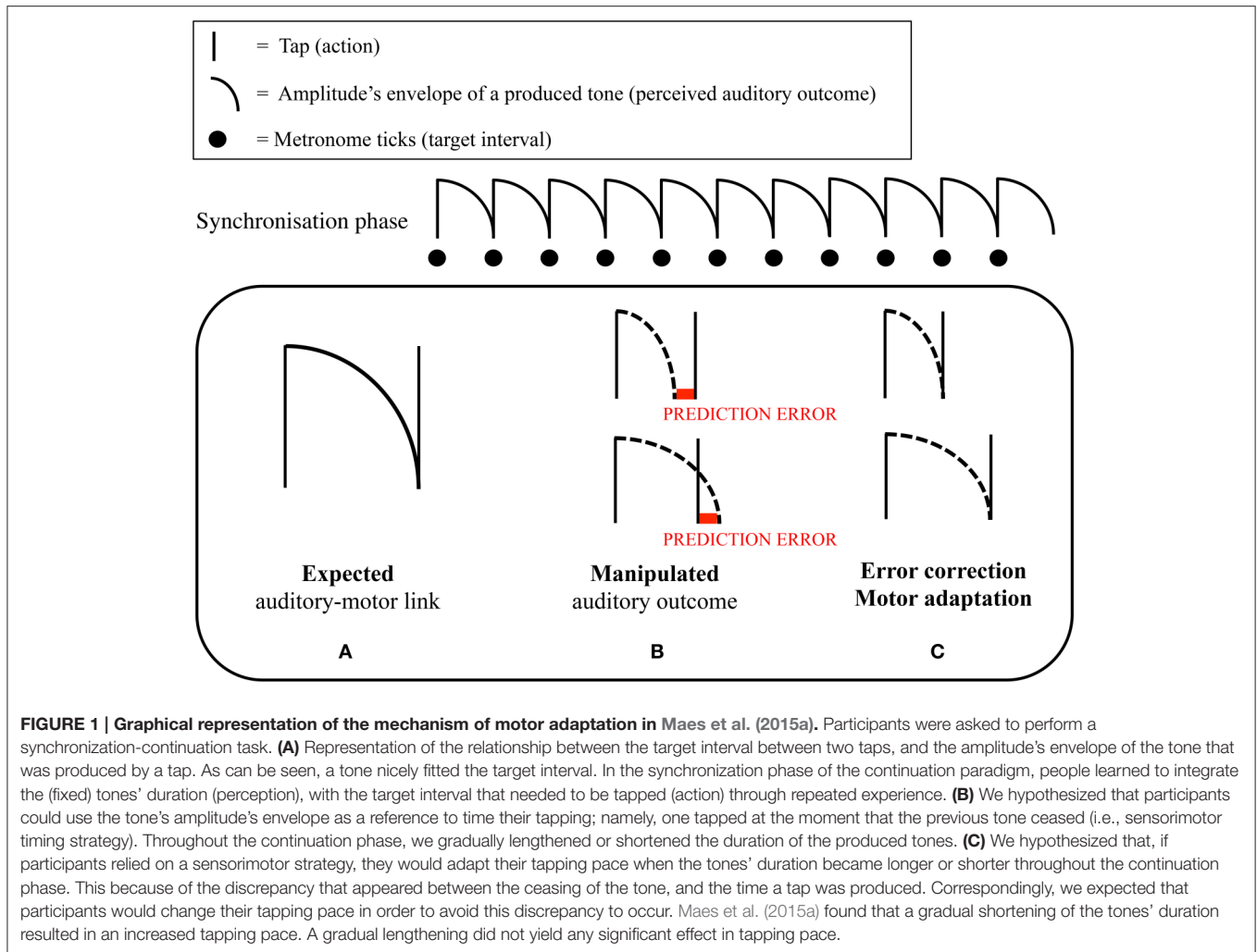

In the experiments, we asked participants—musical novices (Maes et al., 2015a) and professional cellists (Maes et al., 2015b)—to perform melodies consisting of equally spaced notes at a specific target tempo (synchronization-continuation task) while performing an additional cognitive task (cf. dual-task interference paradigm). The main idea was that the production of regular intervals would be relatively unharmed by an additional cognitive load when participants applied a sensorimotor-based timing strategy. The results of these experiments showed that when continuous arm movements could be applied in between tone onsets, production of regular intervals in the continuation task was not affected by an additional cognitive task, suggesting the use of a sensorimotor timing strategy. This in contrast to conditions where no movement was allowed in between tone onsets, leading to a significant increase in variability of produced temporal intervals, suggesting the use of a cognitively-controlled timing strategy. Additionally, participants were generally less accurate—i.e., further apart from the target tempo—when no arm movements were allowed. In this context, Maes et al. (2015a) investigated the role of self-generated auditory feedback. It was

found that participants—in particular when no additional load was present—were better able to keep the target tempo in the continuation phase when key taps produced tones that filled the complete duration of the interval, compared to when tones were short. In another part of the experiment, long tones filling the complete duration of the interval were made gradually shorter or longer throughout the continuation phase. Interestingly, it was found that, when tones were made shorter, participants speeded up their tapping tempo accordingly (i.e., intervals between produced onsets became gradually shorter; see **Figure 1**).

These results indicate that timing in music performance may capitalize directly on the control of movement dynamics and coupled action-perception processes, without the need to explicitly (cognitively) compute time. In other words, our findings showed that musical goals, here regular interval production, may be outsourced to the human sensorimotor system in interaction with (auditory) information accessible in the environment (Clark and Chalmers, 1998; Tylén and McGraw, 2014). Given the inherent limitations of cognitive resources, this concept of *outsourcing*—which shows resemblance with the concept of *auditory scaffolding/latching* (DeNora, 2000; Conway et al., 2009)—enables humans to optimally perform specific tasks in specific contexts, depending on their specific capabilities, state, and intentions. Nonetheless we argue that in order to be effective, sensorimotor timing strategies need to necessarily interact with processes involving associative learning, **prediction**, and **error-correction**. Basically, in synchronizing actions and tones to an external auditory metronome, associative learning facilitates to integrate movement dynamics and dynamic change in auditory information with the temporal intervals to be produced, leading to the development of so-called **internal models**. These enable then to keep the tempo in absence of the external metronome. Characteristic of people is to make predictions about the sensory consequences of planned actions. For instance in the context of our study (Maes et al., 2015a) people expected the onset of a tone to coalesce with the ending of the previous tone. Now, due to the dynamics and uncertainty inherent to internal and external conditions, discrepancies may occur between the expected sensory outcome of actions and the actual outcome. Typical for humans is to correct for occurring discrepancies by spontaneously adapting one's actions or perception, potentially leading to an update of the internal model. In Maes et al. (2015a), this was evidenced by the observation that people made temporal intervals shorter in response to a gradual shortening of the tones

---

KEY CONCEPT 4 | Internal model, prediction, error-correction, and motor adaptation

The temporal integration of actions and their sensory outcome—acquired through systematically repeated sensorimotor experiences—establishes what is typically referred to as an *internal model* (Maes et al., 2014). Internal models contain an *inverse* and *forward* component. Forward models allow predicting the likely sensory outcome of a planned or executed action. A distinct property of forward models is that they allow transforming discrepancies between the *expected* and the *actual* sensory outcome of a performed action into an error signal, which drives changes in motor output (i.e., motor adaptation) in order to reduce sensory prediction errors (Jordan and Rumelhart, 1992; Wolpert et al., 1995; Friston et al., 2006; Lalazar and Vaadia, 2008; Norwich, 2010; Shadmehr et al., 2010; van der Steen and Keller, 2013).

FIGURE 1 | Graphical representation of the mechanism of motor adaptation in Maes et al. (2015a)**.** Participants were asked to perform a synchronization-continuation task. **(A)** Representation of the relationship between the target interval between two taps, and the amplitude's envelope of the tone that was produced by a tap. As can be seen, a tone nicely fitted the target interval. In the synchronization phase of the continuation paradigm, people learned to integrate the (fixed) tones' duration (perception), with the target interval that needed to be tapped (action) through repeated experience. **(B)** We hypothesized that participants could use the tone's amplitude's envelope as a reference to time their tapping; namely, one tapped at the moment that the previous tone ceased (i.e., sensorimotor timing strategy). Throughout the continuation phase, we gradually lengthened or shortened the duration of the produced tones. **(C)** We hypothesized that, if participants relied on a sensorimotor strategy, they would adapt their tapping pace when the tones' duration became longer or shorter throughout the continuation phase. This because of the discrepancy that appeared between the ceasing of the tone, and the time a tap was produced. Correspondingly, we expected that participants would change their tapping pace in order to avoid this discrepancy to occur. Maes et al. (2015a) found that a gradual shortening of the tones' duration resulted in an increased tapping pace. A gradual lengthening did not yield any significant effect in tapping pace.

they produced in the continuation phase. This finding illustrates a fundamental mechanism and powerful strategy to adapt and guide temporal behavior toward specific goals. On top of that, it may evenly contribute to practical applications in the field of sports and motor rehabilitation, where strategies to adapt people's movement behavior instantaneously—and potentially unconsciously—are of high relevance (Moens and Leman, 2015).

4. DISCUSSION

So far, the most frequently cited empirical evidence for the embodiment thesis has been grounded in the observation that music perception automatically engages multi-sensory and motor simulation processes (Schiavio et al., 2015), or that bodily states and movement may influence music perception (Maes et al., 2014). In this article, we advocated for more rigorous evidence to substantiate the "radical" embodiment thesis in the domain of music perception and performance (Mahon and Caramazza, 2008; Chemero, 2011; Wilson and Golonka, 2013; Kiverstein and Miller, 2015). Therefore, we provided a focused

review presenting empirical evidence and computer models demonstrating that music perception and performance may be directly determined by the acoustics of sound and by the natural disposition and dynamics of the human sensory and motor system. On top of that, we have emphasized the role of long-term processes involving learning and prediction in how humans interact with music. At the present, the exact nature of these processes is still a matter of ongoing debate—boldly between inferential and ecological accounts (Orlandi, 2012, 2014a)—yet to be fully determined. However, the collected findings suggest to consider short-term modality-specific processes serving perceptual or sensorimotor functions, and long-term learning and prediction processes as reciprocally determined and interacting; sensorimotor experience may lead to predictions, and predictions may shape sensorimotor engagement with our environment.

In the future, it would be of interest to further extent this sensorimotor-prediction loop with aspects that relate to music expression, emotion, motivation, and social interaction. For instance, the work of Bowling et al. (2010, 2012) indicate that the musical expression and perception of happiness and

sadness has a biological basis in speech. They found that the acoustic frequency spectra of major and minor tone collections, linked to, respectively, happy and sad music, correspond to the frequency spectra found in, respectively, excited and subdued speech. This finding is of particular interest as it provides an explanation why the perception of musical emotion is shared across cultures (Fritz et al., 2009). Further, it is of interest to link musical expression and the experience of affect to neurodynamical processes (Seth, 2013; Flaig and Large, 2014). In addition to musical expression, the study of the sensorimotor-prediction loop can provide deeper insights into aspects of motivation and reward in music. Previous research demonstrated that dopaminergic activity, which relate to feelings of reward, encodes learning prediction errors (Waelti et al., 2001; Schultz, 2007; Hazy et al., 2010). This link of prediction processes to actual physiological responses, in this case dopamine responses, may contribute to our understanding of feelings of pleasure and reward that arise in music-based interactions (Chanda and Levitin, 2013; Zatorre and Salimpoor, 2013). Finally, all of the processes that occur within an individual in interaction with its sensory environment, may be dynamically linked to its social environment, leading to phenomena such as interpersonal coordination and synchronization (Repp and Su, 2013; Moran, 2014; D'Ausilio et al., 2015).

It is important to note that a dynamical, process-based approach to how humans interact with music requires a severe reconsideration of currently dominating analytical tools and methods (behavioral and neurophysiological) that are often reductionistic and focus on generalizations at the expense of variation and change. Therefore, it would be of interest to systematically incorporate methods from within the field of dynamical structure analysis to take into account time-dependent changes, variability, and non-linear complexities (Amelynck et al., 2014; Badino et al., 2014; Demos et al., 2014; Teixeira et al., 2015). A dynamical, processed-based approach, together with appropriate analytical tools may contribute profoundly to our understanding of music, and more importantly, to how and why people interact with music. In turn, this knowledge may be capitalized on by more practical research in various domains, such as sports (Karageorghis and Priest, 2012), motor rehabilitation (Altenmüller et al., 2009; Särkämö and Soto, 2012), developmental disorders (Koelsch, 2009), and well-being (MacDonald et al., 2012).

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

ACKNOWLEDGMENTS

REFERENCES

Aagten-Murphy, D., Cappagli, G., and Burr, D. (2014). Musical training generalises across modalities and reveals efficient and adaptive mechanisms for reproducing temporal intervals. *Acta Psychol.* 147, 25–33. doi: 10.1016/j.actpsy.2013.10.007

Allman, M. J., Teki, S., Griffiths, T. D., and Meck, W. H. (2014). Properties of the internal clock: first-and second-order principles of subjective time. *Annu. Rev. Psychol.* 65, 743–771. doi: 10.1146/annurev-psych-010213-115117

Altenmüller, E., Marco-Pallares, J., Münte, T. F., 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

Amelynck, D., Maes, P. J., Martens, J. P., and Leman, M. (2014). Expressive body movement responses to music are coherent, consistent, and low dimensional. *IEEE Trans. Cybern.* 44, 2288–2301. doi: 10.1109/TCYB.2014.2305998

Ashby, W. R. (1962). "Principles of the self-organising dynamic system," in *Principles of Self-Organization: Transactions of the University of Illinois Symposium*, eds H. Von Foerster and G. W. Zopf (London: Pergamon Press), 255–278.

Badino, L., D'Ausilio, A., Glowinski, D., Camurri, A., and Fadiga, L. (2014). Sensorimotor communication in professional quartets. *Neuropsychologia* 55, 98–104. doi: 10.1016/j.neuropsychologia.2013.11.012

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

Bidelman, G. M. (2013). The role of the auditory brainstem in processing musically relevant pitch. *Front. Psychol.* 4:264. doi: 10.3389/fpsyg.2013.00264

Bidelman, G. M., and Grall, J. (2014). Functional organization for musical consonance and tonal pitch hierarchy in human auditory cortex. *NeuroImage* 101, 204–214. doi: 10.1016/j.neuroimage.2014.07.005

Bigand, E., Delbé, C., Poulin-Charronnat, B., Leman, M., and Tillmann, B. (2014). Empirical evidence for musical syntax processing? Computer simulations reveal the contribution of auditory short-term memory. *Front. Syst. Neurosci.* 8:94. doi: 10.3389/fnsys.2014.00094

Bowers, J. S., and Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychol. Bull.* 138, 389–414. doi: 10.1037/a0026450

Bowling, D. L., Gill, K., Choi, J. D., Prinz, J., and Purves, D. (2010). Major and minor music compared to excited and subdued speech. *J. Acoust. Soc. Am.* 127, 491–503. doi: 10.1121/1.3268504

Bowling, D. L., Sundararajan, J., Han, S., and Purves, D. (2012). Expression of emotion in Eastern and Western music mirrors vocalization. *PLoS ONE* 7:e31942. doi: 10.1371/journal.pone.0031942

Bregman, A. (2008). "Auditory scene analysis," in *The Senses: A Comprehensive Reference*, eds A. I. Basbaum, A. Kaneko, G. M. Shepherd, and G. Westheimer (San Diego, CA: Academic Press), 861–870.

Brown, H., Friston, K., and Bestmann, S. (2011). Active inference, attention, and motor preparation. *Front. Psychol.* 2:218. doi: 10.3389/fpsyg.2011.00218

Buhrmann, T., Di Paolo, E. A., and Barandiaran, X. (2013). A dynamical systems account of sensorimotor contingencies. *Front. Psychol.* 4:285. doi: 10.3389/fpsyg.2013.00285

Chanda, M. L., and Levitin, D. J. (2013). The neurochemistry of music. *Trends Cogn. Sci.* 17, 179–193. doi: 10.1016/j.tics.2013.02.007

Chandrasekaran, B., Skoe, E., and Kraus, N. (2013). An integrative model of subcortical auditory plasticity. *Brain Topogr.* 27, 539–552. doi: 10.1007/s10548-013-0323-9

Chemero, A. (2011). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.

Church, R. M. (1984). Properties of the internal clock. *Ann. N.Y. Acad. Sci.* 423, 566–582. doi: 10.1111/j.1749-6632.1984.tb23459.x

Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M., and Burr, D. C. (2012). Optimal encoding of interval timing in expert percussionists. *J. Neurosci.* 32, 1056–1060. doi: 10.1523/JNEUROSCI.3411-11.2012

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/S0140525X12000477

Clark, A. (2015). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind.* Oxford, UK: Oxford University Press.

Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis* 58, 7–19. doi: 10.1093/analys/58.1.7

Cochrane, T. (2010). Music, emotions and the influence of the cognitive sciences. *Philos. Compass* 5, 978–988. doi: 10.1111/j.1747-9991.2010.00337.x

Collins, T., Tillmann, B., Barrett, F. S., Delbé, C., and Janata, P. (2014). A combined model of sensory and cognitive representations underlying tonal expectations in music: from audio signals to behavior. *Psychol. Rev.* 121, 33–65. doi: 10.1037/a0034695

Collyer, C. E., Broadbent, H. A., and Church, R. M. (1994). Preferred rates of repetitive tapping and categorical time production. *Percept. Psychophys.* 55, 443–453. doi: 10.3758/BF03205301

Conway, C. M., Pisoni, D. B., and Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities the auditory scaffolding hypothesis. *Curr. Dir. Psychol. Sci.* 18, 275–279. doi: 10.1111/j.1467-8721.2009. 01651.x

Cousineau, M., McDermott, J. H., and Peretz, I. (2012). The basis of musical consonance as revealed by congenital amusia. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19858–19863. doi: 10.1073/pnas.1207989109

D'Ausilio, A., Novembre, G., Fadiga, L., and Keller, P. E. (2015). What can music tell us about social interaction? *Trends Cogn. Sci.* 19, 111–114. doi: 10.1016/j.tics.2015.01.005

Deacon, T. W. (2012). *Incomplete Nature: How Mind Emerged from Matter.* New York, NY: W. W. Norton & Company.

Delignières, D., Lemoine, L., and Torre, K. (2004). Time intervals production in tapping and oscillatory motion. *Hum. Mov. Sci.* 23, 87–103. doi: 10.1016/j.humov.2004.07.001

Demos, A. P., Chaffin, R., and Kant, V. (2014). Toward a dynamical theory of body movement in musical performance. *Front. Psychol.* 5:477. doi: 10.3389/fpsyg.2014.00477

DeNora, T. (2000). *Music in Everyday Life.* Cambridge, MA: Cambridge University Press.

Eberhardt, F., and Danks, D. (2011). Confirmation in the cognitive sciences: the problematic case of Bayesian models. *Minds Mach.* 21, 389–410. doi: 10.1007/s11023-011-9241-3

Feldman, H., and Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* 4:215. doi: 10.3389/fnhum.2010.00215

Fischer, K. W., and Bidell, T. R. (2007). "Dynamic development of action and thought," in *Theoretical Models of Human Development. Handbook of Child Psychology 6th Edn.,* Vol. 1, eds W. Damon and R. M. Lerner (New York, NY: Wiley Online Library), 313–399.

Flaig, N. K., and Large, E. W. (2014). Dynamic musical communication of core affect. *Front. Psychol.* 5:72. doi: 10.3389/fpsyg.2014.00072

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Friston, K., and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1211–1221. doi: 10.1098/rstb.2008.0300

Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *J. Physiol.* 100, 70–87. doi: 10.1016/j.jphysparis.2006.10.001

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., et al. (2009). Universal recognition of three basic emotions in music. *Curr. Biol.* 19, 573–576. doi: 10.1016/j.cub.2009.02.058

Geisler, W. S., and Kersten, D. (2002). Illusions, perception and Bayes. *Nat. Neurosci.* 5, 508–510. doi: 10.1038/nn0602-508

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* 84, 279–325. doi: 10.1037/0033-295X.84.3.279

Gill, K. Z., and Purves, D. (2009). A biological rationale for musical scales. *PLoS ONE* 4:e8144. doi: 10.1371/journal.pone.0008144

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

Hazy, T. E., Frank, M. J., and O'Reilly, R. C. (2010). Neural mechanisms of acquired phasic dopamine responses in learning. *Neurosci. Biobehav. Rev.* 34, 701–720. doi: 10.1016/j.neubiorev.2009

Hopson, J. W. (2003). "General learning models: timing without a clock," in *Functional and Neural Mechanisms of Interval Timing,* ed W. H. Meck (Boca Raton, FL: CRC Press), 23–60.

Janata, P., Birk, J. L., Van Horn, J. D., Leman, M., Tillmann, B., and Bharucha, J. J. (2002). The cortical topography of tonal structures underlying western music. *Science* 298, 2167–2170. doi: 10.1126/science.1076262

Janzen, T. B., Thompson, W. F., Ammirante, P., and Ranvaud, R. (2014). Timing skills and expertise: discrete and continuous timed movements among musicians and athletes. *Front. Psychol.* 5:1482. doi: 10.3389/fpsyg.2014.01482

Jones, M., and Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behav. Brain Sci.* 34, 169–188. doi: 10.1017/S0140525X10003134

Jordan, M. I., and Rumelhart, D. E. (1992). Forward models: supervised learning with a distal teacher. *Cogn. Sci.* 16, 307–354. doi: 10.1207/s15516709cog1603_1

Judge, J. (2015). Does the 'missing fundamental' require an inferentialist explanation? *Topoi* 1–11. doi: 10.1007/s11245-014-9298-8. Available online at: http://link.springer.com/article/10.1007%2Fs11245-014-9298-8#/page-1

Kaiser, J. (2015). Dynamics of auditory working memory. *Front. Psychol.* 6:613. doi: 10.3389/fpsyg.2015.00613

Karageorghis, C. I., and Priest, D.-L. (2012). Music in the exercise domain: a review and synthesis, Part I. *Int. Rev. Sport Exerc. Psychol.* 5, 44–66. doi: 10.1080/1750984X.2011.631026

Kay, B. A., Kelso, J. A., Saltzman, E. L., and Schöner, G. (1987). Space–time behavior of single and bimanual rhythmical movements: data and limit cycle model. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 178–192. doi: 10.1037/0096-1523.13.2.178

Kaya, E. M., and Elhilali, M. (2014). Investigating bottom-up auditory attention. *Front. Hum. Neurosci.* 8:327. doi: 10.3389/fnhum.2014.00327

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

Kiverstein, J., and Miller, M. (2015). The embodied brain: towards a radical embodied cognitive neuroscience. *Front. Hum. Neurosci.* 9:237. doi: 10.3389/fnhum.2015.00237

Knill, D. C., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719. doi: 10.1016/j.tins.2004.10.007

Koelsch, S. (2009). A neuroscientific perspective on music therapy. *Ann. N.Y. Acad. Sci.* 1169, 374–384. doi: 10.1111/j.1749-6632.2009.04592.x

Koelsch, S. (2012). *Brain and Music.* Chichester, UK: Wiley & Sons.

Krueger, J. (2013). Affordances and the musically extended mind. *Front. Psychol.* 4:1003. doi: 10.3389/fpsyg.2013.01003

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch.* New York, NY: Oxford University Press.

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

Lange, K. (2013). The ups and downs of temporal orienting: a review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory n1 with opposite effects of attention and prediction. *Front. Hum. Neurosci.* 7:263. doi: 10.3389/fnhum.2013.00263

Large, E. W. (2011a). "A dynamical systems approach to musical tonality," in *Nonlinear Dynamics in Human Behavior, Vol. 328, Studies in Computational Intelligence,* eds R. Huys and V. L. Jirsa (Berlin; Heidelberg: Springer), 193–211.

Large, E. W. (2011b). "Musical tonality, neural resonance and hebbian learning," in *Mathematics and Computation in Music, Vol. 6726, Lecture Notes in Computer Science (LNCS),* eds C. Agon, E. Amiot, M. Andreatta, G. Assayag, J. Bresson, and J. Manderau (Berlin; Heidelberg: Springer), 115–125.

Large, E. W., and Almonte, F. V. (2012). Neurodynamics, tonality, and the auditory brainstem response. *Ann. N.Y. Acad. Sci.* 1252, E1–E7. doi: 10.1111/j.1749-6632.2012.06594.x

Larue, J. (2005). Initial learning of timing in combined serial movements and a no-movement situation. *Music Percept.* 22, 509–530. doi: 10.1525/mp.2005.22.3.509

Lee, K. M., Skoe, E., Kraus, N., and Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *J. Neurosci.* 29, 5832–5840. doi: 10.1523/JNEUROSCI.6133-08.2009

Leman, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Percept.* 17, 481–509. doi: 10.2307/40285830

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

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

Lerdahl, F. (2001). *Tonal Pitch Space*. New York, NY: Oxford University Press.

Lerud, K. D., Almonte, F. V., Kim, J. C., and Large, E. W. (2014). Mode-locking neurodynamics predict human auditory brainstem responses to musical intervals. *Hear. Res.* 308, 41–49. doi: 10.1016/j.heares.2013.09.010

Levitin, D. J. (2006). *This is Your Brain on Music: The Science of a Human Obsession*. New York: NY: Dutton Books.

MacDonald, R., Kreutz, G., and Mitchell, L. (2012). *Music, Health, and Wellbeing*. Oxford, UK: Oxford University Press.

MacDougall, H. G., and Moore, S. T. (2005). Marching to the beat of the same drummer: the spontaneous tempo of human locomotion. *J. Appl. Physiol.* 99, 1164–1173. doi: 10.1152/japplphysiol.00138.2005

Maes, P.-J., Giacofci, M., and Leman, M. (2015a). Auditory and motor contributions to the timing of melodies under cognitive load. *J. Exp. Psychol. Hum. Percept. Perform.* 41, 1336–1352. doi: 10.1037/xhp0000085

Maes, P.-J., Leman, M., Palmer, C., and Wanderley, M. M. (2014). Action-based effects on music perception. *Front. Psychol.* 4:1008. doi: 10.3389/fpsyg.2013.01008

Maes, P.-J., Wanderley, M. M., and Palmer, C. (2015b). The role of working memory in the temporal control of discrete and continuous movements. *Exp. Brain Res.* 233, 263–273. doi: 10.1007/s00221-014-4108-5

Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol.* 102, 59–70. doi: 10.1016/j.jphysparis.2008.03.004

Marcus, G. F., and Davis, E. (2013). How robust are probabilistic models of higher-level cognition? *Psychol. Sci.* 24, 2351–2360. doi: 10.1177/0956797613495418

Matyja, J. R. (2015). The next step: mirror neurons, music, and mechanistic explanation. *Front. Psychol.* 6:409. doi: 10.3389/fpsyg.2015.00409

Mauk, M. D., and Buonomano, D. V. (2004). The neural basis of temporal processing. *Annu. Rev. Neurosci.* 27, 307–340. doi: 10.1146/annurev.neuro.27.070203.144247

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

Moelants, D. (2002). "Preferred tempo reconsidered," in *Proceedings of the 7th International Conference on Music Perception and Cognition (ICMPC)*, eds C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick (Adelaide, SA: Causal Productions), 580–583.

Moens, B., and Leman, M. (2015). Alignment strategies for the entrainment of music and movement rhythms. *Ann. N.Y. Acad. Sci.* 1337, 86–93. doi: 10.1111/nyas.12647

Moran, N. (2014). Social implications arise in embodied music cognition research which can counter musicological "individualism." *Front. Psychol.* 5: 676. doi: 10.3389/fpsyg.2014.00676

Murray, M. P., Drought, A. B., and Kory, R. C. (1964). Walking patterns of normal men. *J. Bone Joint Surg.* 46, 335–360.

Norwich, K. H. (2010). Le Chatelier's principle in sensation and perception: fractal-like enfolding at different scales. *Front. Physiol.* 1:17. doi: 10.3389/fphys.2010.00017

Orlandi, N. (2012). Embedded seeing-as: multi-stable visual perception without interpretation. *Philos. Psychol.* 25, 555–573. doi: 10.1080/09515089.2011.579425

Orlandi, N. (2014a). Bayesian perception is ecological perception. *Minds Online*. Available online at: http://mindsonline.philosophyofbrains.com/author/norlandiucsc-edu/

Orlandi, N. (2014b). *The Innocent Eye: Why Vision is Not a Cognitive Process*. Oxford, UK: Oxford University Press.

Plomp, R., and Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *J. Acoust. Soc. Am.* 38, 548–560. doi: 10.1121/1.1909741

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

Pouget, A., Beck, J. M., Ma, W. J., and Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nat. Neurosci. Rev.* 16, 1170–1178. doi: 10.1038/nn.3495

Rao, R. P. N. (2005). Bayesian inference and attentional modulation in the visual cortex. *Neuroreport* 16, 1843–1848. doi: 10.1097/01.wnr.0000183900.92901.fc

Rao, R. P. N., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580

Repp, B. H., and Su, Y. H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon. Bull. Rev.* 20, 403–452. doi: 10.3758/s13423-012-0371-2

Reybrouck, M. (2005). A biosemiotic and ecological approach to music cognition: event perception between auditory listening and cognitive economy. *Axiomathes* 15, 229–266. doi: 10.1007/s10516-004-6679-4

Robertson, S. D., Zelaznik, H. N., Lantero, D. A., Bojczyk, K. G., Spencer, R. M., Doffin, J. G., et al. (1999). Correlations for timing consistency among tapping and drawing tasks: evidence against a single timing process for motor control. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 1316–1330. doi: 10.1037/0096-1523.25.5.1316

Ross, J. M., and Balasubramaniam, R. (2014). Physical and neural entrainment to rhythm: human sensorimotor coordination across tasks and effector systems. *Front. Hum. Neurosci.* 8:576. doi: 10.3389/fnhum.2014.00576

Sadakata, M., Desain, P., and Honing, H. (2006). The bayesian way to relate rhythm perception and production. *Music Percept.* 23, 269–288. doi: 10.1525/mp.2006.23.3.269

Särkämö, T., and Soto, D. (2012). Music listening after stroke: beneficial effects and potential neural mechanisms. *Ann. N.Y. Acad. Sci.* 1252, 266–281. doi: 10.1111/j.1749-6632.2011.06405.x

Schiavio, A., Menin, D., and Matyja, J. (2015). Music in the flesh: embodied simulation in musical understanding. *Psychomusicol. Music Mind Brain* 24, 340–343. doi: 10.1037/pmu0000052

Schröger, E., Marzecová, A., and SanMiguel, I. (2015). Attention and prediction in human audition: a lesson from cognitive psychophysiology. *Eur. J. Neurosci.* 41, 641–664. doi: 10.1111/ejn.12816

Schultz, W. (2007). Behavioral dopamine signals. *Trends Neurosci.* 30, 203–210. doi: 10.1016/j.tins.2007.03.007

Schulze, K., and Koelsch, S. (2012). Working memory for speech and music. *Ann. N.Y. Acad. Sci.* 1252, 229–236. doi: 10.1111/j.1749-6632.2012.06447.x

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* 17, 565–573. doi: 10.1016/j.tics.2013.09.007

Sethares, W. A. (1993). Local consonance and the relationship between timbre and scale. *J. Acoust. Soc. Am.* 94, 1218–1228. doi: 10.1121/1.408175

Sethares, W. A. (2005). *Tuning, Timbre, Spectrum, Scale*. Berlin; Heidelberg: Springer.

Shadmehr, R., Smith, M. A., and Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* 33, 89–108. doi: 10.1146/annurev-neuro-060909-153135

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

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.

Studenka, B. E., Zelaznik, H. N., and Balasubramaniam, R. (2012). The distinction between tapping and circle drawing with and without tactile feedback: an examination of the sources of timing variance. *Q. J. Exp. Psychol.* 65, 1086–1100. doi: 10.1080/17470218.2011.640404

Styns, F., van Noorden, L., Moelants, D., and Leman, M. (2007). Walking on music. *Hum. Mov. Sci.* 26, 769–785. doi: 10.1016/j.humov.2007.07.007

Summerfield, C., and de Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* 15, 745–756. doi: 10.1038/nrn3838

Teixeira, E. C. F., Yehia, H. C., and Loureiro, M. A. (2015). Relating movement recurrence and expressive timing patterns in music performances. *J. Acoust. Soc. Am.* 138, EL212–EL216. doi: 10.1121/1.4929621

Temperley, D. (2007). *Music and Probability*. Cambridge, MA: MIT Press.

Thelen, E. (1991). "Timing in motor development as emergent process and product," in *Advances in Psychology, Vol. 81, The Development of Timing*

Control and Temporal Organization in Coordinated Action, eds J. Fagard and P. H. Wolff (New York, NY: Elsevier), 201–211.

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 (Hoboken, NJ: Wiley Online Library), 258–312.

Tillmann, B., Bharucha, J. J., and Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. *Psychol. Rev.* 107, 885–913. doi: 10.1037/0033-295X.107.4.885

Torre, K., and Balasubramaniam, R. (2009). Two different processes for sensorimotor synchronization in continuous and discontinuous rhythmic movements. *Exp. Brain Res.* 199, 157–166. doi: 10.1007/s00221-009-1991-2

Turvey, M. T. (1977). "Preliminaries to a theory of action with reference to vision," in *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, eds L. Shaw and J. Bransford (Hillsdale, NJ: Lawrence Erlbaum Associates), 211–265.

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

Tylén, K., and McGraw, J. J. (2014). "Materializing mind: the role of objects in cognition and culture," in *Perspectives on Social Ontology and Social Cognition, Vol. 4, Studies in the Philosophy of Sociality*, eds M. Gallotti and J. Michael (Berlin; Heidelberg: Springer), 135–148.

van der Steen, M. C., and Keller, P. E. (2013). The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Front. Hum. Neurosci.* 7: 253. doi: 10.3389/fnhum.2013.00253

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behav. Brain Sci.* 21, 615–628. doi: 10.1017/S0140525X98001733

van Noorden, L., and Moelants, D. (1999). Resonance in the perception of musical pulse. *J. New Music Res.* 28, 43–66. doi: 10.1076/jnmr.28.1.43.3122

von Helmholtz, H. (1860/1962). *Handbuch der Physiologischen Optik, Vol. 3.* Transl. by J. P. C. Southall in English. New York, NY: Dover.

Vuust, P., and Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Front. Psychol.* 5:1111. doi: 10.3389/fpsyg.2014.01111

Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412, 43–48. doi: 10.1038/35083500

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

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

Winkler, I., and Schröger, E. (2015). Auditory perceptual objects as generative models: setting the stage for communication by sound. *Brain Lang.* 148, 1–22. doi: 10.1016/j.bandl.2015.05.003

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

Zatorre, R. J., and Salimpoor, V. N. (2013). From perception to pleasure: music and its neural substrates. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10430–10437. doi: 10.1073/pnas.1301228110

Zelaznik, H. N., Spencer, R. M., and Ivry, R. B. (2002). Dissociation of explicit and implicit timing in repetitive tapping and drawing movements. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 575–588. doi: 10.1037/0096-1523.28.3.575

Zelaznik, H. N., Spencer, R. M. C., and Ivry, R. B. (2008). "Behavioral analysis of human movement timing," in *Psychology of Time*, ed S. Grondin (Bingley: Emerald Group Publishing Limited), 233–260.