Music in the brain

Peter Vuust[1]✉, Ole A. Heggli[1], Karl J. Friston[2] and Morten L. Kringelbach[1,3,4]

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 brain function. Music processing in the brain — namely, the 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 comprehensive 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.

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

Harmony
The combination of multiple, simultaneously 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.

> *"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 (FIG. 1) — each of which is subserved by overlapping but distinct neural networks[1,2]. These fundamental ingredients may, in rare instances, be experienced in isolation, such as when one is listening to the single melodic lines in a Gregorian chant or the epic drum fill in Phil Collins's 'In the Air Tonight' — introduced unpredictably after three minutes of vocals and keyboard playing. Mostly, though, these constituents interact to create 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[1,3] and emotion[4]. Accordingly, music perception engages brain networks related to action, emotion and learning in addition to the auditory system (FIG. 1d).

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: for example, tonality, which is the experience of a hierarchy of relations pointing towards a tonal centre in melody and harmony; and metre, the experience of regularly recurring rhythmic patterns and accents, which underwrites the way we move regularly to sometimes highly irregular musical rhythms. In this Review, we describe the process of listening to music, in which 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[5].

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 — that guides our perception. Take the example of a repeated syncopated rhythm (FIG. 2), a rhythm wherein one beat is displaced by a fraction. Here we experience an error at the unexpected, syncopated note as proposed by the PCM model. 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 emotional responses to music and musical learning, which updates our underlying predictive model over time. Music is thus a powerful tool for studying the predictive brain, owing to the way its structure licenses anticipation.

In the following sections, we provide an overview of music perception in the brain. We first introduce the PCM model. Then, we discuss the fundamental constituents of music in relation to this model before turning to more complex music processing that entails action,

[1] *Center for Music in the Brain, Aarhus University and The Royal Academy of Music (Det Jyske Musikkonservatorium), Aarhus, Denmark.*

[2] *Wellcome Centre for Human Neuroimaging, University College London, London, UK.*

[3] *Department of Psychiatry, University of Oxford, Oxford, UK.*

[4] *Centre for Eudaimonia and Human Flourishing, Linacre College, University of Oxford, Oxford, UK.*

✉*e-mail: pv@musikkons.dk*

https://doi.org/10.1038/s41583-022-00578-5

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

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 the PCM model to encompass musical interaction and communication in interpersonal relationships and hierarchical organization in groups.

The PCM model

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 their sensory inputs, integrate information with other networks and adapt to new stimuli. Recently, active inference, an influential theory of predictive processing[6], 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 be fully understood only in the light of prediction[7–9]. Music-related predictions are linked to various emotions, and the relationship between musical anticipation

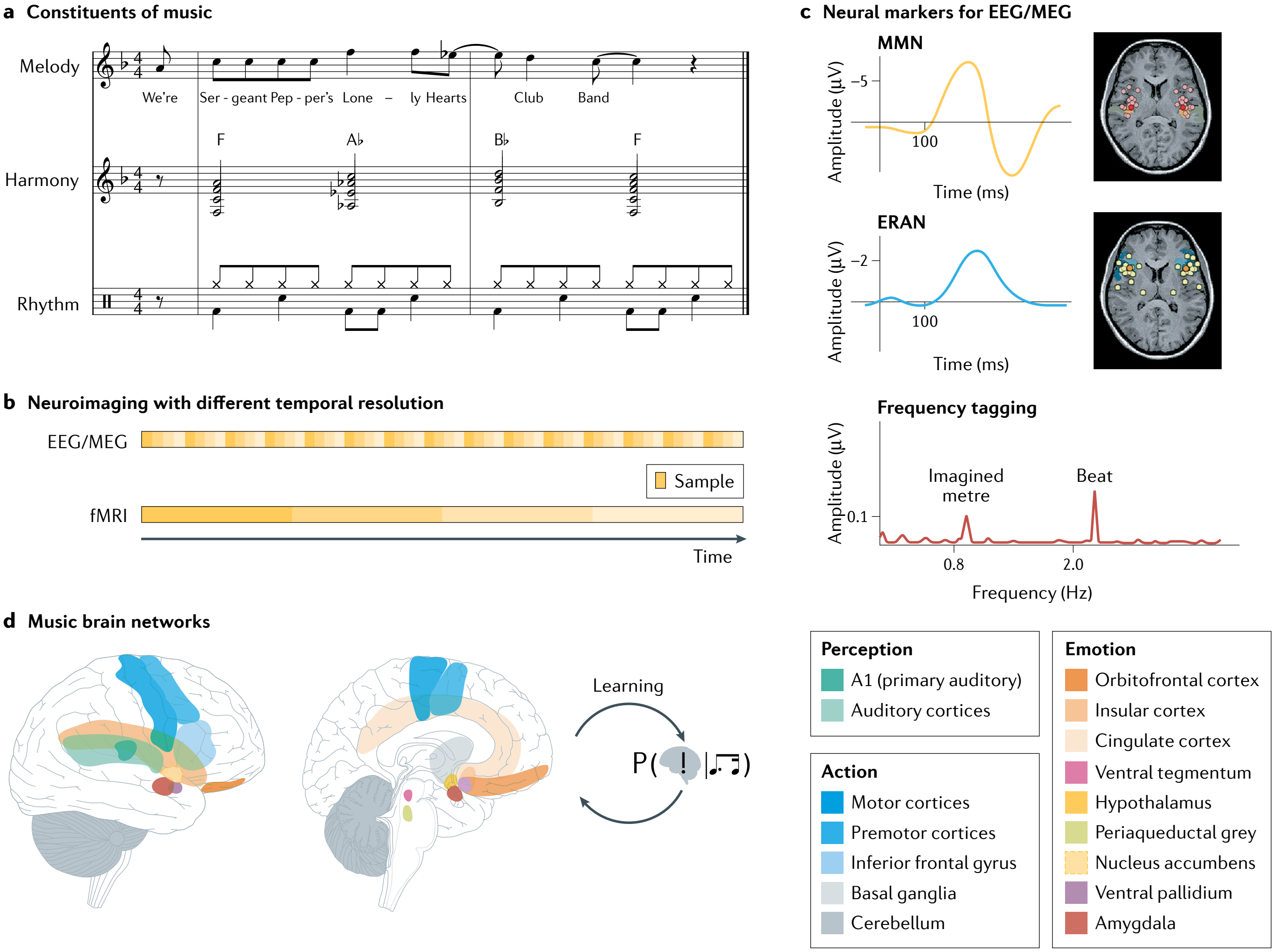


Fig. 1 | **From the structural constituents of music to perception, action and emotion in the brain.** The figure shows the constituent parts of music and their underlying brain bases as established by electrophysiological and neuroimaging techniques. **a** | The melody (the successive pitches notated in the top staff of the score), harmony (the chord progression listed in the middle staff) and rhythm (the percussion notation in the bottom staff) in an excerpt from 'Sgt. Pepper's Lonely Hearts Club Band (Reprise)' by the Beatles. **b** | Brain responses to music can be measured with neuroimaging methods, typically electroencephalography (EEG) or magnetoencephalography (MEG), or functional MRI (fMRI), which have different temporal resolutions. The sampling for EEG and MEG is typically on the scale of 1–10 ms and for fMRI is on the scale of 0.72–3 s. **c** | Neural markers obtained with two analysis methods for EEG/MEG data: event-related potentials and frequency tagging. The perhaps most used event-related potentials are the mismatch negativity (MMN) and the early right anterior negativity (ERAN), which are markers of auditory expectancy violation. The MMN waveform (top) typically occurs at around 110–250 ms, while the ERAN waveform (middle) typically occurs at around 150–200 ms. The images to the right of these panels show how the sources of these signals have been localized in slightly different regions of the brain. Finally, another prominent method, frequency tagging, shows how the beat (here, an unaccented repeated pulse) and an imagined 3/4 metre are represented as peaks in the amplitude spectrum of the EEG recording (bottom). **d** | The brain networks involved in music processing, with the key brain structures related to music perception, action and emotion being highlighted. Learning is illustrated here as the continuous update of real-time predictive brain models through Bayesian inference. For part **a**, 'Sgt. Pepper's Lonely Hearts Club Band (Reprise)' words & music by John Lennon & Paul McCartney copyright 1967 Sony/ATV Music Publishing. All Rights Reserved. International Copyright Secured. Used by Permission of Hal Leonard Europe Ltd. Part **c** adapted, with permission, from REF.[71], Elsevier.

Expectations
Mathematically, the expected values or means of random variables.

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

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

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

Predictions
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.

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

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).

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.

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.

and emotion has been proposed to be associated with survival-related anticipatory brain mechanisms[10]. 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[11] (forming, for example, schematic expectations), familiarity with a particular piece or genre of music[12] (veridical expectations), short-term memory for the immediate musical past (forming dynamic expectations) and deliberate listening strategies[13]. The neuronal mechanisms and functional architectures underlying musical expectation are thus shaped by culture, personal listening history, musical training and biology[14].

Recently, research into music perception has turned towards experiments modelling musical structure, which evinces anticipation[15] and predictive mechanisms. To integrate these approaches in neuromusic research, we developed the PCM model in several recent articles[5,7,16,17] (FIG. 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.

The PCM model proposes that music perception, action, emotion and learning are recursive Bayesian processes, by which the brain attempts to minimize prediction error[18] as formalized in enactive versions of predictive processing (also known as active inference). Accordingly, the processes underlying music perception and action are coupled, such that perception minimizes prediction error by updating the predictions, while action reduces prediction error by generating predicted sensory signals. Emotion, attention and motivation act as Bayes optimal biases to contextualize prediction, thereby guiding behaviour, action and learning.

The notion of the brain as a hierarchical prediction machine — in which 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 later herein. 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 metres or tonalities (FIG. 3). Hence, a Bayesian formulation of predictive coding applies naturally to processes that involve the inference of hidden or latent causes — such as metre and tonality — from the music.

When one is listening to music, attentional selection of which prediction errors to resolve rests on predicting not just the content of sensory streams but also their predictability or precision. Put simply, prediction errors are useful only when things are predictable. The Bayesian belief updating inherent in the PCM model is literally precision engineered, in the sense that it rests on 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 the 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 mismatch negativity (MMN) to an oddball is modulated by its predictability[19]. In music perception, the finding that the amplitude of the MMN — a preattentive marker of prediction error — is reduced (to rhythmic deviations and mistuned pitches) in less predictable contexts offers a clear example of so-called precision-weighted prediction errors[20,21]. These studies

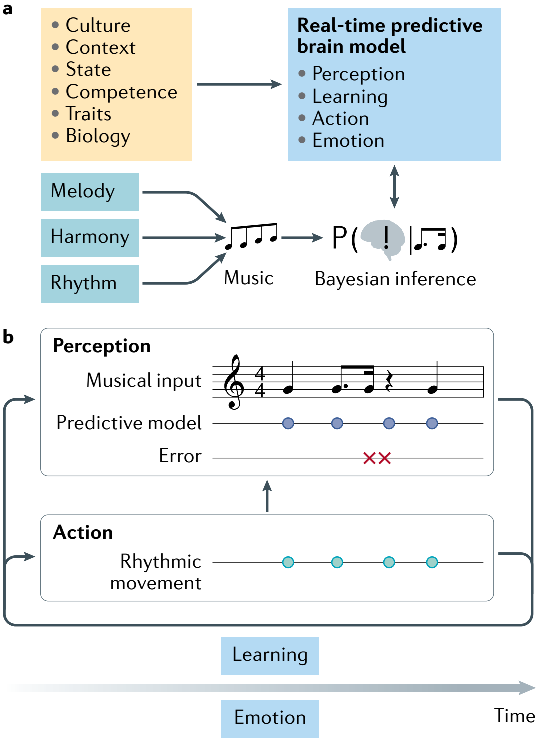
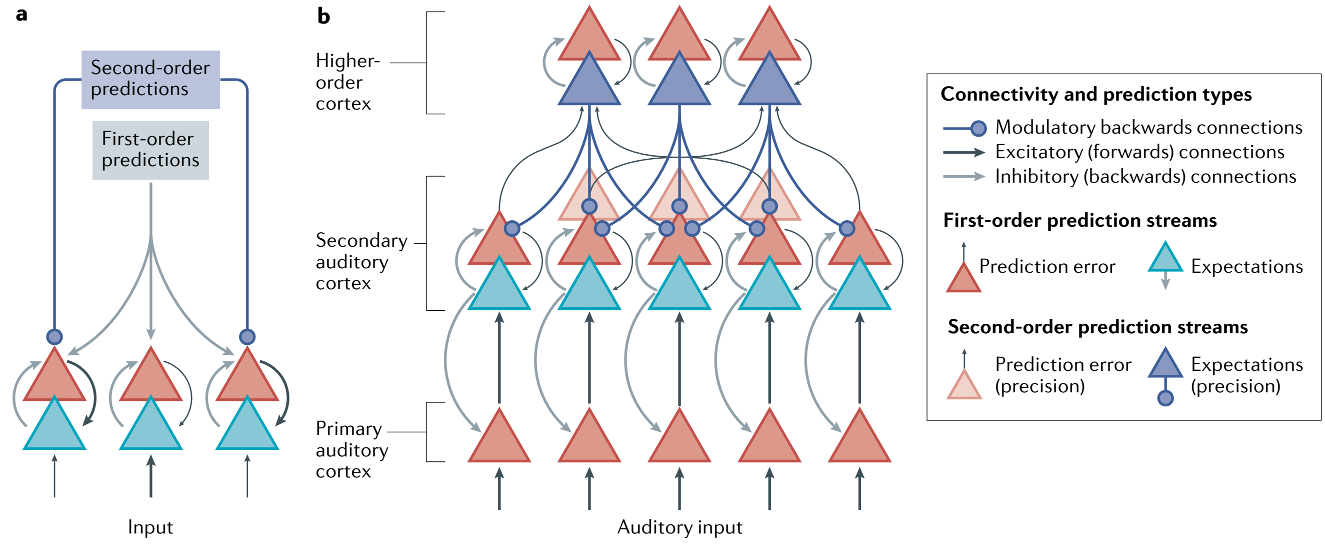


Fig. 2 | **Predictive coding of music. a** | Music perception is guided by the brain's real-time predictive (generative) model (marked with an exclamation mark), which is based on prior experience. The predictive model relies on the cultural background, musical competence, the current context and the brain state, including the attentional state and the emotional state, as well as 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-metre 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 music-related emotions, which in return modify action and perception.

Box 1 | Predictive processing and coding

Predictive processing (also known as active inference) is a general theory of neural processing inspired by research in artificial intelligence, statistical physics and systems neuroscience[6,255,256]. The basic idea can be traced from the students of Plato, through to Kant and Helmholtz[257], and to theories of perception as hypothesis testing[258]. Enactive versions of predictive processing offer integrative accounts of action and perception[259–263] by formalizing how specialized brain networks identify and categorize causes of sensory inputs, integrate information with other networks and actively sample new stimuli[256]. In brief, active inference proposes that perception, action and learning are Bayesian processes by which the brain attempts to minimize hierarchical prediction errors. The figure shows a schematic illustration of the computational architecture of neuronal message passing that underlies predictive coding in the brain. As 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 be of either the input or the precision (that is, the predictability) of that input. These are designated first-order 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. Black arrows indicate forward connections and grey arrows indicate backward connections. In this example, unpredicted auditory input is passed forward to the auditory cortex in the form of ascending prediction errors (for example, from the medial geniculate body). These prediction errors (black arrows) drive posterior expectations (for example, encoded by deep pyramidal neurons) that return descending predictions (grey 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 (for example, superficial pyramidal cells). This enables high levels to select the prediction errors that convey the most precise or predictable information (compare with attentional selection). In short, there are two kinds of descending predictions in predictive coding: first-order prediction of content (grey 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 of certain lower-level features that are consistent with higher-level constructs. This allows the predictive coding of music model to explain figure-ground phenomena in music, such as selecting between different metrical interpretations of a 3 against 4 polyrhythm.



Dynamic expectations
Short-lived expectations that dynamically shift owing to the ongoing musical context, such as when a repeated musical phrase causes the listener to expect similar phrases as the work continues.

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

demonstrate the scope of the PCM model for explaining the fundamental aspects of music processing.

Learning can be cast as minimizing precision-weighted (that is, 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 the early right anterior negativity (ERAN) responses, which are markers of harmonic expectation violation in non-musicians and musicians alike[22,23]. 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 with that in 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 (see later).

Importantly, experiments have shown that the brain architecture subserving precision-weighted prediction errors differs depending on which musical phenomenon — for example, melody, harmony or rhythm — is being studied. For example, a predictive coding-based analysis of the effective connectivity in a melodic oddball functional MRI (fMRI) paradigm revealed that mismatch responses are best explained by a fully connected bilateral auditory network comprising the primary auditory cortices and the planum temporale[24]. Here, the observed increase in excitatory connectivity from the left primary auditory cortex to the planum temporale has been interpreted as the passing of a precise prediction error from lower to higher areas of the hierarchical processing network[25,26], and the observed decrease in inhibitory connectivity within the left primary auditory cortex has been viewed as a precision-related increase in the gain of the superficial pyramidal cells encoding prediction error[27,28]. By contrast, as discussed later,

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

Functional MRI
(fMRI). A neuroimaging technique that images rapid changes in blood oxygenation levels in the brain.

Groove
In the realm of contemporary music, a persistently repeated pattern played by the rhythm section (usually drums, percussion, bass, guitar and/or piano). In music psychology, the pleasurable sensation of wanting to move.

studies of the pleasurable experience of musical harmony have revealed predictive coding mechanisms including precision-weighted prediction errors[29] related to emotion and reward brain networks[29], whereas studies of musical groove have implicated additional motor-related areas.

In the following sections, we review music perception, action, emotion and learning in the light of the PCM model. 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 — that is, the 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[30] and a series of overtones that contributes to its tone quality (timbre), which distinguishes it from other instruments[31,32]. 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[33,34]. 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 (for example, the note C in different octaves). The auditory cortices are central to pitch processing: fMRI suggests that pitch height is processed in the planum temporale posterior to the primary auditory cortex, whereas chroma change is processed in the planum polare, anterior to the primary auditory cortex[35]. 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 (FIG. 1d). A dorsal stream, projecting from primary areas via the planum temporale over the parietal cortex to the dorsolateral prefrontal cortex[36], is supposed to support sensory–motor integration, articulation and memory functions[37] and thereby link the neural apparatus for melody perception and action[38].

Once musical pitches are combined into melodies, global properties emerge, such as melodic contour[39], melodic expectations and tonality[40]. Most melodies point to a certain tonality, even though 12-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 centre, such as a C-major scale, which corresponds to seven white keys on a piano keyboard and excludes the associated black keys. The tonality is not necessarily expressed directly in the auditory input to the ears[5] but is an endogenously generated hierarchical predictive framework that underwrites the perception of melody and harmony (FIG. 3). Listeners, even without explicit musical training, have implicit knowledge of the statistical regularities of melodies of their own culture. This knowledge[41] is constantly applied to form musical expectations by comparing a given note to the given statistical distribution[42,43]. 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[44,45], including the information dynamics of music model, which assigns measures of information content (contextual unexpectedness: see prediction error) and entropy (uncertainty) to each note of a scale using short-term and long-term statistical regularities[46]. 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[47] in particular using the MMN response[48,49] (FIG. 1). Importantly, in unpredictable compared with predictable melodic contexts (as modelled by the information dynamics of music model[21]), the MMN amplitude is reduced, and behavioural deviant detection is impaired with regard to mistuned pitches. In other words, the more precise our melodic predictions — according to culture-dependent statistical learning — the larger the MMNs to surprising notes.

Individual differences have an important role for the predictive processing of melody. In general, musicians score consistently higher than non-musicians on tests that involve distinguishing between different melodies[50]. High scores on such tests are associated with a higher amplitude and shorter latency of the preattentive auditory responses, such as the MMN, to expectancy violations[12,51]. This indicates that musicians develop a more precise predictive model presumably as a result of explicit learning.

Melody perception recruits parts of the brain that are specialized for purposes other than audition, such as motor tasks, as well as emotional and cognitive

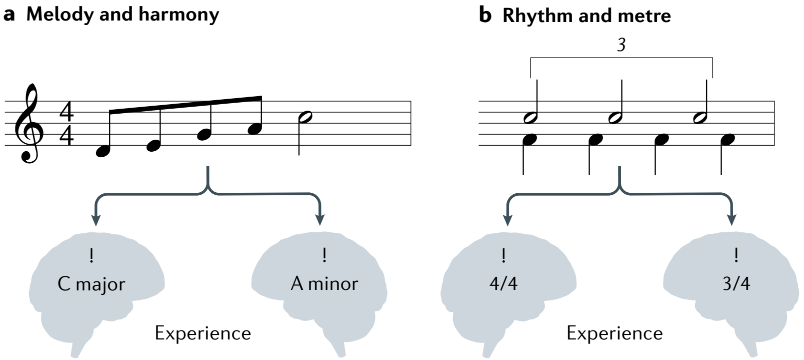


Fig. 3 | **How we may experience the same musical material with different real-time predictive brain models. a** | There can be different interpretations of a simple ambiguous melody according to different tonalities; for example, here C major versus A minor. The melody is compatible with both C major 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 (for example, major or minor). This may lead to quite different experiences of this melody, rendering the melody, for example, 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 metre as the metrical 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.

processes, as shown, for example, in the so-called free listening paradigms[52–54], which combine music information retrieval of real music recordings with fMRI or magnetoencephalography (MEG)[53,54]. Moreover, action and perception can be engaged even without stimuli, as demonstrated by event-related potential, positron emission tomography and fMRI studies showing that the formation of musical mental images engages auditory sensory and premotor areas[37,55,56]. In summary, the study of melody processing has moved towards using naturalistic melodies, emphasizing the coupling of perception and action-related brain mechanisms, in which prediction has a key role.

Harmony. Melodies, especially in Western music, are typically accompanied by harmony that is created by chord progressions (different chords played successively) played either 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 an emotional response. For example, the sound of a C-major triad is perceived as happier than that of a C-minor triad (C, E♭ and G) by Western listeners[57]. Since the ancient Greeks, it has been known that the physical properties of sounds, such as the 'roughness' of chords, 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[58]. This is tied to the experience of sound roughness[59], 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 that are very close in pitch[60,61]. Importantly, however, roughness may be perceived differently according to musical style and culture[62].

Chord progressions establish musical expectations and typically a sense of tonality. Whereas tonality is known in music from all cultures studied[63], neuroscientific studies have concentrated mainly on Western harmony[40,64,65], which follows a specific set of rules, for example, in different types of harmonic cadences. Breaking these rules elicits the ERAN (FIG. 1), which was discovered by Koelsch and colleagues using electroencephalography (EEG) and MEG[66,67]. The ERAN peaks at 150–200 ms after deviant onset, and its latency and amplitude are modulated by attention or knowledge of impending outcomes[68], and musical training[23]. The sources of the ERAN have been localized in the inferior frontal gyrus using MEG as well as fMRI[69,70]. It is most often studied in semi-attended paradigms, in which 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[71–73]. The rules of harmony have often been termed 'musical syntax' or 'musical 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[74,75], 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 inferior frontal gyrus is still unresolved but is probably related to higher-level processing of the temporal order of sequences[76].

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[77–79] and activity in a frontoparietal motor-related network in non-musicians listening to a learned melody after practice playing[80]. For trained pianists, internalized harmonic expectations in auditory and motor networks may furnish modality-specific networks for harmony prediction that interact with the inferior frontal gyrus to optimize action and perception[81]. It is thus likely that musical expertise affects harmony processing by engaging motor mechanisms because of the active experience in auditory–motor association.

As for melody processing, a recent trend has been to model both music and expected brain responses mathematically to enable more ecologically valid approaches to the study of harmony. This allows 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[29] showed maximal pleasure ratings with regard 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, the hippocampus and the nucleus accumbens (NAcc).

The aforementioned 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 Amazonian individuals with limited exposure to Western music do not exhibit a Western-like preference for consonant versus dissonant intervals[62], 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, for example, people are listening to blues versus compositions by Beethoven. The well-known associations between major and minor and happy and sad emotions, respectively, 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.

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

Timbre
Also known as tone colour or tone quality, the perceived sound quality of a sound, including its spectral composition and its additional noise characteristics.

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

Information content
The contextual unexpectedness or surprise associated with an event.

Entropy
In the Shannon sense, the expected surprise or information content (self-information). In other words, it is the uncertainty or unpredictability of a random variable (for example, an event in the future).

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

Event-related potential
A very small electrical voltage generated in the brain structures in response to specific events or stimuli.

Consonant and dissonant intervals
Psychologically, consonance is when two or more notes sound together with an absence of perceived roughness. Dissonance is the antonym of consonance. Western listeners consider intervals produced by frequency ratios such as 1:2 (octave), 3:2 (fifth) or 4:3 (fourth) as consonant. Dissonances are intervals produced by frequency ratios formed from numbers greater than 4.

Harmonic cadences
Stereotypical patterns consisting of two or more chords that conclude a phrase, section or piece of music. They are often used to establish a sense of tonality.

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

This association is not found in all musical cultures[82], 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[83–85]. It is still unclear whether this is a universal principle expressed differentially in different cultures. Nevertheless, basic emotions may still be recognized in realistically sounding music material across cultures[86]. However, this is an ongoing field of study.

In summary, 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. When we listen to 'Blame It on the Boogie' by the Jacksons (FIG. 4), it is difficult to refrain from tapping a foot or bobbing one's head to the beat. This explicit link between the body and the mind, action and perception has recently made musical rhythm a burgeoning topic in cognitive neuroscience.

Rhythm can be produced by the onset of the notes in a melody or without a melody when it is played on designated percussion instruments, where pitch may be less clear[87]. The perception of rhythm usually involves the simultaneous perception of evenly spaced pulses[88] and a metre, which structures this pulse train and its subdivisions into patterns of differentially accented beats. Listening to pulse trains involves the prediction of following events[89], as indicated in studies showing brain responses to omission of a beat or after the end of rhythmic sequences[90–93].

Brochard and colleagues[94] provided strong evidence for metre perception in the simplest possible experimental setting using event-related potentials to show 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 metre are not necessarily expressed directly in the auditory input to the ears[5] but emerge under hierarchal predictive processing that enables the recognition of successive musical events over time[95–97]. The perception of a clear musical metre facilitates rhythm memory[98,99], learning[100] and perceptual sensitivity even at a young age[101].

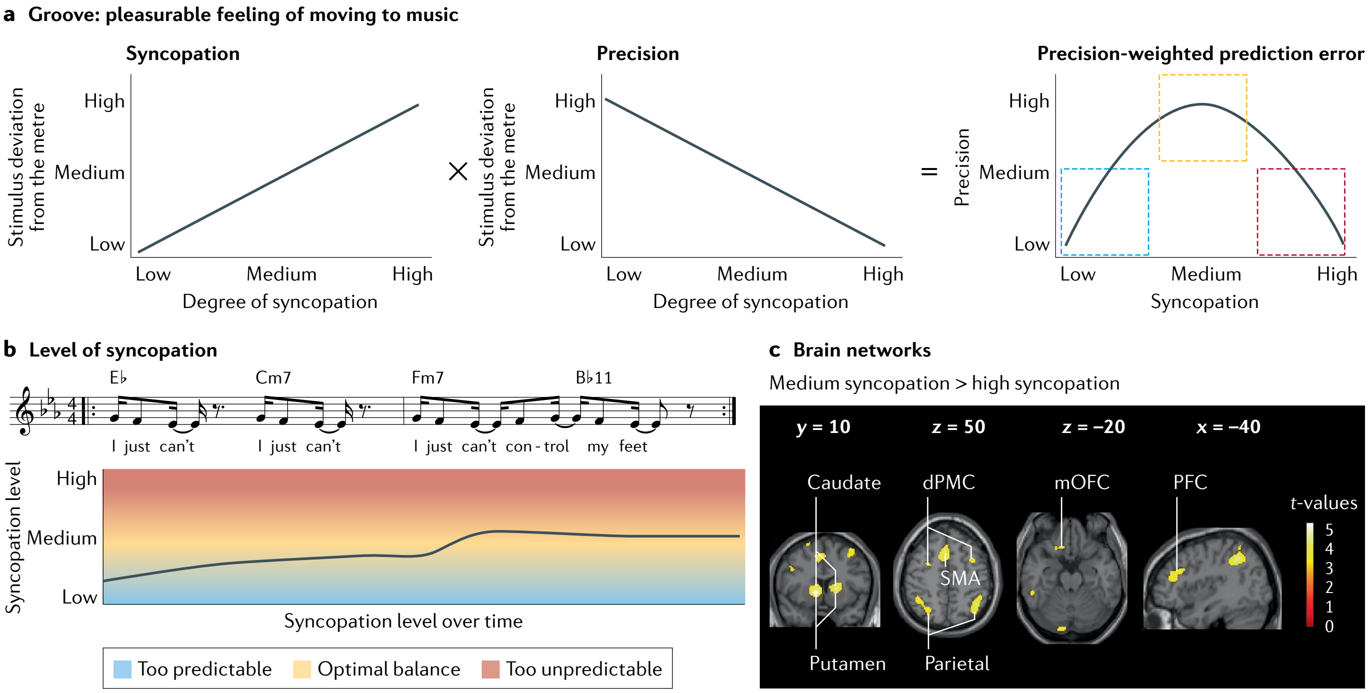


Fig. 4 | **Groove: the pleasurable sensation of wanting to move to music.** **a** | How the inverted U-shaped relationship between rhythmic predictability and the experience of groove observed in groove ratings can be modelled as the product of stimulus syncopation and the precision of the predictions relative to prediction error; that is, the precision of our metrical expectations[16]. 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 trade-off between stimulus complexity (amount of syncopation) and the ability to maintain a sufficiently stable metre for moving in time with music. **b** | In the song 'Blame It on the Boogie' by the Jacksons, 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-related structures (premotor and basal ganglia regions), reward-related structures (orbitofrontal areas and nucleus accumbens) and timing-related brain structures in the basal ganglia when contrasting medium and high syncopation rhythms[136]. dPMC, dorsal premotor cortex; mOFC, medial orbitofrontal cortex; PFC, prefrontal cortex; SMA, supplementary motor area. For part **b**, 'Blame It on the Boogie' words and music by Elmar Krohn, Thomas Meyer, Hans Kampschroer, Michael Jackson Clark and David Jackson Rich copyright 1977 Delay Edition. All Rights Administered by Chrysalis Music Holdings GmbH. All Rights Reserved. International Copyright Secured. Used by Permission of Hal Leonard Europe Ltd. Part **c** adapted, with permission, from REF.[136], Elsevier.

Frequency tagging
A method of analysing 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 FIG. 1c.

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

Despite the possibly innate human ability to synchronize movements to the musical metre, 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 and the other corresponding to the motor system tuned to operate near a double limit cycle bifurcation[102]. This model was able to explain participants' ability to synchronize finger tapping with increasingly syncopated rhythms. Accordingly, frequency tagging (FIG. 1) in electrophysiological recordings shows that even for such rhythms in which the metre is not acoustically accented, the fundamental frequencies of the metre still dominate the signal[103,104]. However, the neural entrainment to rhythm and the different contributions of auditory and motor cortical and subcortical structures in establishing the metre percept are still far from well understood[105], and it is essential to acknowledge a substantial top-down influence on metre perception[18,106]. Metre perception may hence be modulated by cultural[107–109] and innate biological factors (FIG. 2).

Rhythm and metre perception is more developed in musicians, who, as with melody and harmony, score consistently higher than non-musicians on rhythmic ability tests[50]. In addition musicians exhibit a higher amplitude and shorter latency of the MMN to violations[110], suggesting that they deploy more precise predictive models than non-musicians.

The literature on rhythm perception discloses the involvement of the auditory pathway in detecting structural deviations from the metre[110]. These deviations are marked by the MMN, which again is modulated by the predictability of the rhythmic context[20]. Rhythm perception involves large parts of the motor system — the premotor cortex, supplementary motor area, basal ganglia and cerebellum[111–113]. This motor system activity can, to some extent, be attributed to the establishment and maintenance of the musical pulse and metre[114]. These processes, which underlie our ability to dance to music, are measurable already in newborns[115], and can be influenced by training infants to recognize either a duple metre (2/4) or a triple metre (3/4) of the same ambiguous rhythm[116]. Furthermore, synchronizing movements to the same metre may lead to prosocial behaviour[117], as shown in studies where infants exhibit more helpful behaviour after having been bounced in synchrony with an experimenter[118].

The complexity of the brain circuits underlying metre perception may explain why metre 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 finger tapping to a simple metronome at different tempi between approximately 40 and approximately 400 beats per minute depending on musical expertise and do so by predicting the subsequent beats — and may perform the task across modalities[119,120]. By 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[121]. MMN recordings to onbeat and offbeat deviants show that monkeys are sensitive to the isochrony of the stimulus but only humans are sensitive to its metrical structure[122]. Studies in chimpanzees show equally poor results in beat synchronizing to metronomes[123]. 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 summary, the study of musical rhythm demonstrates how sensory input provided by auditory rhythms (bottom-up) are met by predictive models such as the metre (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 section, we consider more complex musical phenomena, which integrate melody, harmony and rhythm and exemplify the crucial role of precision-weighted prediction error.

Action

Why do people rush to the dance floor when hearing the grooves on James Brown's records and move to the music with such apparent pleasure[124–126]? The study of groove and the pleasurable sensation of wanting to move to music[127] is a prime example of how the PCM model — and the concept of precision-weighted prediction error — can inform our understanding of music processing.

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[128]. 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[129]. 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[130] — has been proposed as one of the underlying mechanisms of groove[131] and as one of the reasons why we move to music.

The influence of syncopations on the experience of groove has been described by the predictive coding of rhythmic incongruity (PCRI) model[16], a formal operationalization of the PCM model for rhythm only. According to the PCRI model, brain and behavioural responses to rhythm are modelled in terms of the precision-weighted prediction error; that is, the product of the metrical predictability (precision) and the stimulus deviations from the metre (FIG. 4). Importantly, this model explains the observed inverted U-shaped relationship between the degree of syncopation in and the experience of groove[132,133], 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 the PCRI model, these medium-level 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 the processing hierarchy. In intermediate-level syncopated rhythms, the brain may, according to active inference, resolve prediction error either by revising predictions or through action — for example, by moving the body. Actively resolving prediction errors may explain our drive to reinforce the metre — by moving in time with the beat — while attenuating the precision of proprioceptive and auditory prediction errors. By 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 metre is less precise than the sensory evidence, precluding sensory attenuation and movement.

The inverted U-shaped relationship between syncopation and groove experience has been replicated independently of culture and rhythmic proficiency[128] and using physiological measurements such as pupillometry[134]. It has been tested for rhythm and groove in a within-culture and between-culture approach[128,132,135] and is influenced by musical expertise[135].

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[136] (FIG. 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[16,102]. Importantly, optimal groove experience has been linked to activity in the NAcc and the orbitofrontal cortex, which are key regions of the reward network that are particularly sensitive to the predictability of the consequences of action[136].

It is important to note that the relationship between the rhythmic sensory input and the schematic expectations of the metre is only one of several interacting predictive processes occurring 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[55,137,138]. It is equally important to note that the repeated patterns in many non-Western grooves, such as the Afro-Cuban tumbao, still support a stable metre sensation in experienced listeners, even though they contain few onsets on the most salient metrical positions[139]. There is therefore an ongoing debate as to what extent the metre — 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[140].

In summary, the PCM model proposes that the pleasurable wanting to move is mediated by prediction-based brain mechanisms that 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 and pleasure

A defining feature of music, closely related to theories of its evolutionary origin[141], 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 may provide music-specific experiences, such as the sensation of groove described earlier herein. Even though music is clearly able to give rise to everyday emotions[142,143], and adults listen to music partly to regulate their affective state[144], 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[145,146]. Furthermore, there is a dissociation of valence and pleasure ratings[82] 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 corticothalamostriatal reward circuits than disliked music, regardless of whether the music is sad or happy[147]. Because of this apparent paradox, musical sadness is the subject of several recent studies[148,149], and multiple theories try to explain its existence, often pointing to societal and individual benefits[82]. 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 conduct the translational process from music to emotion can be explained by several psychological mechanisms[150], which typically fall into three categories: hardwired responses, evoking universal survival-related responses such as when brainstem responses to loud sounds trigger fear responses; extramusical associations, in which music links to some extramusical space that carries the particular emotion, such as evaluative conditioning, emotional contagion, visual imagery and episodic memory; and anticipation, when musical structure establishes, fulfils or disappoints expectations that are set up within the music itself. Whereas hardwired responses and extramusical 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 the publication of a seminal book by Meyer[8], it has become increasingly clear that music anticipation may induce various complex emotional responses such as awe, surprise and discomfort and evoke laughter, foot tapping, humming, tears and a lump in the throat[10]. It can give rise to psychogenic responses such as 'shivers down the spine', increased heart rate and increased perspiration[151].

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 and orbitofrontal cortex[4]. An outstanding question is to what extent the emotion-related networks involved in processing of music are mediated by universal, cultural or individual factors: that is, which aspects of music perception are developed only after exposure to a specific musical culture? One pioneering study[86] comparing Western listeners with participants from the African Mafa people pointed towards the above chance level recognition of basic emotions — such as happy, sad and scared or fearful emotions — when they were listening to the other culture's music. Importantly, though, Mafa individuals, who have been culturally isolated from Western music, showed much lower emotion recognition performance with regard to Western music than Western listeners. Furthermore, in a recent large-scale Internet study, US and Chinese listeners identified 13 distinct types of subjective experience associated with music from both cultures[152]. It is, however, unclear to what extent music emotions can be universally recognized[153], or how much they are a result of statistical learning caused by increasingly globalized music listening behaviours[154]. Speaking to the latter, recent modelling approaches highlight the importance of aligned musical priors to the cross-cultural experience of music emotion[155,156]. Predictive coding has therefore become a hot topic in the study of music-related emotions.

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[157]. 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[158]. Recently, these intense experiences have been shown to lead to dopamine release in the striatal system[159,160] with distinct roles for the caudate (anticipation) and the NAcc (reward experience)[161] and related to the degree of emotional arousal[162]. 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[163]. These individuals show reduced NAcc responses and decreased functional connectivity between the right auditory cortex and the ventral striatum — including the NAcc — compared with their responses on a monetary gambling task and compared with other participants with normal or greater than average pleasure responses to music[164].

On the basis of active inference formulations of predictive coding models, Gebauer et al.[165] 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 of and connectivity patterns involving the Nacc — a central component of the mesolimbic system[166]. Furthermore, that 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[29], 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 activity in the NAcc reflected only precision. This is consistent with the optimal zones of predictability and uncertainty in musical pleasure that are found in modelling studies[167].

Closely related to the subject of musical pleasure is musical taste. Why do people with very similar cultural exposures to music often differ greatly in musical preferences? This is a complex question that includes psychological answers[168], such as the well-known mere exposure effect showing increased liking with repeated listening to musical pieces[169]. Other important determinants of musical taste include contextual factors such as important sociological reasons, where music can be seen as a means to express group affiliation[170,171]. 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, for example, by the 'Big Five' or the Zuckermann sensation seeking score, has therefore consistently been related to differences in musical taste[172,173]. As musical pleasure depends on whether culturally learned musical expectancies are fulfilled or violated[174], listeners often exhibit biases favouring music of their native culture, making yet another case for predictive coding as an underlying mechanism of musical taste[168].

In summary, predictive mechanisms in music and the brain are key to understanding complex questions related to music-related emotion. A full description of the precision-weighted prediction errors involved in music-related 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 learning, which is integral to the PCM model. 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, emotion, visual perception and mental imagery[1]. Thus, 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[175]. The differences in cognitive skills relevant to music perception between musicians and non-musicians correspond to differences in both brain structure and brain function between these groups. Classic studies have shown that musicians exhibit morphological differences in the fibre bundle in the corpus callosum[176,177], as well as increases in cerebellar volume[178] and grey matter volume increases in primary motor and somatosensory areas in the left precentral gyrus, premotor areas and left cerebellum[179,180], in areas involved in temporal structuring of language and music[181,182] and in areas involved in auditory perception[183]. Studies have also revealed specific effects of musical training on white matter development[184,185]. Other studies have found functional differences between musicians and non-musicians in auditory and motor areas[1,186] that are dependent on the musical instrument[187,188], practice habits[189], the level of expertise[190] or the style of music being played[191,192].

These functional and structural differences, which are associated 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-term and short-term training on brain anatomy and function and in particular the development of auditory and motor processing, and auditory–motor coupling. In a pioneering study, using direct current EEG analyses, Bangert and Altenmüller[193] showed auditory–motor coupling changes in the cortex of beginners after as little as 20 min of piano training. The enhanced coupling of brain resources for perception and action was recently 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[194]. In addition, a recent study showed increased activity in frontoparietal 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[195].

Viewed in the light of the PCM model, the aforementioned 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 years, 4–5 years and 6–7 years from individuals who attended a musical play school throughout the follow-up period and children with shorter attendance at the same play school[196]. 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 old[197], 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 (for example, see REF.[198]).

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[199–202], cognitive inhibition tasks[176] and music training as a possible supplementary tool for helping children with developmental disorders[203] such as dyslexia[204,205]. 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[206,207].

In summary, cross-sectional and longitudinal studies of musicians and musical learning elucidate how predictive mechanisms for music are shaped by learning[208–210]. It appears that the heightened demands on auditory–motor coupling in music performance shapes brain structure and the ability to form music-related predictions with high precision. The studies discussed above 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 (FIG. 2). These factors are also crucial for how we understand the music of others. In the following section, we propose how music and the PCM model can be extended to encompass the role of communication in dyadic interactions and hierarchical organization in groups.

Musical communication

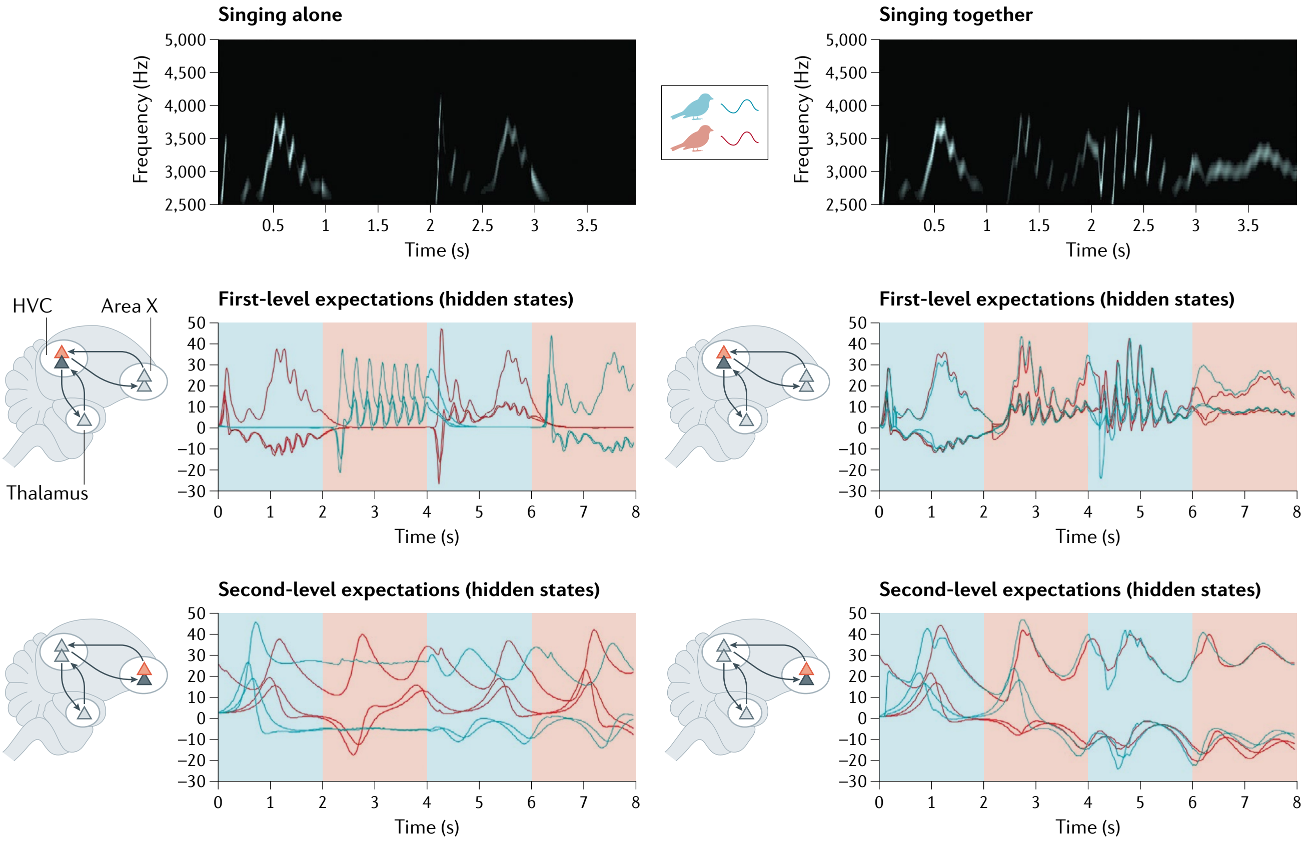
Even though most of the literature reviewed in this Review treats music perception in the individual brain, music is fundamentally a social phenomenon, in that we often make, listen and dance to music together. This makes it a fine-tuned instance of coordinated human interaction that involves interpersonal synchronization, social entrainment, learning, improvisation and communication (see BOX 2 for an example of this in other animals). Recently, the development of research methods such as dual EEG has prompted a line of neuroscientific and behavioural research into musical interaction[211,212]. It shows how competence, social context and mind set, such as empathy perspective taking, may promote interpersonal coordination[213] (FIG. 5).

Musical interactions rely heavily on prediction. While playing, we continuously make predictions about the sensory consequences of our own actions that we generally use to attenuate predicted sensations and amplify those caused by others[214]. 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 so far been found only in humans[215]. Joint action may thus be best understood within a predictive coding framework[216,217], 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 involve a dyad of two individuals who are finger-tapping together under different conditions[120]. Participants are typically placed in separate rooms with headphones and EEG equipment.

Box 2 | Hermeneutics, communication and music

Musical communication is a special case of fundamental communication between conspecifics, ranging from identifying a conspecific[264] through to sharing conceptual narratives[217,265]. Predictive processing 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 an implicit mutual predictability for free. In music, this corresponds to sharing tonality or metre; technically, this mutual predictability can be formalized as predictive coding — or more generally Bayesian belief updating based on shared (exchanged) sensory signals. If we share the same generative model, our neuronal dynamics can harmonize and evince a form of generalized synchrony[266]. 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[217]. 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 2 s 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. In the left panel, because this bird can hear only itself, the sonogram reflects descending proprioceptive predictions based on 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 (that is, 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. In the right panel, the two birds can hear each other. In this instance, the expectations show synchrony at both the sensory and the extrasensory hierarchical levels. Note that the sonogram is continuous over successive 2-s epochs — generated alternately by the first bird and the 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[267,268]. This predictive coding framework provides a powerful model for describing musical communication (FIG. 4). Figure adapted, with permission, from REF.[217], Elsevier.



Some studies are focused on sensorimotor synchronization, where the participants are told to synchronize finger tapping to each other[218], while some studies focus on isochronous self-paced tapping[219], and yet others study synchronization with a computer-generated metronome[220,221] or with piano recordings of self and other[222,223]. In studies where the participants are told to synchronize finger tapping to each other, the tapping analyses reveal that 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. The correlation between the participants' tap sequences demonstrates that the interaction is guided by mutual efforts to reduce prediction error at the millisecond level. This may result in at least three different relationships between participants: leader–follower relations wherein the leader is non-adaptive, which forces the follower to adapt to maintain synchronization[224]; mutual adaptation, in which both participants constantly adapt their taps to their partner's last tap[218]; 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[225]. Importantly, participants adapt differently to each other depending on their underlying internal predictive model. When musicians tap together with different underlying musical metres (for example, 4/4 and 3/4), they initially synchronize their tapping more poorly than when they hold identical musical metres in their minds[225]. This shows that interpersonal

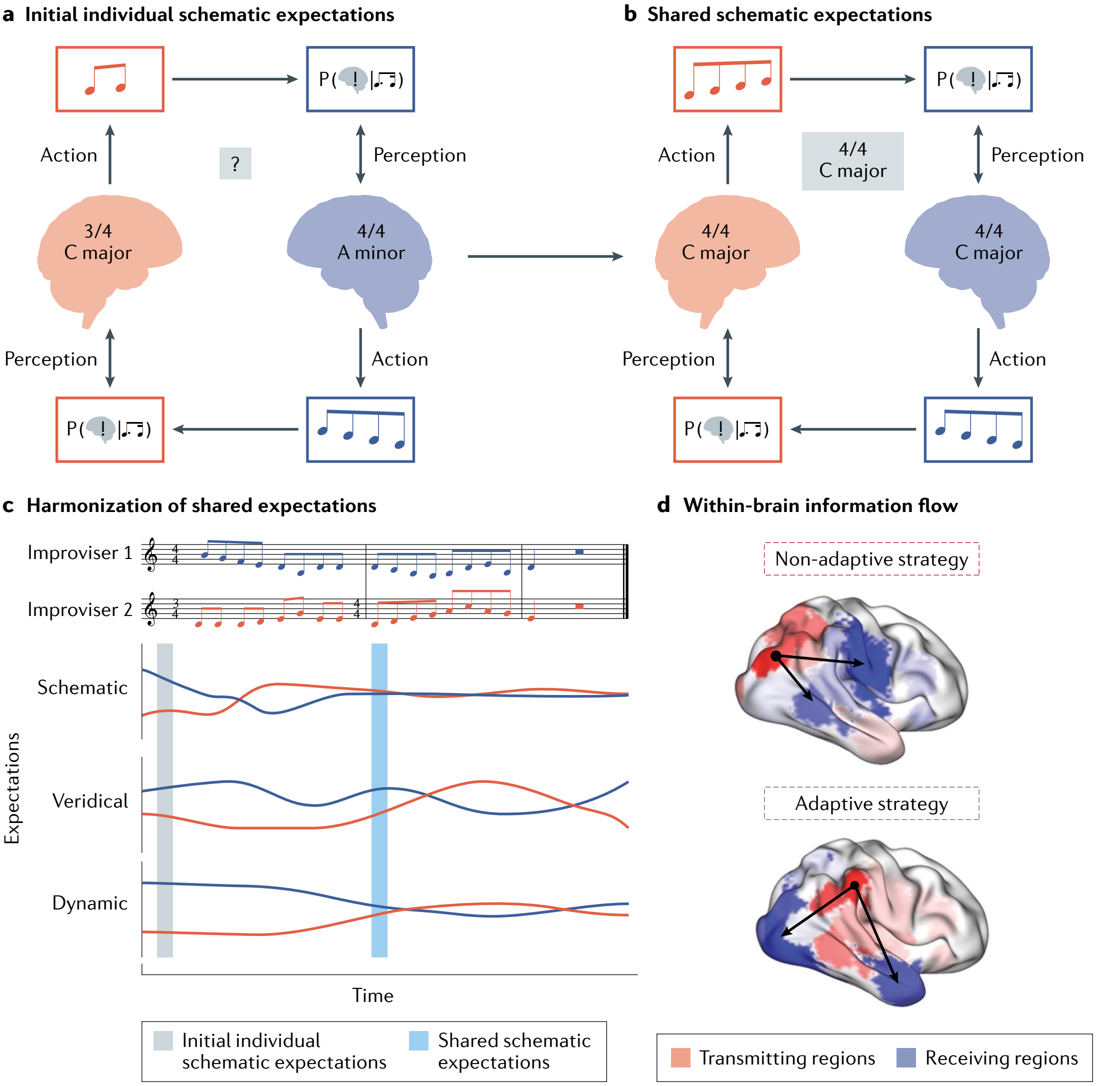


Fig. 5 | **Musical interaction.** The figure presents a model of musical communication inspired by predictive coding showing the continuous and reciprocal process of harmonizing expectations. **a** | Two improvisers (denoted by the red brain and the blue brain) may initially have different schematic expectations; that is, they could experience different metres (a 3/4 and 4/4 metre) 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 of a 4/4 metre and C major through reciprocal predictive coding mechanisms[254]. **c** | 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 part **a**), the schematic expectations (based on experience of metre and tonality) are quite different as illustrated in the music examples. After a while (marked with green, corresponding to part **b**), when a shared predictive metre and tonality model has been established, the schematic expectations of the two improvisers converge. The middle and bottom plots 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 metrical expectations, when two individuals from the same musical background tap a simple rhythm together. The connectivity-based electroencephalography data (computed from recurrent phase-locking patterns[227]) highlight how the information flows differently (as indicated by the arrows) in a non-adaptive musician versus an adaptive musician within a right-lateralized temporoparietal brain network with key nodes in the right somatosensory cortex, right precuneus, right supramarginal gyrus and right middle temporal cortex. Negative values indicating receiving areas are in blue and positive values indicating transmitting areas are in red. Part **d** is adapted, with permission, from REF.[227], OUP.

synchronization is dependent on the tappers' individual predictive coding (FIG. 5).

Differences in dyad tapping behaviour — for example, exhibiting a leader–follower, mutual adaptation or leader–leader tapping pattern — can 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[226]. EEG data furthermore showed that dyad members exhibiting mutual adaptation behaviour evinced intrabrain neural synchronization in an action–perception-related brain network to a higher degree than leader–leader dyads[227] (FIG. 5).

In summary, the dyadic tapping studies illuminate how musical interaction is guided by mutual reduction of prediction errors, in effect rendering them mutually predictable. They may serve as a model for how competence, social context and dyadic interactions depend 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 a window to perhaps the most challenging question about music: how music becomes meaningful[228,229]. Accordingly, we speculate that one of the things that makes music meaningful from a neuroscientific perspective is when musical interaction — in the form of listening to or making music together — 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.

Collective musical improvisation is a particularly demanding example of musical communication, in which musicians' predictive models need to be aligned to a large degree. This is found in many styles of music; for example, 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[230–233]. 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[234–236]. These include prefrontal brain regions, such as the presupplementary area, medial prefrontal cortex, inferior frontal gyrus, dorsolateral prefrontal cortex, dorsal premotor cortex and auditory cortices[237,238]. This is not surprising since improvisation involves several processes simultaneously. A jazz musician has to play, listen to what the other musicians are playing and evaluate how the music sounds as a whole while choosing which direction to take and generate new phrases to play next[239]. Therefore, the neuronal processes underlying musical improvisation must necessarily be predictive in nature and allow dynamic shifts between different networks and brain states[240].

Recent cutting-edge neuroimaging connectivity measures — built on whole-brain computational modelling[241] — have made it possible to understand the changing predictive brain states that 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 networks[230,237,242,243]. 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[244,245]. This is usually understood as carefully creating a sensorium in which the opportunity to resolve (that is, 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[246], 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[7] 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 with regard to melodic expectancy violation[12]. This is coupled with findings from resting-state fMRI that show that improvising musicians exhibit more distributed, globally connected cortical networks than classical musicians, who instead show higher within-network connectivity than the former[247], and morphometric findings of structural differences between these different groups of musicians[248,249]. In a recent study, the amplitudes of event-related potentials in response to chords that varied in expectancy were significantly correlated with behavioural measures of fluency and originality on a divergent thinking task, indicating a putative transfer effect of music skills to more domain-general processes[250]. 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 past 20 years of research into music in the brain has created a foundational understanding of how the brain processes music through predictive coding. The coming years could be dedicated to understanding the way music shapes social interactions and the role of predictive coding in creating shared meaning and perhaps even states of eudaimonia through music.

Eudaimonia
In Aristotelian ethics, refers to a life well lived or human flourishing, and in affective neuroscience, it is often used to describe meaningful pleasure.

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 of 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? As there would be no sensory information with which to compare the internal metre model, the PCM model 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 the different predictive frameworks in which musical events are embedded. It remains for future studies to clarify the interaction between melody, harmony and rhythm — for example, the influence of shifting tonalities or metrical displacement of a given melody — as well as the interaction between lyrics and melody. Whereas this Review has focused mainly 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 auditory prediction processes in the brain.

The PCM model offers a compelling but not exclusive framework for these endeavours. An alternative to the PCM model's probabilistic approach — of modelling hidden reference structures — is an oscillator-based approach simulating perception of metre and tonality in terms of the resonance of coupled nonlinear oscillators[97,102,251,252]. The oscillator approach gives greater weight to stimulus properties than to the top-down effects of learned musical experience. As a result, it will have difficulties in accounting for the range of musical phenomena that can be accounted for by the PCM model or other prediction-based approaches, such as the action simulation for auditory prediction hypothesis, which proposes that the motor system contributes to the accuracy of auditory predictions by providing a periodic temporal framework through these connections[215,253]. 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 hope that the PCM model will continue to shed light on the neural mechanisms underlying music perception, action, emotion and learning and that it will be useful in understanding prediction as a fundamental principle behind brain function. These insights may offer a new pathway to understanding how music becomes meaningful to the individual as well as in musical interactions between people.

1. Zatorre, R. J., Chen, J. L. & Penhune, V. B. When the brain plays music: auditory–motor interactions in music perception and production. *Nat. Rev. Neurosci.* **8**, 547–558 (2007).
A seminal review of auditory–motor coupling in music.
2. Koelsch, S. Toward a neural basis of music perception–a review and updated model. *Front. Psychol.* **2**, 110 (2011).
3. Maes, P. J., Leman, M., Palmer, C. & Wanderley, M. M. Action-based effects on music perception. *Front. Psychol.* **4**, 1008 (2014).
4. Koelsch, S. Brain correlates of music-evoked emotions. *Nat. Rev. Neurosci.* **15**, 170–180 (2014).
In this review, the author shows how music engages phylogenetically old reward networks in the brain to evoke emotions, and not merely subjective feelings.
5. 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).
6. Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).
This review posits that several global brain theories may be unified by the free-energy principle.
7. Koelsch, S., Vuust, P. & Friston, K. Predictive processes and the peculiar case of music. *Trends Cogn. Sci.* **23**, 63–77 (2019).
This review focuses specifically on predictive coding in music.
8. Meyer, L. *Emotion and Meaning in Music* (Univ. of Chicago Press, 1956).
9. Lerdahl, F. & Jackendoff, R. *A Generative Theory of Music* (MIT Press, 1999).
10. Huron, D. *Sweet Anticipation* (MIT Press, 2006).
In this book, Huron draws on evolutionary theory and statistical learning to propose a general theory of musical expectation.
11. Hansen, N. C. & Pearce, M. T. Predictive uncertainty in auditory sequence processing. *Front. Psychol.* https://doi.org/10.3389/fpsyg.2013.01008 (2014).
12. 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).
13. Altenmüller, E. O. How many music centers are in the brain? *Ann. N. Y. Acad. Sci.* **930**, 273–280 (2001).
14. Monelle, R. *Linguistics and Semiotics in Music* (Harwood Academic Publishers, 1992).
15. Rohrmeier, M. A. & Koelsch, S. Predictive information processing in music cognition. A critical review. *Int. J. Psychophysiol.* **83**, 164–175 (2012).
16. Vuust, P., Dietz, M. J., Witek, M. & Kringelbach, M. L. Now you hear it: a predictive coding model for understanding rhythmic incongruity. *Ann. N. Y. Acad. Sci.* https://doi.org/10.1111/nyas.13622 (2018).
17. Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C. & Roepstorff, A. Predictive coding of music–brain responses to rhythmic incongruity. *Cortex* **45**, 80–92 (2009).
18. 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).
This is the foundation for the PCM model used in this Review.
19. Garrido, M. I., Sahani, M. & Dolan, R. J. Outlier responses reflect sensitivity to statistical structure in the human brain. *PLoS Comput. Biol.* **9**, e1002999 (2013).
20. 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. *Soc. Cogn. Affect. Neurosci.* **13**, 877–888 (2018).
21. Quiroga-Martinez, D. R. et al. Musical prediction error responses similarly reduced by predictive uncertainty in musicians and non-musicians. *Eur. J. Neurosci.* https://doi.org/10.1111/ejn.14667 (2019).
22. Koelsch, S., Schröger, E. & Gunter, T. C. Music matters: preattentive musicality of the human brain. *Psychophysiology* **39**, 38–48 (2002).
23. 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).
24. 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. *Hum. Brain Mapp.* **42**, 941–952 (2020).
25. Lieder, F., Daunizeau, J., Garrido, M. I., Friston, K. J. & Stephan, K. E. Modelling trial-by-trial changes in the mismatch negativity. *PLoS Comput. Biol.* **9**, e1002911 (2013).
26. Wacongne, C., Changeux, J. P. & Dehaene, S. A neuronal model of predictive coding accounting for the mismatch negativity. *J. Neurosci.* **32**, 3665–3678 (2012).
27. 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).
28. Feldman, H. & Friston, K. J. Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* **4**, 215 (2010).
29. 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).
This fMRI study ties uncertainty and surprise to musical pleasure.
30. McDermott, J. H. & Oxenham, A. J. Music perception, pitch, and the auditory system. *Curr. Opin. Neurobiol.* **18**, 452–463 (2008).
31. Thoret, E., Caramiaux, B., Depalle, P. & McAdams, S. Learning metrics on spectrotemporal modulations reveals the perception of musical instrument timbre. *Nat. Hum. Behav.* **5**, 369–377 (2020).
32. Siedenburg, K. & McAdams, S. Four distinctions for the auditory "wastebasket" of timbre. *Front. Psychol.* **8**, 1747 (2017).
33. Bendor, D. & Wang, X. The neuronal representation of pitch in primate auditory cortex. *Nature* **436**, 1161–1165 (2005).
34. Zatorre, R. J. Pitch perception of complex tones and human temporal-lobe function. *J. Acoustical Soc. Am.* **84**, 566–572 (1988).
35. 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. USA* **100**, 10038–10042 (2003).
Using fMRI data, this study shows that pitch chroma is represented anterior to the primary auditory cortex, and pitch height is represented posterior to the primary auditory cortex.

36. 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).
37. 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).
38. Houde, J. F. & Chang, E. F. The cortical computations underlying feedback control in vocal production. *Curr. Opin. Neurobiol.* **33**, 174–181 (2015).
39. 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).
40. Janata, P. et al. The cortical topography of tonal structures underlying Western music. *Science* **298**, 2167–2170 (2002).
41. Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical learning by 8-month-old infants. *Science* **274**, 1926–1928 (1996).
42. 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).
43. Krumhansl, C. L. Perceptual structures for tonal music. *Music. Percept.* **1**, 28–62 (1983).
44. Margulis, E. H. A model of melodic expectation. *Music. Percept.* **22**, 663–714 (2005).
45. Temperley, D. A probabilistic model of melody perception. *Cogn. Sci.* **32**, 418–444 (2008).
46. Pearce, M. T. & Wiggins, G. A. Auditory expectation: the information dynamics of music perception and cognition. *Top. Cogn. Sci.* **4**, 625–652 (2012).
47. Sears, D. R. W., Pearce, M. T., Caplin, W. E. & McAdams, S. Simulating melodic and harmonic expectations for tonal cadences using probabilistic models. *J. N. Music. Res.* **47**, 29–52 (2018).
48. Näätänen, R., Gaillard, A. W. & Mäntysalo, S. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.* **42**, 313–329 (1978).
49. Näätänen, R., Paavilainen, P., Rinne, T. & Alho, K. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* **118**, 2544–2590 (2007).
This classic review covers three decades of MMN research to understand auditory perception.
50. Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C. & Vuust, P. The Musical Ear Test, a new reliable test for measuring musical competence. *Learn. Individ. Differ.* **20**, 188–196 (2010).
51. 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).
52. Burunat, I. et al. The reliability of continuous brain responses during naturalistic listening to music. *Neuroimage* **124**, 224–231 (2016).
53. Burunat, I. et al. Action in perception: prominent visuo-motor functional symmetry in musicians during music listening. *PLoS ONE* **10**, e0138238 (2015).
54. Alluri, V. et al. Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage* **59**, 3677–3689 (2012).
A free-listening fMRI study showing brain networks involved in perception of distinct acoustical features of music.
55. Halpern, A. R. & Zatorre, R. J. When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb. Cortex* **9**, 697–704 (1999).
56. 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).
57. Pallesen, K. J. et al. Emotion processing of major, minor, and dissonant chords: a functional magnetic resonance imaging study. *Ann. N. Y. Acad. Sci.* **1060**, 450–453 (2005).
58. McPherson, M. J. et al. Perceptual fusion of musical notes by native Amazonians suggests universal representations of musical intervals. *Nat. Commun.* **11**, 2786 (2020).
59. Helmholtz H. L. F. *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (Cambridge Univ. Press, 1954).
60. Vassilakis, P. N. & Kendall, R. A. in *Human Vision and Electronic Imaging XV*. 75270O (International Society for Optics and Photonics, 2010).
61. Plomp, R. & Levelt, W. J. M. Tonal consonance and critical bandwidth. *J. Acoustical Soc. Am.* **38**, 548–560 (1965).
62. 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).
An ethnomusicology study showing that consonance preference may be absent in people with minimal exposure to Western music.
63. Mehr, S. A. et al. Universality and diversity in human song. *Science* https://doi.org/10.1126/science.aax0868 (2019).
64. Patel, A. D., Gibson, E., Ratner, J., Besson, M. & Holcomb, P. J. Processing syntactic relations in language and music: an event-related potential study. *J. Cogn. Neurosci.* **10**, 717–733 (1998).
This classic study compares responses to syntactic incongruities in both language and Western tonal music.
65. Janata, P. The neural architecture of music-evoked autobiographical memories. *Cereb. Cortex* **19**, 2579–2594 (2009).
66. 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).
67. Koelsch, S. et al. Differentiating ERAN and MMN: an ERP study. *Neuroreport* **12**, 1385–1389 (2001).
Using EEG, the authors show that ERAN and MMN reflect different cognitive mechanisms.
68. Loui, P., Grent-'t-Jong, T., Torpey, D. & Woldorff, M. Effects of attention on the neural processing of harmonic syntax in Western music. *Cogn. Brain Res.* **25**, 678–687 (2005).
69. Koelsch, S., Fritz, T., Schulze, K., Alsop, D. & Schlaug, G. Adults and children processing music: an fMRI study. *Neuroimage* **25**, 1068–1076 (2005).
70. Tillmann, B., Janata, P. & Bharucha, J. J. Activation of the inferior frontal cortex in musical priming. *Ann. N. Y. Acad. Sci.* **999**, 209–211 (2003).
71. 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.* **1389**, 103–114 (2011).
72. Leino, S., Brattico, E., Tervaniemi, M. & Vuust, P. Representation of harmony rules in the human brain: further evidence from event-related potentials. *Brain Res.* **1142**, 169–177 (2007).
73. Sammler, D. et al. Co-localizing linguistic and musical syntax with intracranial EEG. *Neuroimage* **64**, 134–146 (2013).
74. Loui, P., Wessel, D. L. & Hudson Kam, C. L. Humans rapidly learn grammatical structure in a new musical scale. *Music. Percept.* **27**, 377–388 (2010).
75. Loui, P., Wu, E. H., Wessel, D. L. & Knight, R. T. A generalized mechanism for perception of pitch patterns. *J. Neurosci.* **29**, 454–459 (2009).
76. Cheung, V. K. M., Meyer, L., Friederici, A. D. & Koelsch, S. The right inferior frontal gyrus processes nested non-local dependencies in music. *Sci. Rep.* **8**, 3822 (2018).
77. Haueisen, J. & Knosche, T. R. Involuntary motor activity in pianists evoked by music perception. *J. Cogn. Neurosci.* **13**, 786–792 (2001).
78. Bangert, M. et al. Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage* **30**, 917–926 (2006).
79. Baumann, S. et al. A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Res.* **1161**, 65–78 (2007).
80. Lahav, A., Saltzman, E. & Schlaug, G. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J. Neurosci.* **27**, 308–314 (2007).
81. Bianco, R. et al. Neural networks for harmonic structure in music perception and action. *Neuroimage* **142**, 454–464 (2016).
82. Eerola, T., Vuoskoski, J. K., Peltola, H.-R., Putkinen, V. & Schäfer, K. An integrative review of the enjoyment of sadness associated with music. *Phys. Life Rev.* **25**, 100–121 (2018).
83. Huron, D. M. D. The harmonic minor scale provides an optimum way of reducing average melodic interval size, consistent with sad affect cues. *Empir. Musicol. Rev.* **7**, 15 (2012).
84. Huron, D. A comparison of average pitch height and interval size in major-and minor-key themes: evidence consistent with affect-related pitch prosody. **3**, 59-63 (2008).
85. Juslin, P. N. & Laukka, P. Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol. Bull.* **129**, 770 (2003).
86. Fritz, T. et al. Universal recognition of three basic emotions in music. *Curr. Biol.* **19**, 573–576 (2009).
87. London, J. *Hearing in Time: Psychological Aspects of Musical Meter* (Oxford Univ. Press, 2012).
88. Honing, H. Without it no music: beat induction as a fundamental musical trait. *Ann. N. Y. Acad. Sci.* **1252**, 85–91 (2012).
89. Hickok, G., Farahbod, H. & Saberi, K. The rhythm of perception: entrainment to acoustic rhythms induces subsequent perceptual oscillation. *Psychol. Sci.* **26**, 1006–1013 (2015).
90. Yabe, H., Tervaniemi, M., Reinikainen, K. & Näätänen, R. Temporal window of integration revealed by MMN to sound omission. *Neuroreport* **8**, 1971–1974 (1997).
91. 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).
92. 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. *Biol. Psychol.* **75**, 219–228 (2007).
93. 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).
94. 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. *Psychol. Sci.* **14**, 362–366 (2003).
95. Lerdahl, F. & Jackendoff, R. An overview of hierarchical structure in music. *Music. Percept.* **1**, 229–252 (1983).
96. Large, E. W. & Kolen, J. F. Resonance and the perception of musical meter. *Connect. Sci.* **6**, 177–208 (1994).
97. Large, E. W. & Jones, M. R. The dynamics of attending: how people track time-varying events. *Psychol. Rev.* **106**, 119–159 (1999).
98. Cutietta, R. A. & Booth, G. D. The influence of metre, mode, interval type and contour in repeated melodic free-recall. *Psychol. Music* **24**, 222–236 (1996).
99. Smith, K. C. & Cuddy, L. L. Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *J. Exp. Psychol.* **15**, 457–471 (1989).
100. Palmer, C. & Krumhansl, C. L. Mental representations for musical meter. *J. Exp. Psychol.* **16**, 728–741 (1990).
101. Einarson, K. M. & Trainor, L. J. Hearing the beat: young children's perceptual sensitivity to beat alignment varies according to metric structure. *Music. Percept.* **34**, 56–70 (2016).
102. Large, E. W., Herrera, J. A. & Velasco, M. J. Neural networks for beat perception in musical rhythm. *Front. Syst. Neurosci.* **9**, 159 (2015).
103. Nozaradan, S., Peretz, I., Missal, M. & Mouraux, A. Tagging the neuronal entrainment to beat and meter. *J. Neurosci.* **31**, 10234–10240 (2011).
104. Nozaradan, S., Peretz, I. & Mouraux, A. Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *J. Neurosci.* **32**, 17572–17581 (2012).
105. 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).
106. Lenc, T., Keller, P. E., Varlet, M. & Nozaradan, S. Neural and behavioral evidence for frequency-selective context effects in rhythm processing in humans. *Cereb. Cortex Commun.* https://doi.org/10.1093/texcom/tgaa037 (2020).
107. Jacoby, N. & McDermott, J. H. Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Curr. Biol.* **27**, 359–370 (2017).
108. Hannon, E. E. & Trehub, S. E. Metrical categories in infancy and adulthood. *Psychol. Sci.* **16**, 48–55 (2005).
109. Hannon, E. E. & Trehub, S. E. Tuning in to musical rhythms: infants learn more readily than adults. *Proc. Natl Acad. Sci. USA* **102**, 12639–12643 (2005).
110. 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).
111. Grahn, J. A. & Brett, M. Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* **19**, 893–906 (2007).
This fMRI study investigates participants listening to rhythms of varied complexity.

112. Toiviainen, P., Burunat, I., Brattico, E., Vuust, P. & Alluri, V. The chronnectome of musical beat. *Neuroimage* **216**, 116191 (2019).
113. 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).
114. Levitin, D. J., Grahn, J. A. & London, J. The psychology of music: rhythm and movement. *Annu. Rev. Psychol.* **69**, 51–75 (2018).
115. Winkler, I., Haden, G. P., Ladinig, O., Sziller, I. & Honing, H. Newborn infants detect the beat in music. *Proc. Natl Acad. Sci. USA* **106**, 2468–2471 (2009).
116. Phillips-Silver, J. & Trainor, L. J. Feeling the beat: movement influences infant rhythm perception. *Science* **308**, 1430–1430 (2005).
117. Cirelli, L. K., Trehub, S. E. & Trainor, L. J. Rhythm and melody as social signals for infants. *Ann. N. Y. Acad. Sci.* https://doi.org/10.1111/nyas.13580 (2018).
118. Cirelli, L. K., Einarson, K. M. & Trainor, L. J. Interpersonal synchrony increases prosocial behavior in infants. *Dev. Sci.* **17**, 1003–1011 (2014).
119. Repp, B. H. Sensorimotor synchronization: a review of the tapping literature. *Psychon. Bull. Rev.* **12**, 969–992 (2005).
120. Repp, B. H. & Su, Y. H. Sensorimotor synchronization: a review of recent research (2006-2012). *Psychon. Bull. Rev.* **20**, 403–452 (2013).
    This review, and Repp (2005), succinctly covers the field of sensorimotor synchronization.
121. Zarco, W., Merchant, H., Prado, L. & Mendez, J. C. Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *J. Neurophysiol.* **102**, 3191–3202 (2009).
122. 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. *Front. Neurosci.* **12**, 475 (2018).
123. Hattori, Y. & Tomonaga, M. Rhythmic swaying induced by sound in chimpanzees (Pan troglodytes). *Proc. Natl Acad. Sci. USA* **117**, 936–942 (2020).
124. Danielsen, A. *Presence and Pleasure. The Funk Grooves of James Brown and Parliament* (Wesleyan Univ. Press, 2006).
125. 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).
126. 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).
127. Janata, P., Tomic, S. T. & Haberman, J. M. Sensorimotor coupling in music and the psychology of the groove. *J. Exp. Psychol.* **141**, 54 (2012).
    Using a systematic approach, this multiple-studies article shows that the concept of groove can be widely understood as a pleasurable drive towards action.
128. 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. Percept.* **37**, 278–297 (2020).
129. Friston, K., Mattout, J. & Kilner, J. Action understanding and active inference. *Biol. Cybern.* **104**, 137–160 (2011).
130. Longuet-Higgins, H. C. & Lee, C. S. The rhythmic interpretation of monophonic music. *Music. Percept.* **1**, 18 (1984).
131. 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).
132. 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).
133. Kowalewski, D. A., Kratzer, T. M. & Friedman, R. S. Social music: investigating the link between personal liking and perceived groove. *Music. Percept.* **37**, 339–346 (2020).
134. 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. *Front. Neurosci.* **13**, 1039 (2019).
135. 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).
136. Matthews, T. E., Witek, M. A., Lund, T., Vuust, P. & Penhune, V. B. The sensation of groove engages motor and reward networks. *Neuroimage* **214**, 116768 (2020).
    This fMRI study shows that the sensation of groove engages both motor and reward networks in the brain.
137. 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).
138. 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. *J. Cogn. Neurosci.* **8**, 29–46 (1996).
139. Benadon, F. Meter isn't everything: the case of a timeline-oriented Cuban polyrhythm. *N. Ideas Psychol.* **56**, 100735 (2020).
140. 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).
141. Huron, D. Is music an evolutionary adaptation? *Ann. N. Y. Acad. Sci.* **930**, 43–61 (2001).
142. Koelsch, S. Towards a neural basis of music-evoked emotions. *Trends Cogn. Sci.* **14**, 131–137 (2010).
143. Eerola, T. & Vuoskoski, J. K. A comparison of the discrete and dimensional models of emotion in music. *Psychol. Music.* **39**, 18–49 (2010).
144. Lonsdale, A. J. & North, A. C. Why do we listen to music? A uses and gratifications analysis. *Br. J. Psychol.* **102**, 108–134 (2011).
145. Juslin, P. N. & Laukka, P. Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *J. N. Music. Res.* **33**, 217–238 (2004).
146. Huron, D. Why is sad music pleasurable? A possible role for prolactin. *Music. Sci.* **15**, 146–158 (2011).
147. Brattico, E. et al. It's sad but I like it: the neural dissociation between musical emotions and liking in experts and laypersons. *Front. Hum. Neurosci.* **9**, 676 (2015).
148. Sachs, M. E., Damasio, A. & Habibi, A. Unique personality profiles predict when and why sad music is enjoyed. *Psychol. Music* https://doi.org/10.1177/0305735620932660 (2020).
149. Sachs, M. E., Habibi, A., Damasio, A. & Kaplan, J. T. Dynamic intersubject neural synchronization reflects affective responses to sad music. *Neuroimage* **218**, 116512 (2020).
150. Juslin, P. N. & Vastfjall, D. Emotional responses to music: the need to consider underlying mechanisms. *Behav. Brain Sci.* **31**, 559–575 (2008).
    Using a novel theoretical framework, the authors propose that the mechanisms that evoke emotions from music are not unique to music.
151. Rickard, N. S. Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychol. Music.* **32**, 371–388 (2004).
152. 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. USA* **117**, 1924–1934 (2020).
153. Argstatter, H. Perception of basic emotions in music: culture-specific or multicultural? *Psychol. Music.* **44**, 674–690 (2016).
154. Stevens, C. J. Music perception and cognition: a review of recent cross-cultural research. *Top. Cogn. Sci.* **4**, 653–667 (2012).
155. Pearce, M. Cultural distance: a computational approach to exploring cultural influences on music cognition. in *Oxford Handbook of Music and the Brain* Vol. 31 (Oxford Univ. Press, 2018).
156. van der Weij, B., Pearce, M. T. & Honing, H. A probabilistic model of meter perception: simulating enculturation. *Front. Psychol.* **8**, 824 (2017).
157. Kringelbach, M. L. & Berridge, K. C. Towards a functional neuroanatomy of pleasure and happiness. *Trends Cogn. Sci.* **13**, 479–487 (2009).
158. 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. USA* **98**, 11818–11823 (2001).
    This seminal positron emission tomography study shows that the experience of musical chills correlates with activity in the reward system.
159. 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. *Phys. Life Rev.* **10**, 279–280 (2013).
160. Salimpoor, V. N. et al. Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science* **340**, 216–219 (2013).
161. 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).
162. 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).
163. 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).
164. 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. USA* **113**, E7337–E7345 (2016).
165. Gebauer, L. K., M., L. & Vuust, P. Musical pleasure cycles: the role of anticipation and dopamine. *Psychomusicology* **22**, 16 (2012).
166. Shany, O. et al. Surprise-related activation in the nucleus accumbens interacts with music-induced pleasantness. *Soc. Cogn. Affect. Neurosci.* **14**, 459–470 (2019).
167. 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? *J. Neurosci.* **39**, 9397–9409 (2019).
168. Swaminathan, S. & Schellenberg, E. G. Current emotion research in music psychology. *Emot. Rev.* **7**, 189–197 (2015).
169. Madison, G. & Schiölde, G. Repeated listening increases the liking for music regardless of its complexity: implications for the appreciation and aesthetics of music. *Front. Neurosci.* **11**, 147 (2017).
170. Corrigall, K. A. & Schellenberg, E. G. Liking music: genres, contextual factors, and individual differences. in *Art, Aesthetics, and the Brain* (Oxford Univ. Press, 2015).
171. Zentner, A. Measuring the effect of file sharing on music purchases. *J. Law Econ.* **49**, 63–90 (2006).
172. 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).
173. Vuust, P. et al. Personality influences career choice: sensation seeking in professional musicians. *Music. Educ. Res.* **12**, 219–230 (2010).
174. Rohrmeier, M. & Rebuschat, P. Implicit learning and acquisition of music. *Top. Cogn. Sci.* **4**, 525–553 (2012).
175. Münthe, T. F., Altenmüller, E. & Jäncke, L. The musician's brain as a model of neuroplasticity. *Nat. Rev. Neurosci.* **3**, 1–6 (2002).
    This review highlights how professional musicians represent an ideal model for investigating neuroplasticity.
176. Habibi, A. et al. Childhood music training induces change in micro and macroscopic brain structure: results from a longitudinal study. *Cereb. Cortex* **28**, 4336–4347 (2018).
177. Schlaug, G., Jancke, L., Huang, Y., Staiger, J. F. & Steinmetz, H. Increased corpus callosum size in musicians. *Neuropsychologia* **33**, 1047–1055 (1995).
178. Baer, L. H. et al. Regional cerebellar volumes are related to early musical training and finger tapping performance. *Neuroimage* **109**, 130–139 (2015).
179. 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).
180. Gaser, C. & Schlaug, G. Brain structures differ between musicians and non-musicians. *J. Neurosci.* **23**, 9240–9245 (2003).
    Using a morphometric technique, this study shows a grey matter volume difference in multiple brain regions between professional musicians and a matched control group of amateur musicians and non-musicians.
181. 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).
182. 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. *Cereb. Cortex* **27**, 2768–2778 (2017).

183. Schneider, P. et al. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat. Neurosci.* **5**, 688–694 (2002).
184. Bengtsson, S. L. et al. Extensive piano practicing has regionally specific effects on white matter development. *Nat. Neurosci.* **8**, 1148–1150 (2005).
185. 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. *Hum. Brain Mapp.* **38**, 4834–4849 (2017).
186. Kraus, N. & Chandrasekaran, B. Music training for the development of auditory skills. *Nat. Rev. Neurosci.* **11**, 599–605 (2010).
187. Koelsch, S., Schröger, E. & Tervaniemi, M. Superior pre-attentive auditory processing in musicians. *Neuroreport* **10**, 1309–1313 (1999).
188. Münte, T. F., Kohlmetz, C., Nager, W. & Altenmüller, E. Superior auditory spatial tuning in conductors. *Nature* **409**, 580 (2001).
189. Seppänen, M., Brattico, E. & Tervaniemi, M. Practice strategies of musicians modulate neural processing and the learning of sound-patterns. *Neurobiol. Learn. Mem.* **87**, 236–247 (2007).
190. Guillot, A. et al. Functional neuroanatomical networks associated with expertise in motor imagery. *Neuroimage* **41**, 1471–1483 (2008).
191. 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).
192. Vuust, P., Brattico, E., Seppänen, M., Näätänen, R. & Tervaniemi, M. Practiced musical style shapes auditory skills. *Ann. N. Y. Acad. Sci.* **1252**, 139–146 (2012).
193. Bangert, M. & Altenmüller, E. O. Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci.* **4**, 26 (2003).
194. Li, Q. et al. Musical training induces functional and structural auditory-motor network plasticity in young adults. *Hum. Brain Mapp.* **39**, 2098–2110 (2018).
195. 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. *Cereb. Cortex* **26**, 3125–3134 (2016).
196. Putkinen, V., Tervaniemi, M. & Huotilainen, M. Musical playschool activities are linked to faster auditory development during preschool-age: a longitudinal ERP study. *Sci. Rep.* **9**, 11310–11310 (2019).
197. 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. *Dev. Sci.* **17**, 282–297 (2014).
198. Jentschke, S. & Koelsch, S. Musical training modulates the development of syntax processing in children. *Neuroimage* **47**, 735–744 (2009).
199. 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. *Cereb. Cortex* **24**, 956–967 (2014).
200. Moreno, S. et al. Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb. Cortex* **19**, 712–723 (2009).
201. 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).
202. 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).
203. Virtala, P. & Partanen, E. Can very early music interventions promote at-risk infants' development? *Ann. N. Y. Acad. Sci.* **1423**, 92–101 (2018).
204. Flaugnacco, E. et al. Music training increases phonological awareness and reading skills in developmental dyslexia: a randomized control trial. *PLoS ONE* **10**, e0138715 (2015).
205. 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).
206. Schellenberg, E. G. Correlation = causation? music training, psychology, and neuroscience. *Psychol. Aesthet. Creat. Arts* **14**, 475–480 (2019).
207. Sala, G. & Gobet, F. Cognitive and academic benefits of music training with children: a multilevel meta-analysis. *Mem. Cogn.* **48**, 1429–1441 (2020).
208. Saffran, J. R. Musical learning and language development. *Ann. N. Y. Acad. Sci.* **999**, 397–401 (2003).
209. Friston, K. The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* **13**, 293–301 (2009).
210. Pearce, M. T. Statistical learning and probabilistic prediction in music cognition: mechanisms of stylistic enculturation. *Ann. N. Y. Acad. Sci.* **1423**, 378–395 (2018).
211. Novembre, G., Knoblich, G., Dunne, L. & Keller, P. E. Interpersonal synchrony enhanced through 20 Hz phase-coupled dual brain stimulation. *Soc. Cogn. Affect. Neurosci.* **12**, 662–670 (2017).
212. Konvalinka, I. et al. Frontal alpha oscillations distinguish leaders from followers: multivariate decoding of mutually interacting brains. *Neuroimage* **94C**, 79–88 (2014).
213. Novembre, G., Mitsopoulos, Z. & Keller, P. E. Empathic perspective taking promotes interpersonal coordination through music. *Sci. Rep.* **9**, 12255 (2019).
214. Wolpert, D. M., Ghahramani, Z. & Jordan, M. I. An internal model for sensorimotor integration. *Science* **269**, 1880–1882 (1995).
215. 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).
216. Sebanz, N. & Knoblich, G. Prediction in joint action: what, when, and where. *Top. Cogn. Sci.* **1**, 353–367 (2009).
217. Friston, K. J. & Frith, C. D. Active inference, communication and hermeneutics. *Cortex* **68**, 129–143 (2015).
This article proposes a link between active inference, communication and hermeneutics.
218. Konvalinka, I., Vuust, P., Roepstorff, A. & Frith, C. D. Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Q. J. Exp. Psychol.* **63**, 2220–2230 (2010).
219. Wing, A. M. & Kristofferson, A. B. Response delays and the timing of discrete motor responses. *Percept. Psychophys.* **14**, 5–12 (1973).
220. Repp, B. H. & Keller, P. E. Sensorimotor synchronization with adaptively timed sequences. *Hum. Mov. Sci.* **27**, 423–456 (2008).
221. Vorberg, D. & Schulze, H.-H. Linear phase-correction in synchronization: predictions, parameter estimation, and simulations. *J. Math. Psychol.* **46**, 56–87 (2002).
222. 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).
Using dual-EEG, the authors propose alpha oscillations as a candidate for regulating the balance between internal and external information in joint action.
223. 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).
224. 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).
225. Heggli, O. A., Konvalinka, I., Kringelbach, M. L. & Vuust, P. Musical interaction is influenced by underlying predictive models and musical expertise. *Sci. Rep.* **9**, 1–13 (2019).
226. Heggli, O. A., Cabral, J., Konvalinka, I., Vuust, P. & Kringelbach, M. L. A Kuramoto model of self-other integration across interpersonal synchronization strategies. *PLoS Comput. Biol.* **15**, e1007422 (2019).
227. Heggli, O. A. et al. Transient brain networks underlying interpersonal strategies during synchronized action. *Soc. Cogn. Affect. Neurosci.* **16**, 19–30 (2020).
This EEG study shows that differences in interpersonal synchronization are reflected by activity in a temporoparietal network.
228. Patel, A. D. *Music, Language, and the Brain* (Oxford Univ. Press, 2006).
229. Molnar-Szakacs, I. & Overy, K. Music and mirror neurons: from motion to 'e'motion. *Soc. Cogn. Affect. Neurosci.* **1**, 235–241 (2006).
230. Beaty, R. E., Benedek, M., Silvia, P. J. & Schacter, D. L. Creative cognition and brain network dynamics. *Trends Cogn. Sci.* **20**, 87–95 (2016).
231. Limb, C. J. & Braun, A. R. Neural substrates of spontaneous musical performance: an FMRI study of jazz improvisation. *PLoS ONE* **3**, e1679 (2008).
232. Liu, S. et al. Neural correlates of lyrical improvisation: an FMRI study of freestyle rap. *Sci. Rep.* **2**, 834 (2012).
233. Rosen, D. S. et al. Dual-process contributions to creativity in jazz improvisations: an SPM-EEG study. *Neuroimage* **213**, 116632 (2020).
234. Boasen, J., Takeshita, Y., Kuriki, S. & Yokosawa, K. Spectral-spatial differentiation of brain activity during mental imagery of improvisational music performance using MEG. *Front. Hum. Neurosci.* **12**, 156 (2018).
235. Berkowitz, A. L. & Ansari, D. Generation of novel motor sequences: the neural correlates of musical improvisation. *Neuroimage* **41**, 535–543 (2008).
236. Loui, P. Rapid and flexible creativity in musical improvisation: review and a model. *Ann. N. Y. Acad. Sci.* **1423**, 138–145 (2018).
237. Beaty, R. E. The neuroscience of musical improvisation. *Neurosci. Biobehav. Rev.* **51**, 108–117 (2015).
238. Vuust, P. & Kringelbach, M. L. Music improvisation: a challenge for empirical research. in *Routledge Companion to Music Cognition* (Routledge, 2017).
239. Norgaard, M. Descriptions of improvisational thinking by artist-level jazz musicians. *J. Res. Music. Educ.* **59**, 109–127 (2011).
240. Kringelbach, M. L. & Deco, G. Brain states and transitions: insights from computational neuroscience. *Cell Rep.* **32**, 108128 (2020).
241. Deco, G. & Kringelbach, M. L. Hierarchy of information processing in the brain: a novel 'intrinsic ignition' framework. *Neuron* **94**, 961–968 (2017).
242. 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).
243. 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).
244. 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).
245. Beaty, R. E. et al. Robust prediction of individual creative ability from brain functional connectivity. *Proc. Natl Acad. Sci. USA* **115**, 1087–1092 (2018).
246. 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).
247. Belden, A. et al. Improvising at rest: differentiating jazz and classical music training with resting state functional connectivity. *Neuroimage* **207**, 116384 (2020).
248. Arkin, C., Przysinda, E., Pfeifer, C. W., Zeng, T. & Loui, P. Gray matter correlates of creativity in musical improvisation. *Front. Hum. Neurosci.* **13**, 169 (2019).
249. 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. *Sci. Rep.* **6**, 20482 (2016).
250. Przysinda, E., Zeng, T., Maves, K., Arkin, C. & Loui, P. Jazz musicians reveal role of expectancy in human creativity. *Brain Cogn.* **119**, 45–53 (2017).
251. Large, E. W., Kim, J. C., Flaig, N. K., Bharucha, J. J. & Krumhansl, C. L. A neurodynamic account of musical tonality. *Music. Percept.* **33**, 319–331 (2016).
252. Large, E. W. & Palmer, C. Perceiving temporal regularity in music. *Cogn. Sci.* **26**, 1–37 (2002).
This article proposes an oscillator-based approach for the perception of temporal regularity in music.
253. Cannon, J. J. & Patel, A. D. How beat perception co-opts motor neurophysiology. *Trends Cogn. Sci.* **25**, 137–150 (2020).
The authors propose that cyclic time-keeping activity in the supplementary motor area, termed 'proto-actions', is organized by the dorsal striatum to support hierarchical metrical structures.
254. 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 Cross, E. S. & Obhi, S. S.) 280-310 (Cambridge Univ. Press, 2016).
255. 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).
256. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204 (2013).
257. Kahl, R. *Selected Writings of Hermann Helmholtz* (Wesleyan Univ. Press, 1878).
258. Gregory, R. L. Perceptions as hypotheses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **290**, 181–197 (1980).

259. Gibson, J. J. *The Ecological Approach to Visual Perception* (Houghton Mifflin, 1979).
260. Fuster, J. *The Prefrontal Cortex Anatomy, Physiology and Neuropsychology of the Frontal Lobe* (Lippincott-Raven, 1997).
261. Neisser, U. *Cognition and Reality: Principles and Implications of Cognitive Psychology* (W H Freeman/ Times Books/ Henry Holt & Co, 1976).
262. Arbib, M. A. & Hesse, M. B. *The Construction of Reality* (Cambridge Univ. Press, 1986).
263. Cisek, P. & Kalaska, J. F. Neural mechanisms for interacting with a world full of action choices. *Annu. Rev. Neurosci.* **33**, 269–298 (2010).
264. Isomura, T., Parr, T. & Friston, K. Bayesian filtering with multiple internal models: toward a theory of social intelligence. *Neural Comput.* **31**, 2390–2431 (2019).
265. Friston, K. & Frith, C. A duet for one. *Conscious. Cogn.* **36**, 390–405 (2015).
266. Hunt, B. R., Ott, E. & Yorke, J. A. Differentiable generalized synchronization of chaos. *Phys. Rev. E* **55**, 4029–4034 (1997).
267. Ghazanfar, A. A. & Takahashi, D. Y. The evolution of speech: vision, rhythm, cooperation. *Trends Cogn. Sci.* **18**, 543–553 (2014).
268. Wilson, M. & Wilson, T. P. An oscillator model of the timing of turn-taking. *Psychon. Bull. Rev.* **12**, 957–968 (2005).

Acknowledgements

Funding was provided by The Danish National Research Foundation (DNRF117). The authors thank E. Altenmüller and D. Huron for comments on early versions of the manuscript.

Author contributions

The authors contributed equally to all aspects of this article.

Competing interests

The authors declare no competing interests.

Peer review information

Nature Reviews Neuroscience thanks D. Sammler and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note