Music in the Brain

Peter Vuust, Ole A. Heggli, Karl J. Friston & Morten L. Kringelbach

Abstract

Music is ubiquitous across human cultures—as a source of affective and pleasurable experience, moving us both physically and emotionally—and learning to play music shapes both brain structure and function. Music processing in the brain—namely, perception of melody, harmony, and rhythm—has traditionally been studied as an auditory phenomenon using passive listening paradigms. However, when listening to music, we actively generate predictions about what is likely to happen next. This enactive aspect has led to a more complete understanding of music processing involving brain structures implicated in action, emotion, and learning. Here, we review the cognitive neuroscience literature of music perception. We show that music perception, action, emotion, and learning all rest on the human brain's fundamental capacity for prediction—as formulated by the predictive coding of music model. This review elucidates how this formulation of music perception and expertise in individuals can be extended to account for the dynamics and underlying brain mechanisms of collective music making. This in turn has important implications for human creativity as evinced by music improvisation. These recent advances shed new light on what makes music meaningful from a neuroscientific perspective.

Introduction

"Listening is the key to everything great in music."

(Pat Metheny)

Even though music is often described as no more than sounds—or soundscapes—organized intentionally by a composer or performer, it feels meaningful and emotional to most people. From the point of view of music theory, music can be broken down into three fundamental constituents—melody, harmony, and rhythm (**Figure 1**), each of which is subserved by overlapping but distinct neural networks. These fundamental ingredients may, in rare instances, be experienced in isolation, such as when listening to the single melodic lines in Gregorian chant or the genuine effect of surprise created by the epic drum fill in Phil Collins' *In the Air Tonight*—introduced unpredictably after 3 minutes of vocals and keyboard. Mostly, though, these constituents interact in creating unified musical experiences of a unique cognitive and emotional quality. In recent years, it has become increasingly clear that to understand why people are so engaged by music, we need to understand the neuronal underpinnings of music perception, which in turn are closely linked to action in the form of overt or covert movements[1,2] and emotion[3]. Accordingly, music perception engages brain networks related to action, emotion and learning in addition to the auditory system (**Figure 1 D**).

A particularly important feature of music is that its structure often involves patterns that allow listeners to form expectations, based on statistical learning, which may subsequently be fulfilled or betrayed. The experience of music is therefore intimately linked to brain-bound predictive models: e.g., tonality (the experience of a hierarchy of relations pointing towards a tonal center in melody and harmony); and meter, the experience of regularly recurring rhythmic patterns and accents, that underwrites the way we move regularly to sometimes highly irregular musical

rhythms. In this review, we describe the process of listening to music, where we continuously construct predictions of what happens next in a musical piece, and how this process gives rise to perception, action, emotion, and—over time—learning as formulated in the predictive coding of music (PCM) model[4]. In brief, the PCM model states that when we listen to music with melody, harmony, and rhythm, the brain deploys a predictive model—based on prior experience—which guides our **perception**. Take the example of a repeated syncopated rhythm (**Figure 2**), a rhythm wherein one beat feels displaced by a fraction. Here, we experience an error at the unexpected, syncopated note. This may drive an impulse for **action** in the form of enforcing the beat by tapping the foot. This active listening process forms the basis of musical **emotion** and **learning** which updates our underlying predictive model over time. Music is thus a powerful tool for studying the predictive brain, due to the way its structure licenses anticipation.

In the following, we provide an overview of music perception in the brain. We first introduce PCM. We then discuss the fundamental constituents of music in relation to PCM before turning to more complex music processing that entails action, emotion, and learning. Rather than focusing on the link to language – and clinical applications – we focus on the basic neuroscience of music processing in the brain and requisite prediction-based brain mechanisms. Finally, we consider generalizations of PCM to encompass musical interaction and communication in interpersonal relationships and hierarchical organization in groups.

The predictive coding of music (PCM) model - an enactive perspective on music perception

Prediction is increasingly considered a fundamental principle of brain processing. Theories of predictive processing offer explanations for how specialized brain networks can identify and recognize the causes of its sensory inputs, integrate information with other networks, and adapt to new stimuli. Recently, active inference, an influential theory of predictive processing[5] has proposed that perception, action and learning constitute a recursive Bayesian process by which the brain attempts to minimize the prediction error between sensory input and top-down predictions of that input (**Box 1**).

For many years it has been clear that music can only fully be understood in the light of prediction[6-8]. Music-related predictions are linked to various emotions, and the relationship between musical anticipation and emotion has been proposed to be associated with survival-related anticipatory brain mechanisms[9]. Musical expectations are evoked by auditory (bottom-up) sensations on one hand and depend on the brain's (top-down) predictions on the other. Predictive brain mechanisms depend on long-term plasticity and learning[10] (forming e.g. schematic expectations), familiarity with a particular piece or genre of music[11] (veridical expectations), short-term memory for the immediate musical past (forming dynamic expectations), and on deliberate listening strategies[12]. The neuronal mechanisms and functional architectures underlying musical expectation are thus shaped by culture, personal listening history, musical training, and biology[13].

Recently, the research into music perception has turned towards experiments modelling musical structure, which evinces anticipation [14] and predictive mechanisms. To integrate these approaches in neuromusic research, we have developed the predictive coding of music (PCM) model in several recent papers (**Figure 2**). This model is a special case of the predictive processing (**Box 1**) theory of brain function for music, with an explicit focus on the influence of biological, cultural, and contextual factors.

PCM proposes that music perception, action, emotion, and learning are recursive Bayesian processes, by which the brain attempts to minimize prediction error[15] as formalized in enactive versions of predictive processing (a.k.a., active inference). Accordingly, the processes underlying music perception and action are coupled, such that perception minimizes prediction error by updating the predictions, whilst action reduces prediction error by generating predicted sensory signals. Emotion, attention, and motivation act as Bayes optimal biases to contextualize prediction, thereby guiding behavior, action, and learning.

The notion of the brain as a hierarchical prediction machine—where sensory input is constantly held up against the brain's beliefs about the causes of this input—is consistent with music processing, which is not just passive and bottom-up but rests on top-down predictive processes as demonstrated for melody, harmony, and rhythm below. In the case of ambiguous musical stimuli, the listener may—depending on musical training—make the active decision to listen attentively to the same piece of music with different meters or tonalities (**Figure 3**). Hence, a Bayesian formulation of predictive coding, applies naturally to processes that involve the inference of hidden or latent causes —such as meter and tonality—from the musical surface.

As we shall see below for melody, harmony, and rhythm, certain event-related potentials (ERPs) recorded with electro- or magnetoencephalography (EEG or MEG) are clear experimental markers of prediction error: for melody and rhythm we have the mismatch negativity (MMN), for harmony we have the early right anterior negativity (ERAN).

In particular, attentional selection—of which prediction errors to resolve—rests on predicting not just the content of sensory streams but their *predictability or precision*. Put simply, prediction errors are only useful when things are predictable. The Bayesian belief updating inherent in PCM is literally precision engineered, in the sense that it rests upon predictions of predictability.

A crucial concept in predictive coding is the notion that prediction errors are weighted by their expected precision or predictability. In short, the brain has to select the prediction errors that drive Bayesian belief updating and ensuing top-down predictions. This selection can be regarded as a kind of mental or covert action that equips standard predictive coding schemes with an enactive and attentional aspect. The requisite synaptic mechanisms are thought to depend on neuromodulatory synaptic gain control that underwrites sensory attention and attenuation. The importance of precision for predictive processing has been shown in studies of auditory perception, where the MMN to an oddball is modulated by its predictability[16]. In music perception, the findings that MMN amplitude is reduced (to rhythmic deviations and mistuned pitches) in less predictable contexts offer clear examples of so-called *precision-weighted prediction errors*[17,18]. These studies demonstrate the scope of PCM for explaining the fundamental aspects of music processing.

Learning can be cast as minimizing precision-weighted (i.e., predictable) prediction errors over time, via experience-dependent plasticity. This is equally true for implicit and explicit learning. An example of implicit learning is how melodic singing ability or tonality perception schema are learned during repeated exposure—as evinced by ERAN responses to harmonic expectation violation in non-musicians and musicians alike[19,20]. Learning to play an instrument involves implicit learning but introduces additional top-down effects on perception and action. Hence, explicit learning often implies altered processing of musical structure violations—compared to non-musicians—as reflected in enhanced ERAN or MMN responses that could be regarded as reporting precision-weighted prediction errors for melody, harmony, and rhythm (please see below).

Importantly, experiments have shown that the brain architecture subserving precision-weighted prediction errors differs depending on which musical phenomenon e.g., melody, harmony or rhythm that are studied. For example, a predictive coding-based analysis of the effective connectivity in a melodic oddball functional MR imaging (fMRI) paradigm revealed that mismatch responses are best explained by a fully connected bilateral auditory network comprising the primary auditory cortices (A1) and the planum temporale (PT)[21]. Here, the observed increase in excitatory connectivity from left A1 to PT, has been interpreted as the passing of precise prediction error from lower to higher areas of the hierarchical processing network[22,23], and the decrease in inhibitory connectivity within the left A1 as precision-related increase in the gain of the superficial pyramidal cells encoding prediction error[24,25]. In contrast, as discussed later, studies of the pleasurable experience of musical harmony reveals predictive coding

mechanisms including precision-weighted prediction errors[26] related to emotion and reward brain networks[26], whereas studies of musical groove implicate additional motor related areas.

In the following, we review music perception, action emotion and learning in the light of PCM. We begin by summarizing the significant progress in our understanding of music perception of melody, harmony, and rhythm.

Perception of music

Melody

In most musical styles, melody, i.e., patterns of pitched sounds unfolding over time, is an important part of what defines and distinguishes one piece from another. Sing the first eight notes with any rhythm and you will immediately identify Beethoven's fifth symphony.

When you press a key on a piano keyboard, the resulting note comprises a fundamental frequency defining its pitch[27] and a series of overtones that contributes to its tone quality (timbre) which distinguishes it from other instruments[28,29]. A large corpus of research has been devoted to the study of musical pitch, and it is now widely accepted that the brain can extract a single pitch percept from complex tones, even in the absence of the fundamental[30,31]. Pitch perception can be separated into pitch height and pitch chroma. Two different piano notes may have different pitch heights but still be perceived as having the same chroma (e.g., the note C in different octaves). The auditory cortices are central to pitch processing, where fMRI suggests that pitch height is processed in PT posterior to A1, and chroma change in the planum polare, anterior to A1[32]. In general, the recognition of pitch from spectrally complex sounds is thought to be handled by a ventral stream, projecting from primary auditory areas along middle and anterior regions to the inferior frontal gyrus (IFG, **Figure 1 D**). A dorsal stream, projecting from primary areas via the PT over the parietal cortex to the dorsolateral prefrontal cortex (dlPFC)[33], is supposed to support sensory-motor integration, articulation, and memory functions[34] linking neural apparatus for melody perception and action[35].

Once musical pitches are combined into melodies, global properties emerge, such as melodic contour[36], melodic expectations, and tonality[37]. Most melodies point to a certain tonality, even though twelve-tone composers such as Arnold Schoenberg often try to avoid it. In its simplest form, a melody such as 'Frère Jacques" is confined to a subset of pitches—a scale—with a tonal center, such as a C-major scale which corresponds to the seven white keys on a piano keyboard, excluding the black keys. The tonality is not necessarily expressed directly in the auditory input to the ears[4] but an endogenously generated hierarchical predictive framework which underwrites perception of melody and harmony (**Figure 3**). Listeners, even without explicit musical training, have implicit knowledge of the statistical regularities of melodies of their own culture. This knowledge[38] is constantly applied to form musical expectations by comparing a given note to the given statistical distribution[39,40]. The tonality is an example of one predictive model, which underlies melodic expectations.

Importantly, it is possible to model melodic expectation and uncertainty mathematically. Several models exist, e.g., the Information Dynamics Of Music (IDyOM) model which assigns measures of information content (contextual unexpectedness: c.f., prediction error) and entropy (uncertainty: c.f., negative precision) to each note of a scale using short-term and long-term statistical regularities. Mathematical modelling has the advantage that it allows the study of perception of "real" music and obviates the need for tailored musical excerpts, as traditionally used in many experiments. Recently, such models have been used to study the neural processing of melodic expectations[41] in particular using a marker of pre-attentive auditory expectancy violation, the MMN[42,43] (**Figure 1**). Importantly, MMN amplitude and behavioral deviant detection is reduced to mistuned pitches in high-entropy (unpredictable) compared to predictable melodic stimuli as modelled by IDyOM[18]. In other words, the more precise our melodic predictions—according to culture dependent statistical learning—the larger the MMNs to surprising notes.

Individual differences play an important role for the predictive processing of melody. In general, musicians score consistently higher on tests that involve distinguishing between different melodies[44], which correlates with higher amplitude and shorter latency of the pre-attentive the pre-attentive auditory response such as the MMN to expectancy violations[11,45], indicating a more precise predictive model in these individuals presumably as a result of explicit learning.

Melody perception recruits parts of the brain that are specialized for other purposes than audition such as motor, emotion, and cognitive evaluation. as shown e.g., in the so-called *free listening* paradigms[46-48], which combine music information retrieval (MIR) of real music recordings with fMRI or MEG[47,48]. Moreover, action and perception can be engaged even without stimuli, as demonstrated by ERP, positron emission tomography (PET), and fMRI studies showing that the formation of musical mental images engages auditory sensory and premotor areas[34,49,50]. In sum, the study of melody processing has moved towards using naturalistic melodies, emphasizing the coupling of perception and action related brain mechanisms, where prediction plays a key role.

Harmony

Melodies, especially in Western music, are typically accompanied by harmony created by chord progressions played on instruments such as a piano or a guitar or by many instruments playing different notes at the same time. An example of a single chord is a C major triad, which is the combined sound of the three notes C, E and G.

Chords may in themselves give rise to musical emotions, such as the sound of a C major triad which by Western listeners is perceived as happier than a C minor triad (C, Eb and G)[51]. Since the ancient Greeks it has been known that the physical properties, such as the "roughness" of musical harmonies—two or more notes together—are uniquely determined by the integer relationships between the fundamental frequencies of these notes, leading to a differentiation in Western harmony between consonant and dissonant intervals and chords. The smaller this integer relationship is, the more likely the notes are to be misperceived as one single sound, and this trend has been observed also in musically distant cultures[52]. This is tied to the experience of sound roughness[53] which is thought to be related to the bandwidths of critical auditory bands linked to the inability of the basilar membrane in the cochlea to separate notes which are very close[54,55]. Importantly, however, roughness may be perceived differently according to musical style and culture.

Chord progressions (different chords played successively) establish musical expectations and typically a sense of tonality. Whereas tonality is known in music from all studied cultures[56], neuroscientific studies have mainly concentrated on Western harmony[57-59], which follows a set of rules, e.g., in different types of harmonic cadences. Breaking these rules, elicits a specific marker of harmony expectancy violation, the ERAN (**Figure 1**), discovered by Koelsch and colleagues using EEG/MEG[60,61]. The ERAN peaks at 150-200 ms after deviant onset, its latency and amplitude are modulated by attention or knowledge of impending outcomes[62], and musical training[20]. Its sources have been linked to the IFG using MEG as well as fMRI in non-musicians[63,64]. It is most often studied in semi-attended paradigms, where the task requires participants to attend to the musical stimulus but not to the deviating chords. In contrast to the MMN—which can be evoked by a local mistuning of a specific chord— the ERAN amplitude and latency depend strongly on the deviant's position within the chord progression—how well it fits with the rules of harmony[65-67]. The rules of harmony have often been termed musical syntax or grammar and are thought to be encoded through statistical learning. The development of a neural architecture for melody and harmony has been studied in adult non-musicians using artificial experimenter generated musical grammars[68,69] where recognition and liking ratings indicate a generalized probability-based perceptual learning mechanism as the basis for remembering and appreciating music. The precise function of the IFG is still unresolved but is likely related to higher level processing of the temporal order of sequences[70].

With musical training, perception and action networks in the brain become more tightly coupled, which facilitates more precise active inference. Studies have shown motor-related cortical activity in professional pianists listening to piano music[71-73], and activity in a frontoparietal motor-related network in non-musicians listening to a learned melody after practice playing[74]. For trained pianists, internalized harmonic expectations in auditory and motor networks may furnish modality specific networks for harmony prediction that interact with the IFG to optimize action and perception[75]. It is thus likely that musical expertise affects harmony processing by engaging motor mechanisms because of the active experience in auditory-motor association.

Like melody processing, a recent trend has been to model both music and expected brain responses mathematically, enabling more ecologically valid approaches to the study of harmony. This allows for a better understanding of the often-reported link between musical harmony, emotions, and pleasure. Modelling harmony regularities in a corpus of harmonic sequences from the Billboard "Hot 100" pop songs, Cheung and colleagues[26] showed maximal pleasure ratings to expected chords in unpredictable chord sequences and conversely to unexpected chords in predictable chord sequences, and linked this to activity in key limbic and reward related brain structures: the amygdala, hippocampus, and the nucleus accumbens (NAcc).

The above studies highlight the ability of harmony to engage motor, emotion and learning related mechanisms. Cognitive studies of harmony, however, mainly use Western harmony as the source of auditory stimulation which reduces the generalizability of the results. At the level of a single interval native Amazonians with limited exposure to Western music do not exhibit Western-like preference for consonant compared to dissonant intervals[76], indicating that aesthetic preference for certain intervals may be culture-dependent. Importantly, the statistical regularities or the harmonic syntax (captured by ERAN)—which undergird many predictive processes related to harmony—differs between cultures and styles of music, leading to quite different expectations when e.g., listening to blues compared to Beethoven. The well-known associations between major/minor and happy/sad emotions in Western harmony are prime examples of the complexity in determining the influence and interaction between universal and cultural factors in the perception of harmony. This association is not found in all musical cultures[78], which speaks against a universal relationship. However, it has been proposed that the lower pitch intervals, and slightly lower average pitch in the minor mode simulate speech when we are sad[79-81]. It is still unclear if this is a universal principle expressed differently in different cultures. Nevertheless, basic emotions may still be recognized in realistically sounding music material across cultures[77]. However, this is an ongoing field of study.

In sum, predictive structures in musical harmony (often referred to as syntax) have proven an invaluable domain for studying prediction error and its relationship to musical emotion and its relation to musical learning and culture. For harmony and melody, tonality offers a predictive context in the PCM model and this profoundly affects brain processing since both operate and interact in a shared pitch-based domain. The PCM model may explain why a melodic line without harmonies still suggests an underlying harmonic scheme—through the way the brain generates implicit predictions based on harmonic priors.

Rhythm

Perception of musical rhythm is a burgeoning topic in cognitive neuroscience partly because of its ability to link the body and the mind, perception, and action. When we listen to "Blame it on the Boogie" by The Jacksons (**Figure 4**), it is difficult to refrain from tapping a foot or bobbing the head to the beat. Rhythm can be produced by the onset of the notes in a melody or without a melody when played on designated percussion instruments, where pitch may be less clear[82]. Its perception usually involves simultaneous perception of an evenly spaced pulse[83], and a meter, which structures this pulse and its subdivisions into patterns of differentially accented beats. Listening to pulse trains involves prediction of following events[84] as indicated in studies showing brain responses to omission of a beat or after the end of rhythmic sequences[85-88].

Brochard and colleagues[89] provide strong evidence for meter perception in the simplest possible experimental setting, when they showed that listening to an entirely regular and unaccented metronome causes the brain to automatically register some beats as more salient than others, even in the absence of any such structure in the stimulus. Hence, the pulse and the meter are not necessarily expressed directly in the auditory input to the ears[4] but emerge under hierarchal predictive processing which underlies the recognition of successive musical events over time[90-92]. Perception of a clear musical meter facilitates rhythm memory[93,94], learning[95], and perceptual sensibility even at a young age[96].

Despite the possibly innate human ability to synchronize to the musical meter, this ability is not easy to model computationally. Recently, Large and colleagues created a neuronal network model with two hierarchical levels; one corresponding to the sensory system modelled with a simple Hopf bifurcation, the other corresponding to the motor system tuned to operate near a double limit cycle bifurcation[97]. This model was able to explain participants' ability to synchronize with increasingly syncopated rhythms. Accordingly, frequency tagging (**Figure 1**) in electrophysiological recordings show that even for such rhythms in which the meter is not acoustically accented, the fundamental frequencies of the meter still dominate the signal[98,99]. Yet, the neural entrainment to rhythm and the different contributions of auditory and motor cortical and subcortical structures in establishing the meter percept are still far from well understood[100], and it is essential to acknowledge a substantial top-down influence on meter perception[15,101] which may be nuanced by cultural [102-104], and biological factors (**Figure 2**).

As with melody and harmony, musicians score consistently higher on rhythmic ability tests [44], and have higher amplitude and shorter latency of the MMN to violations[105], suggesting that musicians deploy more precise predictive models.

The literature on rhythm perception discloses the involvement of the auditory pathway in detecting structural deviations from the meter[105], marked by the MMN, which again depends on the complexity of the metric context[17] and the involvement of large parts of the motor system, the premotor (PMC), supplementary motor (SMA), basal ganglia, and cerebellum[106,107] for mere listening to auditory rhythms[108]. This motor system activity can—to some extent—be attributed to the establishment and maintenance of the musical pulse/meter[109]. This process, which underlies our ability to dance to music, is measurable already in newborns[110], yet influenced by training, where infants can be trained to either a duple (2/4) or a triple meter (3/4) of the same ambiguous rhythm[111]. Furthermore, synchronizing to the same meter may lead to prosocial behaviour[112].

The complexity of the brain circuits underlying meter perception may explain why meter perception is so rarely observed in non-human animals and never with the same accuracy and flexibility as in humans. Humans have the ability to synchronize to a simple metronome at different tempi between approximately 40 and 400 BPM depending on musical expertise, and do so by predicting the subsequent beats—and may perform the task across modalities[113,114]: In contrast rhesus monkeys can only with great difficulty be trained to follow the beat at different tempi and then tap some hundreds of milliseconds after the beat instead of predicting it[115]. MMN recordings to onbeat and offbeat deviants show that monkeys are sensitive to the isochrony of the stimulus, but only humans to its metrical structure[116]. Studies in chimpanzees show equally poor results in beat synchronizing to metronomes[117]. Even though they may possess the ability to predict the upcoming beats, they lack tempo flexibility. Hence, rhythmic ability for music clearly depends on the expressivity or depth of predictive coding of the human brain.

In sum, the study of musical rhythm demonstrates how sensory input provided by auditory rhythms (bottom-up) are met by predictive models such as the meter (top-down), and how this process gives rise to auditory-motor coupling in the human brain. The involvement of the motor system entails higher precision of the auditory predictions as hypothesized by prediction-based models. In the following we consider more complex musical phenomena, which integrate melody, harmony and rhythm and exemplify the crucial role of precision-weighted prediction error.

Action

A prime example of how PCM—and the concept of precision-weighted prediction error—can inform our understanding of music processing is the study of groove. Groove is defined as the pleasurable sensation of wanting to move to music[118]. Why do people rush to the dance floor when listening to the funky grooves on James Brown's records and move to the music with such apparent pleasure[119-121]? Groove research primarily relates to music originating in the African diaspora, such as soul, funk, disco, Latin, jazz, hip hop, and other dance-related genres[122].

Typically, these styles are characterized by the presence of a rhythm section comprising percussion, bass, and chord instruments. This rhythm section is supposed to keep a constant beat—often taking the form of a constant syncopated rhythmic pattern repeating after one, two or more bars throughout longer parts of the musical form. Groove is a seemingly unique and ubiquitous trait of humans, which emphasizes the link between perception and action formulated in active inference[123]. In this regard, the brain's constant evaluation of prediction error arising from syncopations—defined as the appearance of a beat on a metrically weak accent preceding a rest on a metrically strong accent[124]—has been proposed as one of the underlying mechanism of groove[125] and as one of the reasons why we move to music.

The influence of syncopations on the experience of groove can formally be operationalized through the PCRI (predictive coding of rhythmic incongruity) model[126] where brain and behavioral responses are modelled in terms of the precision-weighted prediction error, i.e. the product of the metrical predictability (precision) and the stimulus deviations from the meter (**Figure 4**). Importantly, this model explains the observed inverted U-shaped relationship between degree of syncopation in and the experience of groove[127,128], where rhythm excerpts with medium levels of syncopations are rated as more pleasurable—and movement inducing—than low and high levels of syncopations.

According to PCRI, these medium syncopated rhythms optimize what the system treats as precision-weighted prediction error, in that both the prediction error and the precision of the prediction are at intermediate levels in processing hierarchy. In intermediate syncopated rhythms the brain may, according to active inference, resolve prediction error by either revising predictions or through action—e.g., by moving the body. Actively resolving prediction errors may explain our drive to reinforce the meter—by moving in time with the beat—while attenuating the precision of proprioceptive and auditory prediction errors. In contrast, rhythms with lower levels of syncopation evince little prediction error and less incentive to move. Conversely, for the highest levels of syncopations our mental model of the meter is less precise than the sensory evidence, precluding sensory attenuation and movement.

The U-shaped relationship between syncopation and groove experience has been replicated independently of culture and rhythmic proficiency[129] and using physiological measurements such as pupillometry[130]. It has been tested for rhythm and groove in a within- and between-culture approach[122,127,131] and is influenced by musical expertise[131].

Optimal levels of the pleasurable sensation of wanting to move have been linked recently to neural activity in the brain's motor and pleasure networks[132] (**Figure 4**), and can thus be seen as a result of precision-weighted prediction error arising from a discrepancy between the syncopation in the auditory input and the motor system's propensity towards isochronism[97,126]. Importantly, optimal groove experience was linked to activity in NAcc and the orbito-frontal cortex (OFC) which are key regions of the reward network that is particularly sensitive to the predictability of the consequences of action.

It is important to note that the relationship between the rhythmic sensory input and the schematic expectations of the meter is only one of several interacting predictive processes taking place contemporaneously. When rhythm section patterns are repeated over and over again, the brain forms short-term rhythmic expectations that—after repeated listening—may turn into veridical expectations about the time course of a specific piece of music[49,133,134]. It is equally important to note that the repeated patterns in many non-Western grooves, such as the Afro-Cuban tumbao, still supports a stable meter sensation in experienced listeners, even though they contain few onsets on the most salient metrical positions. It is therefore an ongoing debate to which extent the meter—which arguably is a construct based on a Western musical tradition, where there is a strong correlation between note frequency and metrical accentuation—can be considered the most important predictive reference structure in other styles of music[135].

In sum, PCM proposes that the pleasurable wanting to move is mediated by prediction-based brain mechanisms which optimize the syncopation related precision-weighted prediction error, thereby engaging the brain's motor and

reward systems. The right level of syncopation offers the opportunity to actively resolve uncertainty by moving—which can lead to the experience of 'pleasure'.

Emotion & Pleasure

A defining feature of music, closely related to theories of its evolutionary origin[136], is its ability to evoke a range of feelings and emotions, which may be similar to everyday emotions, such as happiness, sadness, surprise, and nostalgia, or provide music specific experiences, such as the sensation of groove described above. Even though music is clearly able to express everyday emotions[137,138], and adults listen to music partly to regulate their affective state[139], it is not possible to equate valence and liking. A negatively valenced emotion such as sadness is the eighth most commonly reported emotion induced by music[140,141]. Furthermore, there is a dissociation of valence and pleasure ratings[78] as well as a dissociation of the brain networks underlying the experienced valence and pleasure in sad and happy music. Liked music elicits more activity in the cortico-thalamo-striatal reward circuits than disliked music, regardless of whether the music is sad or happy[142]. Because of this apparent paradox, musical sadness is the subject of several recent studies[143,144], and multiple theories try to explain its existence, often pointing to societal and individual benefits[145]. As an example, the catharsis process by which sad music is seen to provide relief for negative emotions that we all experience in a safe context is thought to promote social cohesion instead of, for instance, aggression.

The different ways in which the human brain might carry out the translational process from music to emotion can be explained by several psychological mechanisms[146] which typically fall into three categories (1) *Hardwired responses*; evoking universal survival-related responses such as for example when brainstem responses to loud sounds trigger fear responses; (2) *Extramusical associations*; where music links to some extra-musical space that carries the particular emotion such as evaluative conditioning, emotional contagion, visual imagery, and episodic memory; and (3) *Anticipation*; where musical structure establishes, fulfils or disappoints expectations which are set up within the music itself. Whereas hardwired responses and extra-musical mechanisms in principle can be elicited by sounds alone, anticipation depends on the organization of sounds into a meaningful succession of events—a defining characteristic of music that is closely connected to predictive coding. Since Leonard B. Meyer's book[7] it has become increasingly clear that music anticipation may induce a variety of complex emotional responses such as awe, surprise and discomfort, evoke laughter, foot tapping, humming, tears, and a lump in the throat[9]. It can give rise to psychogenic responses such as "shivers down the spine", increased heart rate, and increased perspiration[147].

Functional neuroimaging studies of music and emotion show that music perception engages emotion-related brain networks and that music can modulate activity in limbic and paralimbic brain structures such as the amygdala, NAcc, hypothalamus, hippocampus, insula, cingulate cortex as well as the OFC[3]. An outstanding question is to what extent the emotion related networks involved in processing of music are mediated by *universal*, *cultural,* or *individual* mechanisms: i.e., which aspects of music perception are developed only after exposure to a specific musical culture. One pioneering study[77] pointed towards the above chance level recognition of basic emotions—such as happy, sad, and scared/fearful emotions in Western music—in listeners from the African Mafa tribe who were culturally isolated from Western music. Importantly, though, the Mafas showed much lower emotion recognition performance compared to Western listeners. Furthermore, in a recent large scale internet study US and Chinese listeners identified 13 distinct types of subjective experience associated with music in both cultures[148]. It is, however, unclear, to which extent music emotions can be universally recognized[149], or how much they are a result of statistical learning caused by increasingly globalized music listening behaviors[150]. Speaking to the latter, recent modelling approaches highlight the importance of aligned musical priors to the cross-cultural experience of music emotion[151,152]. Predictive coding has therefore become a hot topic in the study of musical emotions.

Musical pleasure: A particularly interesting example of predictive processing of music is the link between musical anticipation and pleasure, similar to the well-established difference between wanting and liking[153]. Music pleasure was originally proposed to be linked to positive reward prediction errors, which arise when what is heard proves to be better than expected. This was first studied through the experience of musical chills that were correlated to activity in the reward system[154]. Recently, these intense experiences have been shown to lead to dopamine release in the striatal system[155,156] with distinct roles for the caudate (anticipation) and the NAcc (reward experience)[157] and related to degree of emotional arousal[158]. The critical role of the interaction between the auditory cortex and the subcortical reward network for the enjoyment of music is further supported by studies on the very few people for whom music holds no reward value—despite normal perceptual ability and normal auditory and musical perceptual abilities—as well as reward-related responses in other domains[159]. These individuals show reduced NAcc responses and decreased functional connectivity between the right auditory cortex and ventral striatum—including the NAcc—compared to their responses on a monetary gambling task and compared to other participants with normal or greater than average pleasure responses to music[160].

Based on active inference formulations of predictive coding models, Gebauer et al.[161] hypothesized that both confirmation and violations of musical expectations are associated with the hedonic response to music via recruitment of the mesolimbic system and its connections with the auditory cortex. This was recently supported by a demonstration of associations between music-induced pleasantness and musical surprises in the activity and connectivity patterns involving the NAcc—a central component of the mesolimbic system[162]. Furthermore, this study found surprise-related activation in the NAcc that was more pronounced among individuals who experienced greater music-induced pleasantness.

A significant contribution to the understanding of the predictive coding mechanisms of musical pleasure was the aforementioned study by Cheung and colleagues[26] who combined computational modelling of expectation in naturalistic chord sequences in songs from the "Billboard Hot 100" with fMRI. They found optimal pleasure to be associated with surprising chords in predictable sequences (high precision, high surprise) and predictable chords in unpredictable sequences (low precision, low surprise) and that this interaction corresponded to activity in the amygdala and hippocampus, whereas the NAcc only reflected precision. This is consistent with optimal zones of predictability and uncertainty in musical pleasure found in modelling studies[163].

Closely related to the subject of musical pleasure is the study of musical taste. Why do people with very similar cultural exposure to music often differ greatly in musical preferences? This is a complex question which includes psychological explanations[164], such as the well-known mere exposure effect showing increased liking with repeated listening to musical pieces[165]. Other important determinants of musical taste are contextual factors such as importantly sociological reasons where music can be seen as a means to express group affiliation[166,167]. In addition, as illustrated by the studies on music anhedonia, individual factors play an important role in music perception and thereby musical taste. Personality as rated e.g., by the 'Big Five' or the Zuckermann sensation seeking score has therefore consistently been related to differences in musical taste[168,169]. Since musical pleasure depends on whether culturally learned musical expectancies are fulfilled or violated[170], listeners often exhibit biases favoring music of their native culture, making yet another case for predictive coding as an underlying mechanism of musical taste[164].

In sum, predictive mechanisms in music and the brain are key to understanding complex questions related to musical emotion. A full description of the precision-weighted prediction errors involved in music emotions still eludes us, but the contribution of predictive coding is becoming clearer.

Learning

One of the best-studied individual factors influencing music perception is musical training and learning, which is integral to PCM. Playing music is a highly specialized skill that places immense demands on the underlying neural

resources. Accordingly, several cross-sectional studies of music perception and performance have indicated training-related changes in networks for auditory processing, motor representations, and for emotion, visual perception, and mental imagery, hence the study of how musicians' brains evolve through daily training is an effective way of gaining insight into the brain's remarkable potential for change during development and training[171]. The differences in musical cognitive skills between musicians and non-musicians can be correlated to differences in both brain structure and function. Classic studies have shown morphological differences in the fiber bundle in the corpus callosum[172,173], increases of cerebellar volume[174], grey matter volume increases in primary motor and somatosensory areas in the left precentral gyrus, premotor areas, and left cerebellum[175,176], in areas involved in temporal structuring of language and music[177,178] and in areas involved in auditory perception[179], as well as specific effects of musical training on white matter development[180,181]. These studies are coupled with functional differences between musicians and non-musicians related to auditory and motor areas[1,182] dependent on musical instrument[183,184], practice habits[185], level of expertise[186], and the style of music they play[187,188].

These functional and structural differences which are correlated with differential music training have been taken as evidence for long-term influence on the brain due to active inference and learning. It is, however, not possible to draw conclusions about causality from cross-sectional approaches. Recently, there has been a growing amount of causal evidence from longitudinal approaches highlighting the influence of long- and short-term training on brain anatomy and function and in particular the development of auditory and motor processing, and the auditory-motor coupling. In a pioneering study, using direct current EEG analyses, Bangert & Altenmüller[189] showed auditory-motor coupling changes in the cortex of beginners after as little as 20 min of musical piano training. The enhanced coupling of brain resources for perception and action has recently been related to increased functional connectivity within the sensorimotor network and increased functional and structural connectivity of the auditory-motor network after 24 weeks of musical training[190]. In addition, a recent study showed increased activity in fronto-parietal and cerebellar areas related to storage of newly learned auditory-motor associations following 6 weeks of piano training when participants were merely listening to the melodies[191].

Viewed in the light of PCM, the above studies indicate that auditory-motor learning leads to increased recruitment and adaptation of higher-order action-related resources (top-down) related to mere listening to music (bottom-up). Targeting the development of auditory predictive coding longitudinally in children, Putkinen and colleagues followed preschool children over several years obtaining measures at ages 2–3, 4–5 and 6–7 years from children who attended a musical play school throughout the follow-up period and children with shorter attendance to the same play school[192]. Their results showed that the musical group activities enhanced the development of the MMN to timbre, melody, mistuning, and rhythm. This was taken as evidence for a facilitation of predictive coding of neural sound discrimination of musical training during early childhood. In later childhood between the ages of 7 and 13 years[193], the MMNs related to deviants in harmony increased more in the music group than in the control group despite lack of evidence for pre-training neural differences between the groups in sound discrimination. These results are consistent with earlier findings from cross-sectional studies of training-related enhanced precision in melody, harmony, and rhythm perception in children (see e.g. [194]). (se

Several cross sectional and longitudinal neuroscientific studies point to a putative transfer effect of musical training to cognitive abilities and brain processing related to reading and language skills[195-198], cognitive inhibition tasks[172], and to music training as a possible supplementary tool for helping children with developmental disorders[199] such as dyslexia[200,201]. While the causal relationship between musical training and music-related brain processing seems well established, it is still controversial to claim that music training has a positive effect on other cognitive abilities[202,203].

In sum, cross-sectional and longitudinal studies of musicians and musical learning, elucidates how predictive mechanisms for music are shaped by learning[204-206]. It appears that the heightened demands on audio-motor coupling in music performance shapes brain structure, and the ability to form music-related predictions with high precision. The studies shed light on how the complex relationship between factors such as musical training, culture,

listening history, music-stylistic preferences, context, personality, and genotype significantly influences the precision and ensuing amplitude of the explainable prediction error, as well as how the brain infers a predictive model from the musical context (**Figure 2**). These factors are also crucial for how we understand the music of others. In the following we propose how music and PCM can be extended to encompass the role of communication in dyadic interactions and hierarchical organization in groups.

Musical communication

Even though most of the literature reviewed in this article treats music perception in the individual brain, music is fundamentally a social phenomenon, in that we make, listen, and dance to music together. This makes it a fine-tuned instance of coordinated human interaction that involves interpersonal synchronization, social entrainment, learning, improvisation, and communication (see **Box 2** for an example of this in other animals). Recently, the development of adequate research methods such as dual EEG has prompted a line of neuroscientific and behavioral research into musical interaction[207,208]. It shows how competence, social context, and mind set, such as empathy perspective taking, may promote interpersonal coordination[209] (**Figure 5**).

Musical interactions rely heavily on prediction. While playing we continuously make predictions about the sensory consequences of our own actions which we generally use to attenuate predicted sensations and amplify those caused by others [210]. This this selective attention and attenuation is found throughout the animal kingdom (**Box 2**) but the more advanced ability for shared predictive processing—needed for the full experience of music—has only so far been found in humans. Joint action may thus be best understood within a predictive coding framework[211,212], where the emphasis is on establishing a shared narrative and mutual predictability. Recent studies have leveraged this perspective looking at musical interactions when two individuals tap together.

These paradigms typically involve two individuals who are placed in separate rooms with headphones and EEG equipment, finger tapping from the perspective of sensorimotor synchronization[114] looking at isochronous self-paced tapping[213], synchronization with a computer-generated metronome[214,215] or with piano recordings of self and other[216,217]. Typically, dyads contain leaders and/or followers who differ in terms of the degree to which they adapt to or rely on the actions of their partner to perform a shared task. Behaviorally, by studying the correlation between the participants' tap sequences, these studies have demonstrated that the interaction is guided by mutual efforts to reduce prediction error at the millisecond level, resulting in at least three different relationships between participants: leader-follower relations[218], mutual adaptation where both participants constantly adapt their taps to their partner's last tap[219], and leader-leader relations which may occur if tappers are highly rhythmically skilled musicians and both follow their own pulse without taking the auditory input from their tapping partner into account[220]. Importantly, participants adapt differently to each other depending on their underlying internal predictive model. When musicians tap together with different underlying musical meters, e.g. 4/4 and 3/4, they initially synchronize poorly compared to when they hold identical musical meters in their minds, even though this condition may be the more difficult[220]. The brain's predictive model directly influences the interaction dynamics.

Differences in dyad tapping behavior—e.g. exhibiting a leader-follower, mutual adaptation, or leader-leader tapping pattern—may be modelled using a coupled oscillator model, which contains one internal and one external Kuramoto oscillator per person, consistent with how the PCM model describes bottom-up and top-down influences on neural processing[221]. This is supported by EEG data showing that dyad members exhibiting mutual adaptation behavior evince intra-brain neural synchronization in an action-perception related brain network to a higher degree than leader-leader dyads[222] (**Figure 5**). These studies shed light on the predictive brain mechanisms underlying human social cognition in general—with specific implications for individuals with impaired or atypical social abilities and may be used for understanding musical interaction in differently organized musical ensembles in particular.

In sum, the dyadic tapping studies illuminate how musical interaction is guided by mutual reduction of prediction errors; in effect rendering themselves mutually predictable. They may serve as a model for how competence, social context, and dyadic interactions rest on predictive brain processing in general and serve as an example of how the PCM model may be extended to communication of musical meaning between individuals. This neuroscience research opens up a window to perhaps the most challenging question about music: *how music becomes meaningful*[223,224]. Accordingly, we speculate that what makes music meaningful from a neuroscientific perspective is when musical interaction over time shapes each of the participating individuals' attentional selection, engendering shared predictions of precision—and the synchronization of joint attention. This gives rise to shared musical expectancies which undergirds music perception, action, emotion, and learning.

Musical improvisation: Collective musical improvisation is a particularly demanding example of musical communication, where musicians' predictive models need to be aligned to a large degree. This is found in many styles of music e.g., in jazz, in which improvisation is the central, defining element, and where one of the most important purposes of compositions, is to serve as a framework for soloists to improvise on. In general, musical improvisation is seen by many researchers as a prime example of human creativity involving moment-to-moment interaction between perception and action[225-228]. The studies on jazz improvisation have consistently implicated several brain regions related to movements, motor sequence generation, attention and executive control, voluntary selection, sensorimotor integration, multimodal sensation, emotional processing, and interpersonal communication[229-231]. These include prefrontal brain regions, such as the pre-SMA, medial prefrontal cortex (mPFC), IFG, dlPFC, the dorsal PMC, and the auditory cortices[232,233]. This is not surprising since improvisation involves several processes simultaneously. A jazz musician has to play, listen to what the other musicians are playing, evaluate how the music sounds as a whole, while at the same time choose which direction to take and generate new phrases to play next[234]. Therefore, the neuronal processes underlying musical improvisation must necessarily be predictive in nature as well as dynamically shifting between different networks and states.

Recent cutting-edge neuroimaging connectivity measures—built on whole-brain computational modelling[235]—has made it possible to understand the changing predictive brain states which underlie communicative creativity in real time. The few studies of musical improvisation from a whole-brain connectivity perspective point to a large repertoire of brain states involving functional brain connectivity among frontal and parietal regions within default, salience, and executive brain systems[225,232,236,237]. Interestingly, this is similar to networks found in more general creativity tasks such as when participants perform the classic divergent thinking tasks—pointing towards musical improvisation as a model for understanding human creativity[238,239]. This is usually understood as carefully creating a sensorium in which the opportunity to resolve (i.e., explain away) prediction error is itself predictable—much like knowing the punchline of a joke resolves uncertainty in an entirely predictable fashion.

Even though musical improvisation involves predictive brain processes[240], there is an apparent paradox; although the primary purpose of the brain is to minimize prediction error, the primary purpose of improvisation is to create something new but aesthetically and emotionally appealing, which will then necessarily create prediction error. The improviser's difficult task is therefore to balance novelty and predictability in a way that generates pleasure responses in listeners[6] or stimulates their cognitive curiosity, in the same way that the dance music producer tries to hit the sweet spot of groove.

Because of the array of skills that are necessary to improvise at a high level, jazz musicians have been shown to outperform other types of musicians in domain specific tasks such as ear training task performance and in quantitative brain measurements to melodic expectancy violation[11]. This is coupled with findings of more distributed, globally-connected cortical networks in improvising musicians compared to higher within-network connectivity in classical musicians in resting state fMRI[241] as well as structural differences between these different groups of musicians[242,243]. In a recent study, the amplitudes of ERPs to chords that varied in expectancy were significantly correlated with behavioral measures of fluency and originality on a divergent thinking task, indicating a

putative transfer effect of music skills in music to more domain-general processes[244]. It remains an open question whether increased creative skills in general can be gained through musical improvisation training, and how this training may alter predictive mechanisms in the brain.

Conclusions and future avenues

The last twenty years of research into music in the brain has created a foundational understanding of how the individual brain processes music through predictive coding. The coming years could be dedicated to understanding the way music shapes interactions, how brains predict and synchronize through music, giving rise to shared meaning and perhaps even states of eudaimonia. As part of this journey there are many unresolved questions. In this vein, we note recent development towards cross-modal paradigms and the need for cross-cultural brain experiments to supplement our current knowledge about music and the brain, which is almost exclusively based on studies of Western music and participants. Another interesting—but so far unanswered—question is whether it is possible to self-generate an MMN during mental imagery of music? Would it be possible to have a pleasurable groove experience by imagining a funky rhythm without moving? Since there would be no sensory information to compare the internal meter model with, PCM would hypothesize that it would be difficult to generate precise prediction errors at least for lower-level predictions, but this is an empirical question that could be tested. Furthermore, only a few studies have considered the influence of different predictive frameworks in which musical events are embedded. It remains to future studies to clarify the interaction between melody, harmony, and rhythm e.g., the influence of shifting tonalities or metric displacement of a given melody, as well as the interaction between or lyrics and melody. Whereas this paper has mainly focused on predictive coding related to expectations in melody, rhythm and harmony, there are presumably also predictive mechanisms at work associated with voice leading, instrumentation, timbre, soundscapes, or musical events such as when there is the so-called "drop" in electronic dance music. These may be related to more abstract and maybe messier auditory prediction processes in the brain.

PCM offers a compelling but not exclusive framework for these endeavors. An alternative to PCM's probabilistic approach— to modelling hidden reference structures—is an oscillator-based approach simulating perception of meter and tonality in terms of the resonance of coupled non-linear oscillators[92,97,245,246]. The oscillator approach gives greater weight to stimulus properties than to the top-down effects of learned musical experience but as a result will have difficulties in accounting for the full range of musical phenomena as PCM or other prediction based approaches such as the "action simulation for auditory prediction" (ASAP) hypothesis, which proposes that the motor system contributes to the accuracy of auditory predictions by providing a periodic temporal framework through these connections[247,248]. However, the two accounts could usefully be combined given their different levels of processing, with the oscillator-based approach providing the basis for internal or generative models the brain uses to elaborate probabilistic predictions (see **Box 2** for an example).

Overall, we believe that understanding the neuroscience of music can provide a scientific foundation for clinical applications of music perception and training to help shape more meaningful lives.

Boxes and figure captions

Box 1 - Predictive processing and coding

Predictive processing (a.k.a. active inference) is a general theory of neural processing inspired by research in artificial intelligence, statistical physics, and systems neuroscience[5,249,250]. The basic idea can be traced back to the students of Plato, through Kant to Helmholtz[251], and to theories of perception as hypothesis testing[252]. Enactive versions offer integrative accounts of action and perception[253-257] by formalizing how specialized brain networks identify and categorize causes of sensory inputs, integrate information with other networks, and actively sample new stimuli[250]. Briefly, active inference proposes that perception, action, and learning are Bayesian processes by which the brain attempts to minimize hierarchical prediction errors. The figure is a schematic illustration of the computational architecture of neuronal message passing that underlies predictive coding in the brain. Since predictive processing is a generic theory of brain function, the precise architecture will vary depending on the functional anatomy in question. Part **A** of the figure shows the basic motif of connections, via which prediction errors are formed by comparing bottom-up input with top-down predictions. Crucially, these predictions can either be of the input or the precision (i.e., predictability) of that input. These are designated first and second order predictions, respectively. Part **B** of the figure describes the resulting hierarchical message passing implicit in predictive coding, in which forward or ascending connections convey prediction errors to higher levels, while backward or descending connections supply the predictions that enable the computation of prediction errors in the lower level. Red arrows indicate forward connections and black arrows indicate backward connections. In this example, unpredicted auditory input is passed forward to the auditory cortex in the form of ascending prediction errors (e.g., from the medial geniculate body). These prediction errors (red arrows) drive posterior expectations (e.g., encoded by deep pyramidal neurons) that return descending predictions (black arrows) to resolve—or explain away—lower-level prediction errors. At the same time, high-level expectations about the context generate predictions of precision (blue arrows) that modulate the gain of cells encoding prediction errors at the lower level (e.g., superficial pyramidal cells). This enables high levels to select the prediction errors that convey the most precise or predictable information (c.f., attentional selection). In short, there are two kinds of descending predictions in predictive coding: first-order prediction of content (black arrows) and second-order predictions of context (blue arrows). Here, context is simply the precision or predictability of prediction errors. The resulting precision-weighted prediction errors therefore mediate the selection the certain lower-level features that are consistent with higher-level constructs. This allows the PCM model to explain figure-ground phenomena in music such as for example selecting between different metrical interpretations of 3 against 4 (see main text).

Box 2 - Hermeneutics, communication, and music

Here, we argue that musical communication is a special case of fundamental communication between conspecifics; ranging from identifying a conspecific[258] through to sharing conceptual narratives[212,259]. The predictive processing gloss here takes a central role in the following sense: if I assume that you are like me, and you assume I am like you, then there is the mutual predictability for free. In music, this corresponds to sharing tonality or meter; technically, this mutual predictability can be formalized as predictive coding—or more generally Bayesian belief updating based upon shared (exchanged) sensory signals. If we share the same generative model, our neuronal dynamics can harmonize and evince a form of generalized synchrony[260]. From a cognitive perspective, this means we are "singing from the same hymn sheet". This enables an elemental theory of mind; enabling me to infer what you are 'singing'. An example of communication using birdsong is shown in the figure (adapted with permission from[212] (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article). It shows a simulation of neuronal hermeneutics; namely, what does this song mean to a bird. Here, two birds with the same generative models—but different initial conditions—sing for two seconds and then listen for a response. The shaded areas indicate which bird is currently singing: red for the first bird and blue for the second bird. When

singing, sensory prediction errors are attenuated so that predictions are realized through action. Conversely, when listening, sensory prediction errors are attended by assigning them high precision. The upper panels show the sonogram heard by the first bird (red lines in the lower panels, note that the timescales differ between the upper and the middle/lower graphs). In the left panel the birds cannot hear each other, while in the right panel they can. The posterior expectations for the first (red) bird are shown in red as a function of time—and the equivalent expectations for the second (blue) bird are shown in blue. Left panel: Because this bird can only hear itself, the sonogram reflects descending proprioceptive predictions based upon expectations in the higher vocal venter (HVC a premotor region, middle panel) and area X (a higher order area, lower panel), which projects to the auditory thalamus. The blue and red lines reporting expectations about underlying causes (i.e., fluctuations in amplitude and frequency) generating the birdsong are shown for the HVC and area X, in the middle and lower panels, respectively. Note that when the birds are listening, their expectations at the first level fall to zero—because they do not hear anything. However, the slower dynamics in area X can generate the song again after the end of each listening period. Right panel: here, the two birds can hear each other. In this instance, the expectations show synchrony at both the sensory and extrasensory hierarchical levels. Note that the sonogram is continuous over successive two second epochs— generated alternately by the first and second bird. The key role of precision emerges again; here, in selectively attending to sensory streams—generated by the birds—in a coordinated way that enables turn taking and communication[261,262]. This predictive coding framework provides a powerful model for describing musical communication (Figure 4).

Figure 1 - Music: from music structural constituents to perception, action, and emotion in the brain

The figure shows the constituent parts of music and their underlying brain bases as measured with established electrophysiological and neuroimaging techniques. **A)** This panel shows the melody, harmony, and rhythm elements in an excerpt from *The Reprise of Sgt. Pepper's Lonely Hearts Club Band* by the Beatles. **B)** Brain responses to music can be measured with neuroimaging methods, typically EEG/MEG and fMRI, which have different temporal resolution. The sampling for EEG/MEG is typically on the scale of 1-10 ms and for fMRI between 0.72-3 seconds. **C)** Neural markers obtained with two analysis methods for EEG/MEG data: event-related potentials (ERPs) and frequency tagging. . The perhaps most used ERPs are MMN and ERAN, which are markers of auditory expectancy violation. The MMN waveform (top panel) typically occurs around 110-250ms (here adapted from a study on melody perception[11]), while the ERAN waveform (middle panel) typically occurs around 150-200ms. The panel shows how the sources of these signals have been localized in slightly different regions of the brain (top and middle panels to the right of the waveforms, from a study on musical harmony[65]). Finally, another prominent method, the *frequency tagging* shows how the beat (here, an unaccented repeated pulse) and an imagined 3/4 meter are represented as peaks in the amplitude spectrum of the EEG (adapted from a study on musical rhythm[98]). **D)** The evidence for brain networks involved in music is shown with the key brain structures related to music perception, action, and emotion. Learning is here illustrated as the continuous update of real-time predictive brain models through Bayesian inference.

Figure 2 - Predictive Coding of Music

A) Music perception is guided by the brain's real time predictive (a.k.a., generative) model—marked with (!)— which is based on prior experience. The predictive model relies on cultural background, musical competence, the current context, and brain state including attentional and emotional state, individual traits, and innate biological factors. The brain constantly attempts to minimize prediction error at all levels of the brain hierarchy through the process of Bayesian inference. **B)** The music example shows a syncopated rhythm to which the brain may apply a 4/4-meter model. The syncopated (unexpected) note provokes a prediction error between the sensory input and the top-down predictions. This process may lead to an impulse for action, in the form of rhythmic movement, such as tapping the foot, to produce proprioceptive sensations that conform to the predictive model—and attenuate

(auditory) prediction errors that do not. The recursive arrows indicate that this process is iterated every time the rhythm repeats. Over time, this forms the basis of learning and evolving musical emotions, which in return modify action and perception.

Figure 3 - How we may experience the same musical material with different real time predictive brain models

A) There can be different interpretations of a simple ambiguous melody according to different tonalities, e.g., here C major versus A minor. The melody is compatible with both C major and A minor, and an individual's perception of harmonic context relies on top-down processes that depend on prior experience, culture, competence, context, the current state, personal traits, or an active decision to listen from a certain viewpoint e.g., major, or minor. This may lead to quite different experiences of this melody e.g., rendering the melody happy (major) or sad (minor). **B)** Similarly, a polyrhythm may equally well be heard from the point of view of a 3/4 or 4/4 as the metric predictive model. The temporal predictions in these two cases will be very different, and the rhythm can thus be experienced as a waltz (3/4) or a march (4/4) even by the same individual.

Figure 4 - Groove: the pleasurable sensation of wanting to move to music

A) The figure shows how the inverted U-shaped relationship between rhythmic predictability and the experience of groove observed in groove ratings can be modelled as the product of stimulus syncopation and the precision of the predictions relative to prediction error, i.e., the precision of our metrical expectations[126]. The U-shape implies that there is a sweet spot at which we experience the pleasurable experience of wanting to move. Hence, the experience of groove is a tradeoff between stimulus complexity (amount of syncopation) and the ability to maintain a sufficiently stable meter for moving in time with music. **B)** In The Jackson's song *Blame it on the Boogie* the rhythm of the melody quickly reaches and remains at a medium level of syncopation corresponding to the groove sweet spot. **C)** Activity in motor-, reward-, and timing-related brain structures when contrasting medium and high syncopation rhythms[132]. SMA=supplementary motor area; dPMC=dorsal premotor cortex; mOFC=medial orbitofrontal cortex; PFC=prefrontal cortex.

Figure 5 – Musical interaction.

The figure presents a model of musical communication inspired by predictive coding showing the continuous and reciprocal process of harmonizing expectations. **A)** Two improvisers (red and blue brain) may initially (marked with a yellow square) have different schematic expectations— i.e., they could experience different meters (a 3/4 and 4/4 meter) and tonalities (C-major and A-minor) while playing together. **B)** Over the course of the interaction these models may become harmonized into a shared experience (marked with a green square) of a 4/4 meter and C-major through reciprocal predictive coding mechanisms[263]. **C)** The panel shows how a simulated interaction between two improvisers may evolve over time with three different types of simultaneously occurring and interacting musical expectations: schematic, veridical, and dynamic expectations. Initially (marked with yellow, corresponding to **5A**), the schematic expectations (based on experience of meter and tonality) are quite different as illustrated in the music examples. After a while (marked with green, corresponding to **5B**), when a shared predictive meter and tonality model has been established, the schematic expectations of the two improvisers converge. The middle and bottom panels illustrate that veridical expectations (of familiar musical material) and dynamic (short term) expectations will be more but not fully harmonized after the shared schematic expectations are established. **D)** A data-based example of synchronization of dynamic metric expectations, when two individuals from the same musical background tap a simple rhythm together. The connectivity-based EEG data (recurrent phase-locking patterns (PL states))[222] highlights how the information flows differently (as indicated by the arrows) in a non-adaptive as compared to an adaptive musician differ within a right-lateralized temporoparietal brain network consisting of the right

somatosensory cortex, right precuneus, right supramarginal gyrus and right middle temporal cortex. Negative values are in blue and positive values in red.

Glossary

Pitch

Pitch is the perceptual correlate of periodicity in sounds that allows their ordering on a frequency related musical scale.

Chroma

The pitch class containing all pitches separated by an integer number of octaves. Humans perceive a similarity between notes having the same chroma.

Timbre

Timbre, also known as "tone color" or "tone quality", is the perceived sound quality of a sound, including its spectral composition and its additional noise characteristics.

Melody

Patterns of pitched sounds unfolding over time, in accordance with cultural conventions and constraints.

Harmony

The combination of multiple simultaneous pitched sounds to form a chord, and subsequent chord progressions, a fundamental building block of Western music. The rules of harmony are the hierarchically organized expectations for chord progressions.

Consonant/dissonant intervals

Psychologically, a harmonious sounding together of two or more notes, that is with an 'absence of roughness', 'relief of tonal tension' or the like are considered consonant by Western listeners. Dissonance is the antonym to consonance. Consonant intervals are produced by frequency ratios like 1: 2 (octave) 3:2, (fifth) or 4:3 (fourth) respectively Dissonances are intervals produced by frequency ratios formed from numbers greater than 4.

Harmonic cadence

A harmonic cadence is a stereotypical pattern consisting of two or more chords that concludes a phrase, section, or a piece of music. It is often used to establish a sense of tonality.

Rhythm

Here defined as the structured arrangement of successive sound events over time, a primary parameter of musical structure. Rhythm perception is based on perception of duration and grouping of these events and can be achieved even if sounds are not discrete, such as amplitude-modulated sounds.

Meter

A predictive framework governing the interpretation of regularly recurring patterns and accents in rhythm.

Tonality/Tonal Centre

In Western music, the organization of melody and harmony in a hierarchy of relations, often pointing towards a referential pitch (the tonal center/the tonic).

Atonal music

Music that lacks a tonal center

Schematic expectations

Expectations of musical events based on prior knowledge of regularities and patterns in musical sequences, such as melodies and chords.

Veridical expectations

Expectations of specific events or patterns in a familiar musical sequence—such as familiarity with a particular musical work.

Dynamic expectations

Short-lived expectations which dynamically shifts due to the ongoing musical context such as when a repeated musical phrase causes the listener to expect similar phrases as the work continues.

Statistical learning

The ability to extract statistical regularities from the world to learn about the environment.

Prediction error

A quantity used in predictive coding to denote the difference between an observation or point estimate and its predicted value. Predictive coding uses precision-weighted prediction errors to update expectations that generate predictions.

Information content

The contextual unexpectedness or surprisal associated with an event.

Entropy

In the Shannon sense, entropy is defined as the expected surprise or information content (a.k.a., self-information). In other words, it is the uncertainty or unpredictability of a random variable (e.g., an event in the future).

Precision

The inverse variance or negative entropy of a random variable. It corresponds to a second-order statistic (e.g., second order moment) of the variable's probability distribution or density. This can be contrasted with the mean or expectation that constitutes a first-order statistic (e.g., first-order moment).

Active Inference

An enactive generalization of predictive coding that casts both action and perception as minimizing surprise or prediction error (active inference is considered a corollary of the free energy principle).

Frequency tagging

A method of analyzing steady-state evoked potentials arising from stimulation or aspects of stimulation repeated at a fixed rate. An example of frequency tagging analysis is shown in Figure 1, panel C.

Syncopation

A shift of rhythmic emphasis from metrically strong accents to weak accents, a characteristic of multiple musical genres such as funk, jazz, and hiphop.

Anticipation

The subjective experience accompanying a strong expectation that a particular event will occur.

Expectation

Mathematically, an expectation is the expected value or mean of a random variable.

Prediction

The output of a model generating outcomes from their causes. In predictive coding, the prediction is generated from expected states of the world and compared with observed outcomes to form a prediction error.

MEG

Magnetoencephalography (MEG) is a neuroimaging technique that measures the magnetic fields produced by naturally occurring electrical activity in the brain.

EEG

Electroencephalography (EEG) is an electrophysiological method that measures electrical activity of the brain.

ERP

Event-related potentials (ERPs) are very small electrical voltages generated in the brain structures in response to specific events or stimuli.

MMN

The mismatch negativity (MMN) is a component of the auditory ERP recorded with EEG or MEG related to change in different sound features such as pitch, timbre, location of sound source, intensity, and rhythm. It peaks approximately 110-250 ms after change onset and is typically recorded while participants' attention is distracted from stimulus, usually by watching a silent movie or reading a book. The amplitude and latency of the MMN depends on deviation magnitude such that larger deviations in the same context yield larger and faster MMNs.

fMRI

Functional magnetic resonance imaging (fMRI) is a neuroimaging technique that images rapid changes in blood oxygenation (BOLD) levels in the brain.

Eudaimonia

In Aristotelian ethics, a life well lived or human flourishing, in affective neuroscience often described as meaningful pleasure.

References

1   Zatorre, R. J., Chen, J. L. & Penhune, V. B. When the brain plays music: auditory–motor interactions in music perception and production. *Nature reviews neuroscience* **8**, 547-558, (2007).

2   Maes, P. J., Leman, M., Palmer, C. & Wanderley, M. M. Action-based effects on music perception. *Front Psychol* **4**, (2014).

3   Koelsch, S. Brain correlates of music-evoked emotions. *Nat Rev Neurosci* **15**, 170-180, (2014).

4   Vuust, P. & Witek, M. A. Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Front Psychol* **5**, 1111, (2014).

5   Friston, K. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience* **11**, 127-138, (2010).

6   Koelsch, S., Vuust, P. & Friston, K. Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences* **23**, 63-77, (2019).

7   Meyer, L. *Emotion and Meaning in Music*. (University of Chicago Press., 1956).

8   Lerdahl, F. & Jackendoff, R. *A Generative Theory of Music*. Vol. MIT Press, Cambridge, Massachusetts (MIT Press 1999).

9   Huron, D. *Sweet Anticipation*. (The MIT Book, 2006).

10   Hansen, N. C. & Pearce, M. T. Predictive uncertainty in auditory sequence processing. *Front Psychol* **5**, (2014).

11   Vuust, P., Brattico, E., Seppanen, M., Naatanen, R. & Tervaniemi, M. The sound of music: Differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. *Neuropsychologia* **50**, 1432-1443, (2012).

12   Altenmüller, E. O. How many music centers are in the brain? *Annals of the New York Academy of Sciences* **930**, 273-280, (2001).

13   Monelle, R. *Linguistics and Semiotics in Music*. (Harwood Academic Publishers, 1992).

14   Rohrmeier, M. A. & Koelsch, S. Predictive information processing in music cognition. A critical review. *International journal of psychophysiology* **83**, 164-175, (2012).

15   Vuust, P. & Frith, C. Anticipation is the key to understanding music and the effects of music on emotion. *Behav.Brain Res.* **31**, 599-600, (2008).

16   Garrido, M. I., Sahani, M. & Dolan, R. J. Outlier responses reflect sensitivity to statistical structure in the human brain. *PLoS computational biology* **9**, e1002999, (2013).

17   Lumaca, M., Baggio, G., Brattico, E., Haumann, N. T. & Vuust, P. From random to regular: neural constraints on the emergence of isochronous rhythm during cultural transmission. *Social cognitive and affective neuroscience* **13**, 877-888, (2018).

18   Quiroga-Martinez, D. R. *et al.* Musical prediction error responses similarly reduced by predictive uncertainty in musicians and non-musicians. *European Journal of Neuroscience*, (2019).

19   Koelsch, S., Schröger, E. & Gunter, T. C. Music matters: Preattentive musicality of the human brain. *Psychophysiology* **39**, 38-48, (2002).

20   Koelsch, S., Schmidt, B.-h. & Kansok, J. Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology* **39**, 657-663, (2002).

21   Lumaca, M., Dietz, M. J., Hansen, N. C., Quiroga-Martinez, D. R. & Vuust, P. Perceptual learning of tone patterns changes the effective connectivity between Heschl's gyrus and planum temporale. *Human Brain Mapping*, (2020).

22   Lieder, F., Daunizeau, J., Garrido, M. I., Friston, K. J. & Stephan, K. E. Modelling trial-by-trial changes in the mismatch negativity. *PLoS computational biology* **9**, e1002911, (2013).

23   Wacongne, C., Changeux, J. P. & Dehaene, S. A neuronal model of predictive coding accounting for the mismatch negativity. *J Neurosci* **32**, 3665-3678, (2012).

24   Kiebel, S. J., Garrido, M. I. & Friston, K. J. Dynamic causal modelling of evoked responses: the role of intrinsic connections. *Neuroimage* **36**, 332-345, (2007).

25   Feldman, H. & Friston, K. J. Attention, uncertainty, and free-energy. *Frontiers in human neuroscience* **4**, 215, (2010).

26   Cheung, V. K. M. *et al.* Uncertainty and Surprise Jointly Predict Musical Pleasure and Amygdala, Hippocampus, and Auditory Cortex Activity. *Curr Biol* **29**, 4084-4092 e4084, (2019).

27   McDermott, J. H. & Oxenham, A. J. Music perception, pitch, and the auditory system. *Current opinion in neurobiology* **18**, 452-463, (2008).

28   Thoret, E., Caramiaux, B., Depalle, P. & McAdams, S. Learning metrics on spectrotemporal modulations reveals the perception of musical instrument timbre. *Nat Hum Behav*, (2020).

29   Siedenburg, K. & McAdams, S. Four Distinctions for the Auditory "Wastebasket" of Timbre. *Front Psychol* **8**, 1747, (2017).

30   Bendor, D. & Wang, X. The neuronal representation of pitch in primate auditory cortex. *Nature* **436**, 1161-1165, (2005).

31   Zatorre, R. J. Pitch perception of complex tones and human temporal-lobe functiona. *Journal of the Acoustical Society of America* **84**, 566-572, (1988).

32   Warren, J. D., Uppenkamp, S., Patterson, R. D. & Griffiths, T. D. Separating pitch chroma and pitch height in the human brain. *Proc Natl Acad Sci U S A* **100**, 10038-10042, (2003).

33   Rauschecker, J. P. & Scott, S. K. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* **12**, 718-724, (2009).

34   Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R. & Rauschecker, J. P. Brain activation during anticipation of sound sequences. *J Neurosci* **29**, 2477-2485, (2009).

35   Houde, J. F. & Chang, E. F. The cortical computations underlying feedback control in vocal production. *Current opinion in neurobiology* **33**, 174-181, (2015).

36   Lee, Y. S., Janata, P., Frost, C., Hanke, M. & Granger, R. Investigation of melodic contour processing in the brain using multivariate pattern-based fMRI. *Neuroimage* **57**, 293-300, (2011).

37   Janata, P. *et al.* The cortical topography of tonal structures underlying Western music. *Science* **298**, 2167-2170, (2002).

38   Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical learning by 8-month-old infants. *Science* **274**, 1926-1928, (1996).

39   Saffran, J. R., Johnson, E. K., Aslin, R. N. & Newport, E. L. Statistical learning of tone sequences by human infants and adults. *Cognition* **70**, 27-52, (1999).

40   Krumhansl, C. L. Perceptual Structures for Tonal Music. *Music Perception* **1**, 28-62, (1983).

41   Sears, D. R. W., Pearce, M. T., Caplin, W. E. & McAdams, S. Simulating melodic and harmonic expectations for tonal cadences using probabilistic models. *Journal of New Music Research* **47**, 29-52, (2018).

42   Näätänen, R., Gaillard, A. W. & Mäntysalo, S. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.(Amst)* **42**, 313-329, (1978).

43   Näätänen, R., Paavilainen, P., Rinne, T. & Alho, K. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology* **118**, 2544-2590, (2007).

44   Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C. & Vuust, P. The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences* **20**, 188-196, (2010).

45   Tervaniemi, M. *et al.* Top-down modulation of auditory processing: effects of sound context, musical expertise and attentional focus. *Eur J Neurosci* **30**, 1636-1642, (2009).

46   Burunat, I. *et al.* The reliability of continuous brain responses during naturalistic listening to music. *Neuroimage* **124**, 224-231, (2016).

47   Burunat, I. *et al.* Action in Perception: Prominent Visuo-Motor Functional Symmetry in Musicians during Music Listening. *PLoS One* **10**, e0138238, (2015).

48   Alluri, V. *et al.* Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *NeuroImage* **59**, 3677-3689, (2012).

49   Halpern, A. R. & Zatorre, R. J. When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex* **9**, 697-704, (1999).

50   Herholz, S. C., Halpern, A. R. & Zatorre, R. J. Neuronal correlates of perception, imagery, and memory for familiar tunes. *J Cogn Neurosci* **24**, 1382-1397, (2012).

51   Pallesen, K. J. *et al.* Emotion processing of major, minor, and dissonant chords: a functional magnetic resonance imaging study. *Annals of the New York Academy of Sciences* **1060**, 450-453, (2005).

52   McPherson, M. J. *et al.* Perceptual fusion of musical notes by native Amazonians suggests universal representations of musical intervals. *Nat Commun* **11**, 2786, (2020).

53   Hermann, L. Helmholtz, On the Sensations of Tone as a Physiological Basis for the Theory of Music. *Trans. Alexander J. Ellis (New York: Dover, 1954)* **7**, (1954).

54   Vassilakis, P. N. & Kendall, R. A. in *Human Vision and Electronic Imaging XV.* 75270O (International Society for Optics and Photonics).

55   Plomp, R. & Levelt, W. J. M. Tonal consonance and critical bandwidth. *The journal of the Acoustical Society of America* **38**, 548-560, (1965).

56   Mehr, S. A. *et al.* Universality and diversity in human song. *Science* **366**, (2019).

57   Patel, A. D., Gibson, E., Ratner, J., Besson, M. & Holcomb, P. J. Processing syntactic relations in language and music: An event-related potential study. *Journal of cognitive neuroscience* **10**, 717-733, (1998).

58   Janata, P. The neural architecture of music-evoked autobiographical memories. *Cereb Cortex* **19**, 2579-2594, (2009).

59   Janata, P. *et al.* The cortical topography of tonal structures underlying Western music. *science* **298**, 2167-2170, (2002).

60   Maess, B., Koelsch, S., Gunter, T. C. & Friederici, A. D. Musical syntax is processed in Broca's area: an MEG study. *Nat.Neurosci.* **4**, 540-545, (2001).

61   Koelsch, S. *et al.* Differentiating ERAN and MMN: an ERP study. *Neuroreport* **12**, 1385-1389, (2001).

62   Loui, P., Grent-'t-Jong, T., Torpey, D. & Woldorff, M. Effects of attention on the neural processing of harmonic syntax in Western music. *Cognitive Brain Research* **25**, 678-687, (2005).

63   Koelsch, S., Fritz, T., Schulze, K., Alsop, D. & Schlaug, G. Adults and children processing music: an fMRI study. *Neuroimage* **25**, 1068-1076, (2005).

64   Tillmann, B., Janata, P. & Bharucha, J. J. Activation of the inferior frontal cortex in musical priming. *Annals of the New York Academy of Sciences* **999**, 209-211, (2003).

65   Garza-Villarreal, E. A., Brattico, E., Leino, S., Ostergaard, L. & Vuust, P. Distinct neural responses to chord violations: A multiple source analysis study. *Brain Res*, (2011).

66   Leino, S., Brattico, E., Tervaniemi, M. & Vuust, P. Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain Research* **1142**, 169-177, (2007).

67   Sammler, D. *et al.* Co-localizing linguistic and musical syntax with intracranial EEG. *Neuroimage* **64**, 134-146, (2013).

68   Loui, P., Wessel, D. L. & Hudson Kam, C. L. Humans rapidly learn grammatical structure in a new musical scale. *Music Perception* **27**, 377-388, (2010).

69   Loui, P., Wu, E. H., Wessel, D. L. & Knight, R. T. A generalized mechanism for perception of pitch patterns. *Journal of Neuroscience* **29**, 454-459, (2009).

70   Cheung, V. K. M., Meyer, L., Friederici, A. D. & Koelsch, S. The right inferior frontal gyrus processes nested non-local dependencies in music. *Scientific reports* **8**, 3822, (2018).

71   Haueisen, J. & Knosche, T. R. Involuntary motor activity in pianists evoked by music perception. *J Cogn Neurosci* **13**, 786-792, (2001).

72   Bangert, M. *et al.* Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage* **30**, 917-926, (2006).

73   Baumann, S. *et al.* A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Res* **1161**, 65-78, (2007).

74   Lahav, A., Saltzman, E. & Schlaug, G. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience* **27**, 308-314, (2007).

75   Bianco, R. *et al.* Neural networks for harmonic structure in music perception and action. *NeuroImage* **142**, 454-464, (2016).

76   McDermott, J. H., Schultz, A. F., Undurraga, E. A. & Godoy, R. A. Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature* **535**, 547-550, (2016).

77   Fritz, T. *et al.* Universal recognition of three basic emotions in music. *Curr Biol* **19**, 573-576, (2009).

78   Eerola, T., Vuoskoski, J. K., Peltola, H.-R., Putkinen, V. & Schäfer, K. An integrative review of the enjoyment of sadness associated with music. *Physics of life reviews* **25**, 100-121, (2018).

79   Huron, D. M. D. The Harmonic Minor Scale Provides an Optimum Way of Reducing Average Melodic Interval Size, Consistent with Sad Affect Cues. *Empirical Musicology Review* **7**, 15, (2012).

80   Huron, D. A comparison of average pitch height and interval size in major-and minor-key themes: Evidence consistent with affect-related pitch prosody. (2008).

81   Juslin, P. N. & Laukka, P. Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological bulletin* **129**, 770, (2003).

82   London, J. *Hearing in time: Psychological aspects of musical meter*. (Oxford University Press, 2012).

83   Honing, H. Without it no music: beat induction as a fundamental musical trait. *Ann N Y Acad Sci* **1252**, 85-91, (2012).

84   Hickok, G., Farahbod, H. & Saberi, K. The rhythm of perception: entrainment to acoustic rhythms induces subsequent perceptual oscillation. *Psychological science* **26**, 1006-1013, (2015).

85   Yabe, H., Tervaniemi, M., Reinikainen, K. & Näätänen, R. Temporal window of integration revealed by MMN to sound omission. *Neuroreport* **8**, 1971-1974, (1997).

86   Andreou, L.-V., Griffiths, T. D. & Chait, M. Sensitivity to the temporal structure of rapid sound sequences—an MEG study. *Neuroimage* **110**, 194-204, (2015).

87   Jongsma, M. L., Meeuwissen, E., Vos, P. G. & Maes, R. Rhythm perception: Speeding up or slowing down affects different subcomponents of the ERP P3 complex. *Biological psychology* **75**, 219-228, (2007).

88   Graber, E. & Fujioka, T. Endogenous Expectations for Sequence Continuation after Auditory Beat Accelerations and Decelerations Revealed by P3a and Induced Beta-Band Responses. *Neuroscience* **413**, 11-21, (2019).

89   Brochard, R., Abecasis, D., Potter, D., Ragot, R. & Drake, C. The "ticktock" of our internal clock: direct brain evidence of subjective accents in isochronous sequences. *Psychological Science* **14**, 362-366, (2003).

90   Lerdahl, F. & Jackendoff, R. An overview of hierarchical structure in music. *Music Perception* **1**, 229-252, (1983).

91   Large, E. W. & Kolen, J. F. Resonance and the perception of musical meter. *Connection Science* **6**   177-208, (1994).

92   Large, E. W. & Jones, M. R. The dynamics of attending: How people track time-varying events. *Psychological Review* **106**, 119-159, (1999).

93   Cutietta, R. A. & Booth, G. D. The influence of metre, mode, interval type and contour in repeated melodic free-recall. *Psychology of Music* **24**, 222-236, (1996).

94   Smith, K. C. & Cuddy, L. L. Effects of Metric and Harmonic Rhythm on the Detection of Pitch Alterations in Melodic Sequences. *Journal of Experimental Psychology: Human Perception and Performance* **15**, 457-471, (1989).

95   Palmer, C. & Krumhansl, C. L. Mental Representations for Musical Meter. *Journal of Experimental Psychology: Human Perception and Performance* **16**, 728-741, (1990).

96   Einarson, K. M. & Trainor, L. J. Hearing the beat: Young children's perceptual sensitivity to beat alignment varies according to metric structure. *Music Perception* **34**, 56-70, (2016).

97   Large, E. W., Herrera, J. A. & Velasco, M. J. Neural Networks for Beat Perception in Musical Rhythm. *Front Syst Neurosci* **9**, 159, (2015).

98   Nozaradan, S., Peretz, I., Missal, M. & Mouraux, A. Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience* **31**, 10234-10240, (2011).

99   Nozaradan, S., Peretz, I. & Mouraux, A. Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience* **32**, 17572-17581, (2012).

100   Nozaradan, S., Schonwiesner, M., Keller, P. E., Lenc, T. & Lehmann, A. Neural bases of rhythmic entrainment in humans: critical transformation between cortical and lower-level representations of auditory rhythm. *Eur J Neurosci* **47**, 321-332, (2018).

101   Lenc, T., Keller, P. E., Varlet, M. & Nozaradan, S. Neural and Behavioral Evidence for Frequency-Selective Context Effects in Rhythm Processing in Humans. *Cerebral Cortex Communications*, (2020).

102   Jacoby, N. & McDermott, J. H. Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction. *Curr Biol* **27**, 359-370, (2017).

103   Hannon, E. E. & Trehub, S. E. Metrical categories in infancy and adulthood. *Psychol Sci* **16**, 48-55, (2005).

104   Hannon, E. E. & Trehub, S. E. Tuning in to musical rhythms: infants learn more readily than adults. *Proc Natl Acad Sci U S A* **102**, 12639-12643, (2005).

105   Vuust, P. *et al.* To musicians, the message is in the meter pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *Neuroimage* **24**, 560-564, (2005).

106   Grahn, J. A. & Brett, M. Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience* **19**, 893-906, (2007).

107   Toiviainen, P., Burunat, I., Brattico, E., Vuust, P. & Alluri, V. The chronnectome of musical beat. *Neuroimage*, 116191, (2019).

108   Chen, J. L., Penhune, V. B. & Zatorre, R. J. Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *J Cogn Neurosci* **20**, 226-239, (2008).

109   Levitin, D. J., Grahn, J. A. & London, J. The Psychology of Music: Rhythm and Movement. *Annu Rev Psychol* **69**, 51-75, (2018).

110   Winkler, I., Haden, G. P., Ladinig, O., Sziller, I. & Honing, H. Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 2468-2471, (2009).

111   Phillips-Silver, J. & Trainor, L. J. Feeling the beat: Movement influences infant rhythm perception. *Science* **308**, 1430-1430, (2005).

112   Cirelli, L. K., Trehub, S. E. & Trainor, L. J. Rhythm and melody as social signals for infants. *Ann N Y Acad Sci*, (2018).

113   Repp, B. H. Sensorimotor synchronization: a review of the tapping literature. *Psychon Bull Rev* **12**, 969-992, (2005).

114   Repp, B. H. & Su, Y. H. Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin and Review* **20**, 403-452, (2013).

115   Zarco, W., Merchant, H., Prado, L. & Mendez, J. C. Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology* **102**, 3191-3202, (2009).

116   Honing, H., Bouwer, F. L., Prado, L. & Merchant, H. Rhesus Monkeys (Macaca mulatta) sense isochrony in rhythm, but not the beat: Additional support for the gradual audiomotor evolution hypothesis. *Frontiers in neuroscience* **12**, (2018).

117   Hattori, Y. & Tomonaga, M. Rhythmic swaying induced by sound in chimpanzees (Pan troglodytes). *Proceedings of the National Academy of Sciences of the United States of America* **117**, 936-942, (2020).

118   Janata, P., Tomic, S. T. & Haberman, J. M. Sensorimotor coupling in music and the psychology of the groove. *Journal of experimental psychology: general* **141**, 54, (2012).

119   Danielsen, A. *Presence and Pleasure. The funk grooves of James Brown and Parliament*. (Wesleyan University Press, 2006).

120   Madison, G., Gouyon, F., Ullen, F. & Hornstrom, K. Modeling the tendency for music to induce movement in humans: first correlations with low-level audio descriptors across music genres. *J Exp Psychol Hum Percept Perform* **37**, 1578-1594, (2011).

121   Stupacher, J., Hove, M. J., Novembre, G., Schutz-Bosbach, S. & Keller, P. E. Musical groove modulates motor cortex excitability: a TMS investigation. *Brain Cogn* **82**, 127-136, (2013).

122   Witek, M. A. *et al.* A Critical Cross-cultural Study of Sensorimotor and Groove Responses to Syncopation Among Ghanaian and American University Students and Staff. *Music Perception* **37**, 278-297, (2020).

123   Friston, K., Mattout, J. & Kilner, J. Action understanding and active inference. *Biological cybernetics* **104**, 137-160, (2011).

124   Longuet-Higgins, H. C. & Lee, C. S. The rhythmic interpretation of monophonic music. *Music Perception* **1**, 18, (1984).

125   Sioros, G., Miron, M., Davies, M., Gouyon, F. & Madison, G. Syncopation creates the sensation of groove in synthesized music examples. *Front Psychol* **5**, 1036, (2014).

126   Vuust, P., Dietz, M. J., Witek, M. & Kringelbach, M. L. Now you hear it: a predictive coding model for understanding rhythmic incongruity. *Annals of the New York Academy of Sciences*, (2018).

127   Witek, M. A., Clarke, E. F., Wallentin, M., Kringelbach, M. L. & Vuust, P. Syncopation, body-movement and pleasure in groove music. *PLoS One* **9**, e94446, (2014).

128   Kowalewski, D. A., Kratzer, T. M. & Friedman, R. S. Social Music: Investigating the Link Between Personal Liking and Perceived Groove. *Music Perception* **37**, 339-346, (2020).

129   Witek, M. A. G. *et al.* The brain dynamics of musical groove: whole-brain modelling of effective connectivity reveals increased metastability of reward and motor networks. *Nature Communications* **In Review**, (2019).

130   Bowling, D. L., Ancochea, P. G., Hove, M. J. & Tecumseh Fitch, W. Pupillometry of groove: Evidence for noradrenergic arousal in the link between music and movement. *Frontiers in neuroscience* **13**, (2019).

131   Matthews, T. E., Witek, M. A. G., Heggli, O. A., Penhune, V. B. & Vuust, P. The sensation of groove is affected by the interaction of rhythmic and harmonic complexity. *PLoS One* **14**, e0204539, (2019).

132   Matthews, T. E., Witek, M. A., Lund, T., Vuust, P. & Penhune, V. B. The sensation of groove engages motor and reward networks. *NeuroImage*, 116768, (2020).

133   Vaquero, L., Ramos-Escobar, N., François, C., Penhune, V. & Rodríguez-Fornells, A. White-matter structural connectivity predicts short-term melody and rhythm learning in non-musicians. *NeuroImage* **181**, 252-262, (2018).

134   Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E. & Evans, A. C. Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience* **8**, 29-46, (1996).

135   London, J., Polak, R. & Jacoby, N. Rhythm histograms and musical meter: A corpus study of Malian percussion music. *Psychon Bull Rev* **24**, 474-480, (2017).

136   Huron, D. Is music an evolutionary adaptation? *Ann.N.Y.Acad.Sci.* **930**, 43-61, (2001).

137   Koelsch, S. Towards a neural basis of music-evoked emotions. *Trends Cogn Sci* **14**, 131-137, (2010).

138   Eerola, T. & Vuoskoski, J. K. A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music* **39**, 18-49, (2010).

139   Lonsdale, A. J. & North, A. C. Why do we listen to music? A uses and gratifications analysis. *British Journal of Psychology* **102**, 108-134, (2011).

140   Juslin, P. N. & Laukka, P. Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening. *Journal of New Music Research* **33**, 217-238, (2004).

141   Huron, D. Why is sad music pleasurable? A possible role for prolactin. *Musicae Scientiae* **15**, 146-158, (2011).

142   Brattico, E. *et al.* It's Sad but I Like It: The Neural Dissociation Between Musical Emotions and Liking in Experts and Laypersons. *Frontiers in human neuroscience* **9**, 676, (2015).

143   Sachs, M. E., Damasio, A. & Habibi, A. Unique personality profiles predict when and why sad music is enjoyed. *Psychology of Music*, (2020).

144   Sachs, M. E., Habibi, A., Damasio, A. & Kaplan, J. T. Dynamic intersubject neural synchronization reflects affective responses to sad music. *NeuroImage* **218**, (2020).

145   Eerola, T., Vuoskoski, J. K., Peltola, H. R., Putkinen, V. & Schafer, K. An integrative review of the enjoyment of sadness associated with music. *Physics of life reviews* **25**, 100-121, (2018).

146   Juslin, P. N. & Vastfjall, D. Emotional responses to music: the need to consider underlying mechanisms. *Behav.Brain Sci* **31**, 559-575, (2008).

147   Rickard, N. S. Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of music* **32**, 371-388, (2004).

148   Cowen, A. S., Fang, X., Sauter, D. & Keltner, D. What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proc Natl Acad Sci U S A* **117**, 1924-1934, (2020).

149   Argstatter, H. Perception of basic emotions in music: Culture-specific or multicultural? *Psychology of Music* **44**, 674-690, (2016).

150   Stevens, C. J. Music Perception and Cognition: A Review of Recent Cross-Cultural Research. *Topics in Cognitive Science* **4**, 653-667, (2012).

151   Pearce, M. *Cultural Distance: A Computational Approach to Exploring Cultural Influences on Music Cognition*.in *Oxford Handbook of Music and the Brain*   (ed Oxford University Press) 31 (Oxford University Press, 2018).

152   van der Weij, B., Pearce, M. T. & Honing, H. A Probabilistic Model of Meter Perception: Simulating Enculturation. *Front Psychol* **8**, (2017).

153   Kringelbach, M. L. & Berridge, K. C. Towards a functional neuroanatomy of pleasure and happiness. *Trends Cogn Sci.* **13**, 479-487, (2009).

154   Blood, A. J. & Zatorre, R. J. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc.Natl.Acad.Sci.U.S.A* **98**, 11818-11823, (2001).

155   Salimpoor, V. N. & Zatorre, R. J. Complex cognitive functions underlie aesthetic emotions: comment on "From everyday emotions to aesthetic emotions: towards a unified theory of musical emotions" by Patrik N. Juslin. *Physics of life reviews* **10**, 279-280, (2013).

156   Salimpoor, V. N. *et al.* Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science* **340**, 216-219, (2013).

157   Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A. & Zatorre, R. J. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat Neurosci* **14**, 257-262, (2011).

158   Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R. & Zatorre, R. J. The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS One* **4**, e7487, (2009).

159   Mas-Herrero, E., Zatorre, R. J., Rodriguez-Fornells, A. & Marco-Pallares, J. Dissociation between musical and monetary reward responses in specific musical anhedonia. *Curr Biol* **24**, 699-704, (2014).

160   Martinez-Molina, N., Mas-Herrero, E., Rodriguez-Fornells, A., Zatorre, R. J. & Marco-Pallares, J. Neural correlates of specific musical anhedonia. *Proc Natl Acad Sci U S A* **113**, E7337-E7345, (2016).

161   Gebauer, L. K., M.L.; Vuust, P. Musical Pleasure Cycles: The role of anticipation and dopamine. *Psychomusicology* **22**, 16, (2012).

162   Shany, O. *et al.* Surprise-related activation in the nucleus accumbens interacts with music-induced pleasantness. *Social cognitive and affective neuroscience* **14**, 459-470, (2019).

163 Gold, B. P., Pearce, M. T., Mas-Herrero, E., Dagher, A. & Zatorre, R. J. Predictability and Uncertainty in the Pleasure of Music: A Reward for Learning? *The Journal of neuroscience : the official journal of the Society for Neuroscience* **39**, 9397-9409, (2019).

164 Swaminathan, S. & Schellenberg, E. G. Current emotion research in music psychology. *Emotion Review* **7**, 189-197, (2015).

165 Madison, G. & Schiölde, G. Repeated Listening Increases the Liking for Music Regardless of Its Complexity: Implications for the Appreciation and Aesthetics of Music. *Frontiers in neuroscience* **11**, (2017).

166 Corrigall, K. A. & Schellenberg, E. G. *Liking music: Genres, contextual factors, and individual differences*.in *Art, Aesthetics, and the Brain* (Oxford University Press, 2015).

167 Zentner, A. Measuring the effect of file sharing on music purchases. *Journal of Law and Economics* **49**, 63-90, (2006).

168 Rentfrow, P. J. & Gosling, S. D. The do re mi's of everyday life: the structure and personality correlates of music preferences. *J.Pers.Soc.Psychol.* **84**, 1236-1256, (2003).

169 Vuust, P. *et al.* Personality influences career choice: Sensation seeking in professional musicians. *Music Education Research* **12**, 219-230, (2010).

170 Rohrmeier, M. & Rebuschat, P. Implicit Learning and Acquisition of Music. *Topics in Cognitive Science* **4**, 525-553, (2012).

171 Münthe, T. F., Altenmüller, E. & Jäncke, L. The musician's brain as a model of Neuroplasticity. *Nat.Rev.Neurosci.* **3**, 1-6, (2002).

172 Habibi, A. *et al.* Childhood music training induces change in micro and macroscopic brain structure: Results from a longitudinal study. *Cerebral Cortex* **28**, 4336-4347, (2018).

173 Schlaug, G., Jancke, L., Huang, Y., Staiger, J. F. & Steinmetz, H. Increased corpus callosum size in musicians. *Neuropsychologia* **33**, 1047-1055, (1995).

174 Baer, L. H. *et al.* Regional cerebellar volumes are related to early musical training and finger tapping performance. *Neuroimage* **109**, 130-139, (2015).

175 Kleber, B. *et al.* Voxel-based morphometry in opera singers: Increased gray-matter volume in right somatosensory and auditory cortices. *Neuroimage* **133**, 477-483, (2016).

176 Gaser, C. & Schlaug, G. Brain structures differ between musicians and non-musicians. *Journal of Neuroscience* **23**, 9240-9245, (2003).

177 Sluming, V. *et al.* Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *Neuroimage.* **17**, 1613-1622, (2002).

178 Palomar-García, M.-Á., Zatorre, R. J., Ventura-Campos, N., Bueichekú, E. & Ávila, C. Modulation of functional connectivity in auditory–motor networks in musicians compared with nonmusicians. *Cerebral Cortex* **27**, 2768-2778, (2017).

179 Schneider, P. *et al.* Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature neuroscience* **5**, 688-694, (2002).

180   Bengtsson, S. L. *et al.* Extensive piano practicing has regionally specific effects on white matter development. *Nature neuroscience* **8**, 1148-1150, (2005).

181   Zamorano, A. M., Cifre, I., Montoya, P., Riquelme, I. & Kleber, B. Insula-based networks in professional musicians: Evidence for increased functional connectivity during resting state fMRI. *Human brain mapping* **38**, 4834-4849, (2017).

182   Kraus, N. & Chandrasekaran, B. Music training for the development of auditory skills. *Nat Rev Neurosci* **11**, 599-605, (2010).

183   Koelsch, S., Schröger, E. & Tervaniemi, M. Superior pre-attentive auditory processing in musicians. *Neuroreport* **10**, 1309-1313, (1999).

184   Münte, T. F., Kohlmetz, C., Nager, W. & Altenmüller, E. Neuroperception. Superior auditory spatial tuning in conductors. *Nature* **409**, 580, (2001).

185   Seppänen, M., Brattico, E. & Tervaniemi, M. Practice strategies of musicians modulate neural processing and the learning of sound-patterns. *Neurobiology of Learning and Memory* **87**, 236-247, (2007).

186   Guillot, A. *et al.* Functional neuroanatomical networks associated with expertise in motor imagery. *Neuroimage* **41**, 1471-1483, (2008).

187   Bianco, R., Novembre, G., Keller, P. E., Villringer, A. & Sammler, D. Musical genre-dependent behavioural and EEG signatures of action planning. A comparison between classical and jazz pianists. *Neuroimage* **169**, 383-394, (2018).

188   Vuust, P., Brattico, E., Seppänen, M., Näätänen, R. & Tervaniemi, M. Practiced musical style shapes auditory skills. *Annals of the New York Academy of Sciences* **1252**, 139-146, (2012).

189   Bangert, M. & Altenmüller, E. O. Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC neuroscience* **4**, (2003).

190   Li, Q. *et al.* Musical training induces functional and structural auditory-motor network plasticity in young adults. *Human Brain Mapping* **39**, 2098-2110, (2018).

191   Herholz, S. C., Coffey, E. B. J., Pantev, C. & Zatorre, R. J. Dissociation of neural networks for predisposition and for training-related plasticity in auditory-motor learning. *Cerebral Cortex* **26**, 3125-3134, (2016).

192   Putkinen, V., Tervaniemi, M. & Huotilainen, M. Musical playschool activities are linked to faster auditory development during preschool-age: a longitudinal ERP study. *Scientific reports* **9**, 11310-11310, (2019).

193   Putkinen, V., Tervaniemi, M., Saarikivi, K., Ojala, P. & Huotilainen, M. Enhanced development of auditory change detection in musically trained school-aged children: a longitudinal event-related potential study. *Developmental science* **17**, 282-297, (2014).

194   Jentschke, S. & Koelsch, S. Musical training modulates the development of syntax processing in children. *Neuroimage* **47**, 735-744, (2009).

195   Chobert, J., François, C., Velay, J. L. & Besson, M. Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex* **24**, 956-967, (2014).

196   Moreno, S. *et al.* Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex* **19**, 712-723, (2009).

197   Putkinen, V., Huotilainen, M. & Tervaniemi, M. Neural Encoding of Pitch Direction Is Enhanced in Musically Trained Children and Is Related to Reading Skills. *Front Psychol* **10**, 1475, (2019).

198   Wong, P. C., Skoe, E., Russo, N. M., Dees, T. & Kraus, N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* **10**, 420-422, (2007).

199   Virtala, P. & Partanen, E. Can very early music interventions promote at-risk infants' development? *Annals of the New York Academy of Sciences* **1423**, 92-101, (2018).

200   Flaugnacco, E. *et al.* Music Training Increases Phonological Awareness and Reading Skills in Developmental Dyslexia: A Randomized Control Trial. *PLoS One* **10**, e0138715, (2015).

201   Fiveash, A. *et al.* A stimulus-brain coupling analysis of regular and irregular rhythms in adults with dyslexia and controls. *Brain Cogn* **140**, 105531, (2020).

202   Schellenberg, E. G. Correlation = Causation? Music Training, Psychology, and Neuroscience. *Psychology of Aesthetics, Creativity, and the Arts*, (2019).

203   Sala, G. & Gobet, F. Cognitive and academic benefits of music training with children: A multilevel meta-analysis. *Memory & Cognition*, (2020).

204   Saffran, J. R. Musical learning and language development. *Annals of the New York Academy of Sciences* **999**, 397-401, (2003).

205   Friston, K. The free-energy principle: a rough guide to the brain? *Trends Cogn Sci* **13**, 293-301, (2009).

206   Pearce, M. T. Statistical learning and probabilistic prediction in music cognition: mechanisms of stylistic enculturation. *Ann N Y Acad Sci*, (2018).

207   Novembre, G., Knoblich, G., Dunne, L. & Keller, P. E. Interpersonal synchrony enhanced through 20 Hz phase-coupled dual brain stimulation. *Social cognitive and affective neuroscience* **12**, 662-670, (2017).

208   Konvalinka, I. *et al.* Frontal alpha oscillations distinguish leaders from followers: Multivariate decoding of mutually interacting brains. *Neuroimage* **94C**, 79-88, (2014).

209   Novembre, G., Mitsopoulos, Z. & Keller, P. E. Empathic perspective taking promotes interpersonal coordination through music. *Scientific reports* **9**, 12255, (2019).

210   Wolpert, D. M., Ghahramani, Z. & Jordan, M. I. An internal model for sensorimotor integration. *Science* **269**, 1880-1882, (1995).

211   Sebanz, N. & Knoblich, G. Prediction in Joint Action: What, When, and Where. *Topics in Cognitive Science* **1**, 353-367, (2009).

212   Friston, K. J. & Frith, C. D. Active inference, communication and hermeneutics. *Cortex* **68**, 129-143, (2015).

213   Wing, A. M. & Kristofferson, A. B. Response delays and the timing of discrete motor responses. *Perception & Psychophysics* **14**, 5-12, (1973).

214   Repp, B. H. & Keller, P. E. Sensorimotor synchronization with adaptively timed sequences. *Human Movement Science* **27**, 423-456, (2008).

215   Vorberg, D. & Schulze, H.-H. Linear Phase-Correction in Synchronization: Predictions, Parameter Estimation, and Simulations. *Journal of Mathematical Psychology* **46**, 56-87, (2002).

216   Novembre, G., Sammler, D. & Keller, P. E. Neural alpha oscillations index the balance between self-other integration and segregation in real-time joint action. *Neuropsychologia* **89**, 414-425, (2016).

217   Keller, P. E., Knoblich, G. & Repp, B. H. Pianists duet better when they play with themselves: on the possible role of action simulation in synchronization. *Conscious Cogn* **16**, 102-111, (2007).

218   Fairhurst, M. T., Janata, P. & Keller, P. E. Leading the follower: an fMRI investigation of dynamic cooperativity and leader-follower strategies in synchronization with an adaptive virtual partner. *Neuroimage* **84**, 688-697, (2014).

219   Konvalinka, I., Vuust, P., Roepstorff, A. & Frith, C. D. Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology*, (2010).

220   Heggli, O. A., Konvalinka, I., Kringelbach, M. L. & Vuust, P. Musical interaction is influenced by underlying predictive models and musical expertise. *Scientific reports* **9**, 1-13, (2019).

221   Heggli, O. A., Cabral, J., Konvalinka, I., Vuust, P. & Kringelbach, M. L. A Kuramoto model of self-other integration across interpersonal synchronization strategies. *PLoS computational biology* **15**, (2019).

222   Heggli, O. A. *et al.* Transient brain networks underlying interpersonal strategies during synchronized action. *Social cognitive and affective neuroscience*, (2020).

223   Patel, A. D. *Music, Language, and the Brain*. (Oxford University Press, 2006).

224   Molnar-Szakacs, I. & Overy, K. Music and mirror neurons: from motion to'e'motion. *Social cognitive and affective neuroscience* **1**, 235-241, (2006).

225   Beaty, R. E., Benedek, M., Silvia, P. J. & Schacter, D. L. Creative Cognition and Brain Network Dynamics. *Trends in Cognitive Sciences* **20**, 87-95, (2016).

226   Limb, C. J. & Braun, A. R. Neural substrates of spontaneous musical performance: an FMRI study of jazz improvisation. *PLoS.ONE.* **3**, e1679, (2008).

227   Liu, S. *et al.* Neural correlates of lyrical improvisation: an FMRI study of freestyle rap. *Scientific reports* **2**, 834, (2012).

228   Rosen, D. S. *et al.* Dual-process contributions to creativity in jazz improvisations: An SPM-EEG study. *Neuroimage* **213**, 116632, (2020).

229   Boasen, J., Takeshita, Y., Kuriki, S. & Yokosawa, K. Spectral-Spatial Differentiation of Brain Activity During Mental Imagery of Improvisational Music Performance Using MEG. *Frontiers in human neuroscience* **12**, 156, (2018).

230   Berkowitz, A. L. & Ansari, D. Generation of novel motor sequences: the neural correlates of musical improvisation. *Neuroimage* **41**, 535-543, (2008).

231   Loui, P. Rapid and flexible creativity in musical improvisation: review and a model. *Ann NY Acad Sci* **1423**, 138-145, (2018).

232   Beaty, R. E. The neuroscience of musical improvisation. *Neuroscience & Biobehavioral Reviews* **51**, 108-117, (2015).

233   Vuust, P. & Kringelbach, M. L. *Music Improvisation: A Challenge for Empirical Research*.in *Routledge Companion To Music Cognition*   (Routledge, 2017).

234   Norgaard, M. Descriptions of improvisational thinking by artist-level jazz musicians. *Journal of Research in Music Education* **59**, 109-127, (2011).

235   Deco, G. & Kringelbach, M. L. Hierarchy of Information Processing in the Brain: A Novel 'Intrinsic Ignition' Framework. *Neuron* **94**, 961-968, (2017).

236   Pinho, A. L., de Manzano, O., Fransson, P., Eriksson, H. & Ullen, F. Connecting to create: expertise in musical improvisation is associated with increased functional connectivity between premotor and prefrontal areas. *J Neurosci* **34**, 6156-6163, (2014).

237   Pinho, A. L., Ullen, F., Castelo-Branco, M., Fransson, P. & de Manzano, O. Addressing a Paradox: Dual Strategies for Creative Performance in Introspective and Extrospective Networks. *Cereb Cortex* **26**, 3052-3063, (2016).

238   de Manzano, O. & Ullen, F. Activation and connectivity patterns of the presupplementary and dorsal premotor areas during free improvisation of melodies and rhythms. *Neuroimage* **63**, 272-280, (2012).

239   Beaty, R. E. *et al.* Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 1087-1092, (2018).

240   Daikoku, T. Entropy, Uncertainty, and the Depth of Implicit Knowledge on Musical Creativity: Computational Study of Improvisation in Melody and Rhythm. *Front Comput Neurosci* **12**, 97, (2018).

241   Belden, A. *et al.* Improvising at rest: Differentiating jazz and classical music training with resting state functional connectivity. *Neuroimage* **207**, 116384, (2020).

242   Arkin, C., Przysinda, E., Pfeifer, C. W., Zeng, T. & Loui, P. Gray Matter Correlates of Creativity in Musical Improvisation. *Frontiers in human neuroscience* **13**, 169, (2019).

243   Bashwiner, D. M., Wertz, C. J., Flores, R. A. & Jung, R. E. Musical Creativity "Revealed" in Brain Structure: Interplay between Motor, Default Mode, and Limbic Networks. *Scientific reports* **6**, 20482, (2016).

244   Przysinda, E., Zeng, T., Maves, K., Arkin, C. & Loui, P. Jazz musicians reveal role of expectancy in human creativity. *Brain and Cognition* **119**, 45-53, (2017).

245   Large, E. W., Kim, J. C., Flaig, N. K., Bharucha, J. J. & Krumhansl, C. L. A neurodynamic account of musical tonality. *Music Perception: An Interdisciplinary Journal* **33**, 319-331, (2016).

246   Large, E. W. & Palmer, C. Perceiving temporal regularity in music. *Cognitive Science* **26**, 1-37, (2002).

247   Patel, A. D. & Iversen, J. R. The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Front Syst Neurosci* **8**, 57, (2014).

248   Cannon, J. J. & Patel, A. D. How Beat Perception Co-opts Motor Neurophysiology. *Trends in Cognitive Sciences*, (2020).

249   Rao, R. P. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat.Neurosci.* **2**, 79-87, (1999).

250   Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* **36**, 181-204, (2013).

251   Kahl, R. Selected Writings of Hermann Helmholtz. (1878).

252   Gregory, R. L. Perceptions as hypotheses. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **290**, 181-197, (1980).

253   Gibson, J. J. *The ecological approach to visual perception*. (Houghton Mifflin, 1979).

254   Fuster, J. *The Prefrontal Cortex Anatomy, Physiology and Neuropsychology of the Frontal Lobe*. (1997).

255   Neisser, U. *Cognition and reality: Principles and implications of cognitive psychology*. (W H Freeman/Times Books/ Henry Holt & Co, 1976).

256   Arbib, M. A. & Hesse, M. B. *The Construction of Reality*. (Cambridge University Press, 1986).

257   Cisek, P. & Kalaska, J. F. Neural Mechanisms for Interacting with a World Full of Action Choices. *Annual Review of Neuroscience* **33**, 269-298, (2010).

258   Isomura, T., Parr, T. & Friston, K. Bayesian Filtering with Multiple Internal Models: Toward a Theory of Social Intelligence. *Neural Comput* **31**, 2390-2431, (2019).

259   Friston, K. & Frith, C. A Duet for one. *Conscious Cogn* **36**, 390-405, (2015).

260   Hunt, B. R., Ott, E. & Yorke, J. A. Differentiable generalized synchronization of chaos. *Physical Review E* **55**, 4029-4034, (1997).

261   Ghazanfar, A. A. & Takahashi, D. Y. The evolution of speech: vision, rhythm, cooperation. *Trends Cogn Sci* **18**, 543-553, (2014).

262   Wilson, M. & Wilson, T. P. An oscillator model of the timing of turn-taking. *Psychon Bull Rev* **12**, 957-968, (2005).

263   Keller, P. E., Novembre, G. & Loehr, J. *Musical Ensemble Performance: Representing Self, Other and Joint Action Outcomes*.in *Shared Representations: Sensorimotor Foundations of Social Life Cambridge Social Neuroscience* (eds Emily S. Cross & Sukhvinder S. Obhi) 280-310 (Cambridge University Press, 2016).