What is “Predictive Coding”?

In recent years, predictive processes have been recognized as an integral part of both perception and action [1,2]. Predictive coding (PC) furnishes a compelling framework for understanding such processes in audition and language comprehension [3-6]. According to the predictive coding framework, backward predictions are passed down cortical hierarchies to resolve or ‘explain away’ **prediction errors** (see Glossary) at lower levels, i.e., sensory information that does not match a prediction (Figure 1, Key Figure). These prediction errors then ascend the hierarchy to evince better predictions [7]. With regard to auditory perception, the levels of this hierarchy could comprise the auditory brainstem and thalamus, primary auditory cortex, auditory association cortex, and frontal cortex. This distributed system is part of a hierarchy that ultimately subsumes other systems such as sensorimotor, visceromotor (autonomic), and memory systems. By formulating motor and autonomic reflexes as the resolution of prediction errors, one arrives at **active inference**; namely, an enactive generalization of PC. This generalization means that perceptual processes are not just passive and bottom-up but are constructive, where perception is actively driven by top-down processes. Thus, active inference implies that *cognitive* processes on the one hand (e.g., prediction, planning, action, learning and the like) and *perceptual* processes on the other are not serial or separable, but intertwined in a hierarchical cascade of prediction error reducing dynamics.

Music offers a most illuminating paradigm to understand the fundaments of the predictive brain, largely because every type of music is based on predictable regularities, e.g., temporal, melodic, and – depending on the musical culture or musical tradition – harmonic, timbral, and textural structure [8]. In contrast to an unpredictable (e.g., random) sensory stream, musical structure affords competing hypotheses or predictions about possible outcomes [9] and then dispels uncertainty by confirming a particular prediction, given a particular musical tradition or culture. We will first consider the most straightforward example of musical predictions, namely predictions of musical events in time (i.e., predictions about when musical events occur). We then turn to the conundrum that acoustically irregular events, as well as irregular harmonies, elicit error signals even when we know in advance they will occur. Finally, we suggest how this conundrum can be explained by considering how second-order predictions (i.e., predictions about predictability) can modulate the processing of error signals on different levels of the processing hierarchy.

Predictions and the drive to move to a musical groove

In most musical traditions, the temporal structure of music is based on a pulse. The perception of a musical pulse usually engenders a synchronization of movements to that pulse; e.g., by tapping the foot or moving the entire body [10]. Usually, this pulse is hierarchically structured into the so-called *meter*, comprising evenly spaced and differentially accented beats [11-14] (i.e. 'strong' and 'weak' beats), which provides a prior context for perceiving the succession of musical events over time (e.g. 'Waltz' vs. 'March') [15-18]. However, music features events that challenge metrical expectations, e.g. in the case of syncopations (i.e. musical accents falling outside of the musical pulse). Curiously, such syncopations lead to a drive to move ('wanting to move'), especially when used repeatedly. This creates so-called musical *grooves* that are typical of contemporary dance-related music. Given a constant tempo, there is an inverted U-shaped relationship (reminiscent of the *Wundt-curve* [19,20]) between the amount of syncopation in a given rhythm, our 'wanting to move' – and accompanying feelings of pleasure [21] (see **Box 1**).

This special aspect of music perception appears to entail an active engagement with the sensorium – both in terms of proprioception (i.e., wanting to move) and interoception (i.e., feeling pleasure) [22].

Clearly, generating music is quintessentially enactive. Here, we go further and suggest that musical *perception* itself is an active process, much like the move from perceptual synthesis in vision to the notion of active vision [23,24] or active sensing [25]. The explanation we offer for this phenomenon is that, much like language, we predict music in terms of how we might generate it ourselves. On this view, we may feel the drive to move our bodies to the metrical beat to establish a metric model that generates the right sort of auditory predictions. Put simply, the 'act of listening' entails an internal generation of music (e.g. singing or playing an instrument), in which overt action is suppressed by ignoring evidence that one is not actually playing or singing, while attending to the auditory consequences of the internal act [26]. In short, we might generate sensorimotor predictions of hand or foot-movements – without actually moving – when establishing a cognitive representation of the musical pulse; especially when this representation is challenged by syncopations (i.e., prediction errors).

To understand 'ignoring evidence' and 'attention' in predictive coding, one has to make a key distinction between predictions of *content* and predictions of *context*, associating content with expectations about what will be heard and context with the **precision** or confidence placed in those expectations (statistically speaking, this distinction is between the first and second-order statistical moments of probabilistic predictions – see Figure 1). In other words, there is a fundamental distinction between being able to predict *what* will be heard next and predictions of *whether* it is possible to predict that content with any precision. For example, when reading the sentence

"The hills are alive with the sound of

of music."

we predict the content of the last word "music". However, we also predict its predictability in the sense that we believe any other word is highly unlikely. Conversely, when reading an unfamiliar sentence such as:

"Her resolved failed when she heard the sound of

of music."

one's prediction of predictability changes profoundly; in the sense the final word could have been sampled from a larger repertoire of plausible denouements (e.g., "crying", "war", "rain", etc.). Because the last word now resolves more uncertainty – about what it could have been – we deploy more attention to the final word. In predictive coding terms, this involves affording sensory prediction errors more precision or weight, to complement the loss of precision in our predictions. Indeed, this selective attention to uncertainty resolving cues may have caused you to miss the repeated "of" in both sentences – because it is not salient.

Similarly, when listening to tonal music, at the end of a cadence (e.g., after a subdominant and a dominant seventh chord in root position), I know that the next chord will either be a tonic, or a submediant, or another chord. This is the prediction of content. In addition, I know that by far the most likely chord to follow is the tonic, thus I expect the occurrence of the tonic, and I am quite confident that the tonic will occur. In other words, I can make a prediction about my prediction: I can predict that my prediction is likely to be correct. This is the precision, or prediction of context. A high precision (i.e. low **entropy** or uncertainty) thus means that an event is on average predictable.

In predictive coding, ascending prediction errors are therefore assigned a precision, such that they have the right sort of influence on higher levels of processing. Precise prediction errors induce belief updating higher in the hierarchy, whereas imprecise prediction errors are effectively ignored (i.e., when we expect them to be imprecise or unreliable, such as in a noisy bar). Heuristically, precision therefore selects salient sensory information that resolves uncertainty. Crucially, this precision weighting brings something new to the game. Specifically, *precision itself has to be predicted*. This means that there are two sorts of *descending predictions:* (first-order) predictions of perceptual content and (second-order) predictions about the precision that should be ascribed to first-order predictions. Heuristically, the predicted precision corresponds to predictions about the (second-order) statistics of perceptual (first-order) content, thus to 'expected uncertainty' or the 'known unknowns': see [22,27]. The key insight here is that a fully-fledged predictive coding scheme must be equipped with models that generate both first-order predictions (content) and second-order predictions (precision).

Computationally, ascribing a high precision to prediction errors increases their gain (c.f., attentional gain), so that they have a greater effect on subsequent processing. Physiologically, this gain control is thought to be mediated by neuromodulatory mechanisms that control the postsynaptic excitability of neuronal populations encoding prediction error. These modulatory effects are indicated schematically in Figure 1 using blue connections. Psychologically, this 'selection' means we have attentional control over what features to select, thereby equipping us with the active choice to attend or ignore prediction errors at different levels in the auditory hierarchy.

Predictive processes and the mismatch negativity

The distinction between predictions of content and precision is nicely illustrated by studies of the mismatch negativity (MMN), when subjects are aware that they are exposed to an unpredictable auditory context. The MMN is a specific brain-electric response that is elicited by auditory 'deviants' or 'oddballs' that are presented among a series of repeating standard sounds. For example, in a sequence of tones like ccccdcccccdccc, the d's elicit an MMN. Interestingly, several studies suggest the MMN is unaffected when people know a deviant tone is about to occur. For example, a visual cue – signaling the occurrence of a duration [28] or pitch [29] deviant – does not influence the amplitude (nor the latency) of the MMN. Even when participants generate deviant tones themselves (by pressing a button), the amplitude and latency of the MMN remain unchanged [30]. Thus, it appears that, with regard to the MMN, prior knowledge or beliefs do *not* modulate the processing of prediction errors [31,32]. On the other hand, the MMN is followed by later positive electrophysiological responses when the deviants automatically attract attention (eliciting a P3a – an electrophysiological response that characteristically occurs around 250-300 ms at anterior electrodes), or when individuals consciously detect the deviants (eliciting a P3b that typically occurs around 300 ms or later at posterior electrodes). Because the P3a and P3b have longer latencies than the MMN, they reflect processes originating from higher levels in the auditory processing hierarchy (see Figure 1 and [33]). These processes are clearly influenced by prior knowledge: for example, it has been demonstrated that prior knowledge significantly reduced the P3b [28] and a significant P3a was only observed when deviants occurred without prior warning [29].

Within the PC framework, we suggest a novel formulation of predictive (i.e., precision) filtering to explain these results; specifically, the conundrum that the prediction error signals reflected in the MMN are not modulated by prior knowledge. With repeated standards, sensory learning (i.e., short-term plasticity) changes both the predictions of content and precision of those predictions. Under the architecture in Figure 1, neuronal populations encoding the auditory standard are released from (lateral) precision constraints, so that prediction errors engendered by standard stimuli can eliminate themselves more efficiently, via excitation of the appropriate representation (i.e., expectation). This elimination has two components. First, sensory learning enables more accurate predictions of content. Second, the precision of lateral constraints decreases (or the precision of ascending prediction errors increases), such that the ascending prediction errors selectively engage the veridical representation. In the classical MMN literature this dual aspect of sensory learning is closely related to *model adjustment* and *adaptation* mechanisms, respectively [33,34]. When an oddball is encountered, the ensuing prediction errors 'fall on deaf ears', because the corresponding representations are attenuated with high precision. The prediction error cannot resolve itself and will keep 'knocking on the door'; thereby eliciting an MMN. Please see Figure 1 for a detailed explanation.

In terms of the P3b, electrophysiological studies [35,36] and predictive coding simulations [37] suggest that these long latency waveforms reflect a change in the context or predictability (i.e., a change in predictions of precision). The ensuing changes in precision are manifest on subsequent trials, where the MMN amplitude is significantly attenuated when the oddball is repeated immediately [38]. In other words, alternative representations are released from precision constraints and become available to explain away prediction errors elicited by novel stimuli. This is consistent with the observation that prior knowledge attenuates the P3b in the above experiments. In short, an unpredicted stimulus may be surprising in terms of its content but not in terms of the context – a stimulus can be predictably unpredictable.

At first glance, this presents a problem for our account of music perception. In line with the pre-attentive nature of the MMN, it seems as if early sensory processing is impervious to attention. In other words, just knowing about an impending deviant does not necessarily affect our predictive processing: we have to first experience the irregularity before updating our beliefs about precision (and subsequent attenuation of the MMN). The resistance of the classical MMN to explicit knowledge is consistent with the fact that we cannot willfully attenuate some forms of sensory processing; for example, we cannot reverse saccadic suppression during eye movements (for further discussion see [37,39,40]).

So where is the role for attention (i.e., mental action) in music perception? In what follows, we ask whether there is evidence for attentional modulation of early auditory processing when prior beliefs about predictability are not confounded by sensory learning (as in the oddball paradigm). In brief, we consider the early right anterior negativity (ERAN – an early response to music-syntactic irregularities) that has a remarkable behaviour. Crucially, the irregularities eliciting the ERAN rest upon (usually implicit) knowledge of **musical syntax**. Acquisition of such knowledge requires substantial exposure to music [41], whereas the events eliciting the MMN are inferred in real-time from the acoustic environment. We will show that the MMN and ERAN dissociate in two revealing ways: In contrast to the MMN, which attenuates with repeated acoustical irregularities, and which is relatively impervious to attention or predictions of predictability, the ERAN persists with repeated exposure to syntactical irregularities, but appears to be somewhat sensitive to knowledge about impending outcomes.

Predictive processes in the perception of harmony

Imagine a musical passage such as the beginning of Mozart's Symphony No. 31 (**Figure 2**). Because we are familiar with tonal music (and thus the major scale), the last note of the scale is both predicted and predictable. This means we can predict it with high precision (i.e., with high confidence), given the preceding context of a scale. Later in the movement (after the beginning of the development section), the same passage is modified in a way that the last note of the scale is out of key – and is therefore irregular or improbable given the preceding syntactic context (see also **Figure 2, and Supplementary Sound Files 1 & 2**). Several *cognitive* and *emotional* effects are elicited by this irregularity in anyone familiar with tonal music; c.f., the evocative use of an irregular or low-frequency word in a sentence. The *emotional effects* of such irregular events, e.g. surprise, anticipation, and tension, have a strong tradition in music psychology [9,42-44]. However, we first focus on *cognitive aspects* of harmonic-predictive processes and return to emotional concomitants later.

Music-syntactically irregular events (such as those shown in Figures 2 & 3) elicit an early right anterior negativity (ERAN), or 'music-syntactic MMN' [45-54]. In contrast to the classical MMN [28-30] – that depends on regularities in ongoing auditory input – the ERAN depends on syntactic knowledge that transcends current auditory sensations. This is because music-syntactic regularities are represented in long-term memory. As with the MMN, the amplitude of the ERAN does not change when individuals know that a syntactic irregularity is pending. For example, in a supervised learning paradigm [55], individuals were presented 10 times with stimuli similar to those described in Figure 2a (embedded in longer phrases): 5 stimuli were regular and 5 featured a music-syntactically irregular harmony. Participants were always told whether the next stimulus would be regular or irregular. Figure 2b shows that irregular events elicited an ERAN of typical amplitude and frontal distribution – despite the fact that subjects knew these events would be irregular. The ERAN was followed by a P3a and a P3b (Figure 2b). The amplitudes of ERAN, P3a and P3b were calculated separately for the first presentation of an irregular event, the second presentation, and so forth. Whereas P3a and P3b declined systematically across presentations, there was no systematic attenuation of the ERAN amplitude (Figure 2c).

Even if musicians are presented repeatedly with only two chord sequences (presented in different tonal keys) – ending either on a regular harmony (a tonic following a dominant-seventh chord) or on an irregular harmony (a double dominant following a dominant-seventh chord, see Figure 3 and Supplementary Sound Files 3 & 4) – the ERAN amplitude remains virtually unchanged [56]. In this study, musicians were told (in 144 trials) whether the final chord of a sequence would be regular or irregular, and (in 144 different trials) they were not informed. When musicians were not told about the sequence ending, a typical ERAN was elicited, which did not differ in amplitude from the ERAN elicited when participants knew how the sequence would end (Figure 3a & b, recall that only the two sequence types shown in Figure 3, transposed to different tonal keys, were used in this experiment). In contrast, later positive potentials differed between conditions: when participants had no knowledge about the sequence ending, the irregular chords elicited a P3a (but no P3b), and when participants were told about the sequence endings, irregular endings elicited a more parietal, P3b-like potential (but no P3a; see also Figure 3a & b).

Although the ERAN amplitude was unaffected by prior knowledge, the peak latency of the ERAN was about 10 ms shorter when participants were informed about the upcoming chord and could therefore predict the (regular or irregular) sequence ending (between 160-200 ms), compared to when participants could not predict the ending (between 150-190 ms). A similar (though smaller) latency difference was observed in an independent group of non-musicians (see bar chart in Figure 3). The important result here is that knowing about an impending syntactic deviant reduced the latency of prediction error responses. This is important because the speed of evidence accumulation is determined by the precision afforded to sensory prediction errors [57-59]. In turn, the above latency reduction suggests that prior beliefs about context (here a precise and unambiguous irregularity) *can* increase sensory precision – and evidently do so in musical processing (see also [60]).

In summary, the MMN and ERAN dissociate in two revealing ways. The MMN attenuates with repeated acoustical (non-syntactical) irregularities. Furthermore, the MMN is relatively impervious to attention or predictions of predictability (i.e., predictions of precision or context). In contrast, the ERAN persists with repeated exposure to syntactical irregularities – that are sampled from a limited and learned repertoire (e.g., a repertoire of harmonic successions with different degrees of regularity). Furthermore, the latency of the ERAN appears to be sensitive to attentional set or knowledge about impending outcomes.

On the predictive coding view, this is entirely sensible given that knowing a sequence of repeating auditory events can be violated by any kind of oddball does not help predict the nature of the violation (because there are too many possibilities). Conversely, knowing that an irregularity of sequential musical structure is approaching, enables the brain to selectively attend to (i.e., afford precision to) auditory features that will resolve uncertainty about the experienced sensory sequence (e.g., an irregular ending). This distinction rests upon long-term knowledge about musical syntax – enabling selective modulation by attentional set – and constitutes electrophysiological evidence for an active (attentional) process in listening to music.

Precision, attention and mental action

In the following, we unpack the relationship between mental action, selective attention and predictions of precision or predictability. In the studies reported above [55,56] the genesis of predictions about precision was ascribed to the P3b, because the P3b indexes changes in predictability (see also Figure 2c). The effects of these precision predictions may then be manifest in terms of the modulation of early (sensory) prediction error or deviant responses. The influence of precision on the processing of syntactic error signals, and its hierarchical balance is illustrated in Figure 4.

This formulation enjoys support from a recent study [45], in which the *precision* of predictions was manipulated experimentally. Participants listened to melodies with or without out-of-key notes and were asked if the melody contained an anomalous note – and how confident they were in their judgment. In some blocks, participants were – unbeknownst to them – provided with random feedback about their judgments. Consequently, participants were less accurate and had lower confidence in their judgments. Crucially, in this low confidence condition, the amplitude of the ERAN increased. This suggests that while the predictive music-syntactic processes reflected in the ERAN amplitude are not modulated by first order predictions, they are in fact modulated by second order predictions about the precision that should be ascribed to the first-order predictions. Furthermore – on the precision account of attention – an increase in the amplitude of the ERAN speaks to an appropriate revision of attentional set.

The generalization of predictive coding to incorporate predicted precision is brought into sharp focus when we consider the perception of music. Although the brain is in the game of minimizing prediction errors, we find the violations of musical predictions appealing when listening to musical compositions. That is, we appear to actively seek out or attend to musical sensations that are, by their very nature, predictably surprising (mathematically, these events have a high **salience**, information gain or relative entropy). The concomitant emotional effects of predictions, their precision, violation, and fulfillment are summarized under the term *musical tension*, which is a central principle underlying the evocation of emotion in Western music [61]. Musical tension highlights an apparent paradox in music perception: we derive pleasure from musical prediction errors – even if we know a musical piece – because they invariably resolve uncertainty about what we might have heard.

The epistemic offering of music

Because behaviors or policies that minimize expected **surprise** or prediction errors are the policies we select, one could wonder why we choose to listen to music featuring expectancy violations. This apparent paradox can easily be resolved under active inference. Expected surprise (or expected information content) is, mathematically, uncertainty or entropy. This means we should sample or attend to salient sensory cues that we expect to reduce uncertainty (in the visual search and saliency literature, this is known as **Bayesian surprise**). When listening to music, we entertain a number of predictions, or hypotheses, about future musical events (e.g., in terms of meter, rhythm, melody and harmony) [22,27], which are resolved in the near future, usually within the next few tones. Music thus provides the opportunity to continuously resolve uncertainty over such hypotheses (c.f., perception as hypothesis testing [62]). This *epistemic offering* of music is referred to as epistemic affordance, or **epistemic value** in active inference [26]; namely, the intrinsic value or motivation to resolve uncertainty [63-65] (note that we do not use the term "epistemic" in the tradition of the philosophy of the mind, but with regard to perceptual information that confirms or refutes competing hypotheses).

Clearly, listening to music is not an overt behavior; however, deploying attention or precision can be regarded as a mental action [66,67] . On this view, attending to a predictable syntactic irregularity could be regarded as a covert 'auditory saccade'. This suggests that musical tension calls upon exactly the same principles that underlie the epistemic foraging seen in overt exploratory behavior [68]. It is tempting to speculate that the epistemic offering of music – i.e., the recurrent resolution of uncertainty – activates reward-networks that underwrite the pleasure induced by listening to music [9,44] (**Box 2** deals with the question how auditory predictions can influence affective, or interoceptive precision; see also **Outstanding Questions Box**). This formulation would predict that even prediction errors elicit activity in reward-networks, e.g. because they help to improve our learned model of musical regularities. Interestingly, fluctuations in the (un)certainty of predictions create an 'entropic flux' [9] that contributes to the aesthetic appreciation of music. That is, it appears that we appreciate the epistemic offering of music in particular when it is provided in an aesthetically pleasing flux of predictable unpredictability or entropy.

Technically, expectations about precision have always been an integral part of predictive coding (e.g., the Kalman gain in Kalman filtering formulations of predictive coding). As noted above, they represent a key aspect of the functional architectures that may be involved in attention [69], motor control [70], and interoception [71-74]. The argument here is that a substantial part of perception (and music is an exemplary case) calls on the deployment of sensory precision (associated with attentional gain) via a sophisticated generative model of when and where precision should be deployed. This becomes particularly relevant if we want to understand the interplay between predictions at a subjective (i.e., consciously controlled), in relation to predictions at a sensory (i.e., subpersonal) level. For example, what sorts of predictions characterize the perceptual set induced when participants are told that they will hear an irregular chord at the end of the sequence (as in [55,56])? Participants may not predict exactly *what* they will hear; however, they can predict *that* some precise, uncertainty reducing sensory input is impending. This opportunity to resolve uncertainty augments the precision (and thus attentional gain) of sensory prediction errors; accelerating their accumulation and reducing the latency of prediction error (ERAN) responses. This ‘precision engineered’ aspect of predictive coding, we submit, explains a paradoxical aspect of predictive coding in music; namely, the allure of the predictably implausible.

Concluding Remarks

In conclusion, one of the major allures of music rests upon its special epistemic offering. This involves an enactive view of music perception, in which mental action selects the right kind of attentional set that optimizes the accumulation of sensory evidence – to resolve uncertainty about the inferred musical narrative. The concomitant belief updating can be cast in terms of attentional or precision control and the selection of attentional sets with epistemic value, which appears to fit comfortably with electrophysiological, affective (interoceptive) and evocative (proprioceptive) concomitants of music perception. Music offers a powerful tool to investigate predictive coding in the brain, because the statistical regularities in music are so well defined. We have seen that the PC framework can account for several key phenomena in auditory processing. First, the classical MMN (elicited by acoustical deviants) is unaffected by explicit knowledge about upcoming deviants – due to high precision of ascending prediction errors. Thus, the MMN appears to be impervious to second-order predictions. In contrast, the ERAN (elicited by music-syntactic deviants) latency decreases, if individuals know about an upcoming syntactic deviant. Furthermore, the ERAN amplitude can be modulated by selective manipulations of confidence (i.e., precision). In short, the ERAN latency is sensitive to attentional set or knowledge about impending outcomes and can be modulated by second order predictions about the precision that should be ascribed to first-order predictions. Whether this sort of dissociation (as reflected in the MMN and ERAN) is conserved in other forms of auditory processing remains to be seen – or heard (see **Outstanding Questions Box**).

Glossary

Active inference: an enactive generalisation of predictive processing that casts both action and perception as minimising surprise or prediction error (active inference is considered a corollary of the free energy principle)

Bayesian surprise: A measure of salience based on the Kullback-Leibler divergence between a recognition density (which encodes posterior beliefs) and a prior density. It measures the information gain afforded by some data.

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

Epistemic value: Also known as intrinsic value [26,75,76] or the value of information [64,77]. The epistemic value or affordance of a covert or overt action corresponds to the salience or expected information gain. Here, epistemic value attaches to precise beliefs – as opposed to the philosophical use of epistemic, which is the kind of value that attaches to true beliefs and entails some form of propositional knowledge [78].

Expectation: The mean or average (i.e., first order moment) of a probability distribution or density over a random variable.

Musical syntax: Refers to the principles underlying the sequential organization of sounds into musical sequences. Models of cognition employ music-syntactic probabilities, e.g. *n*-Gram models, and hidden Markov models [22,27,79,80]. In several musical traditions, musical syntax also allows for nesting of phrases (centre-embedding). Thus, one may argue that human musical capacities exceed Markovian and finite-state representation [81].

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

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.

Salience: The expected reduction of uncertainty as measured by the relative entropy between posterior and prior beliefs (i.e., after and before an observation) it is also known as expected Bayesian surprise or information gain [82,83]. The salience of active sampling corresponds to its epistemic value or affordance. Salience is an attribute of probability distributions; in other words, it is 'about something' (e.g., melody, harmony, rhythm, meter, loudness, timbre, *etc*).

Surprise: Surprise, surprisal, self-information or information content is the negative logarithm of an observation or outcome. This self-information corresponds to the information content and is measured in bits or nats (depending upon the base of the logarithm) [27,84,85]. The average surprise corresponds to uncertainty; namely, entropy.

BOX 1. Predictive coding of rhythm, meter and groove

The inverted U-shaped relationship between the pleasurable drive to move and prediction-error provoking syncopations in musical groove pose a particularly interesting probe of PC [86]. The notion of precision-weighted prediction error may help us understand why we prefer to move to rhythms with intermediate levels of syncopations compared to high and low levels of syncopations. **Figure I(a)** shows how the inverted U-shape can be explained as a function of the level of syncopation and the precision assigned to the ensuing prediction errors.

Humans tend to move in synchrony with the main pulse of the music embedded within the so-called *meter*. A meter consists of differentially accented groupings and subdivisions of that pulse, and thus can be represented mentally as a predictive model [13,14]. **Figure I(b)** shows how the main pulse of the 4/4 meter is divided into strong (S̲), weak (W), less strong (s), and very weak (w) beats. Syncopations are placements of rhythmic stresses or accents falling between the beats of the musical pulse. They can be calculated directly from the score [18]. **Figure I(b)** shows the syncopations marked with red circles in the Pharrell Williams song “Happy”. Regularly organized rhythms with lower levels of syncopation feed forward only little prediction error. For the highest levels of syncopation, the meter becomes obscured, which subverts the precision of predictions. In contrast, what the system experiences as precision-weighted prediction error is highest at intermediate levels of syncopation for which both prediction error and the precision of the prediction are moderate (left of **Figure 1(a)**). According to active inference, the brain can minimize prediction error by revising predictions or through action; e.g. by moving the body in a way that changes the proprioceptive and sensory input to make it more like the predicted input [87]. In the context of musical groove, we may feel the drive to move our bodies to a regular metrical beat – at least at a subpersonal level – to suppress or attenuate the precision of prediction errors arising from syncopations.

Hence, as the level of syncopation in a groovy rhythm increases, the metrical precision decreases, marked by a decrease in sensorimotor synchronization in response to an increase in syncopation – as assessed in tapping studies and motion capture studies of musical groove [88,89]. This suggests that the internal metrical model does not fit the sensory input for the highest levels of syncopation. Thus, in addition to large prediction errors, the brain's predictive model – by which it explains away prediction error – is compromised for high levels of syncopation, because it no longer considers the sensory evidence to be sufficiently precise. In contrast, for intermediate levels of syncopation, we may experience a strong drive to reinforce the meter by moving in time with the beat. In this mode, we can elect to ignore irregularities by attenuating or suppressing their sensory precision. This account rests upon the formulation of sensory attenuation that accompanies the consequences of action. In other words, it is necessary to suspend attention – to the consequences of action – by attenuating sensory precision to realize proprioceptive predictions of the sort involved in dancing [39].

BOX 2: Does the precision hierarchy lose domain specificity at higher levels?

An interesting implication of the above formulation is the existence of a hierarchical generative model of precision. This suggests that precision control is mediated within a hierarchy that loses domain specificity at higher levels. In other words, if we are equipped with hierarchical models that generate descending predictions of precision to specific modalities, at some level in the hierarchy, predictions and predicted precision must transcend any single modality. For example, if a context gives rise to (first-order) auditory predictions and (second-order) predictions about both exteroceptive (e.g., auditory) and interoceptive precision, then there will be a necessary conflation of changes in auditory and interoceptive precision (irrespective of whether these changes are evoked by auditory or interoceptive cues). This is interesting because much of recent theorizing about emotion and selfhood rests upon interoceptive inference and, in particular, the precision of interoceptive prediction errors [90,91].

This issue is also relevant for the peculiar ability of music to induce a ‘wanting to move’ and evoke emotions. The ‘mental actions’ implicit in deploying precision to attend to salient, uncertainty resolving musical narratives may also entail (attenuated) interoceptive and proprioceptive predictions of the sort we would encounter when generating music ourselves. On this view, music perception becomes an inherently enactive process, more akin to language processing or subliminal dancing than a passive appreciation of our auditory sensorium.

Acknowledgement: The authors thank Geraint A. Wiggins for valuable comments on an earlier version of this manuscript.

REFERENCES

1. Pickering, M. J. & Clark, A. Getting ahead: forward models and their place in cognitive architecture. *Trends in Cognitive Sciences* **18**, 451–456 (2014).
2. FitzGerald, T. H., Dolan, R. J. & Friston, K. J. Model averaging, optimal inference, and habit formation. *Frontiers in Human Neuroscience* **8** (2014).
3. Bendixen, A. Predictability effects in auditory scene analysis: a review. *Frontiers in Neuroscience* **8** (2014).
4. Schröger, E., Marzecová, A. & SanMiguel, I. Attention and prediction in human audition: A lesson from cognitive psychophysiology. *European Journal of Neuroscience* **41**, 641–664 (2015).
5. Lewis, A. G. & Bastiaansen, M. A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex* **68**, 155–168 (2015).
6. Lupyan, G. & Clark, A. Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science* **24**, 279–284 (2015).
7. Friston, K. J., Stephan, K. E., Montague, R. & Dolan, R. J. Computational psychiatry: the brain as a phantastic organ. *The Lancet Psychiatry* **1**, 148–158 (2014).
8. Rohrmeier, M. A. & Koelsch, S. Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology* **83**, 164–175 (2012).
9. Koelsch, S. Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience* **15**, 170–180 (2014).

10. Forth, J., Agres, K., Purver, M. & Wiggins, G. A. Entraining IDyOT: Timing in the information dynamics of thinking. *Frontiers in Psychology* **7**, 1575 (2016).
11. Palmer, C. & Krumhansl, C. L. Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance* **16**, 728–741 (1990).
12. Honing, H. Structure and interpretation of rhythm in music. Deutsch, D., editor, *Psychology of Music*, 369–404. Amsterdam: Academic Press (2013).
13. London, J. *Hearing in time: Psychological aspects of musical meter*. Oxford University Press (2012).
14. Lerdahl, F. The sounds of poetry viewed as music. Zatorre, R. J. & Peretz, I., editors, *The Biological Foundations of Music*, vol. 930. New York: The New York Academy of Sciences (2001).
15. Large, E. W. & Kolen, J. F. Resonance and the perception of musical meter. *Connection Science* **6**, 177–208 (1994).
16. Brochard, R., Abecasis, D., Potter, D., Ragot, R. & Drake, C. The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science* **14**, 362–366 (2003).
17. Vuust, P., Roepstorff, A., Wallentin, M., Mouridsen, K. & Østergaard, L. It don’t mean a thing...: Keeping the rhythm during polyrhythmic tension activates language areas (BA47). *Neuroimage* **31**, 832–841 (2006).
18. Witek, M. A., Clarke, E. F., Kringelbach, M. L. & Vuust, P. Effects of polyphonic context, instrumentation, and metrical location on syncopation in music. *Music Perception: An Interdisciplinary Journal* **32**, 201–217 (2014).

19. Margulis, E. H. & Beatty, A. P. Musical style, psychoaesthetics, and prospects for entropy as an analytic tool. *Computer Music Journal* **32**, 64–78 (2008).

20. Wundt, W. *Grundzüge der physiologischen Psychologie [Foundations of Physiological Psychology]*. Leipzig: Engelmann (1911).

21. 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).

22. Pearce, M. T. & Wiggins, G. A. Expectation in melody: The influence of context and learning. *Music Perception: An Interdisciplinary Journal* **23**, 377–405 (2006).

23. Ognibene, D. & Baldassare, G. Ecological active vision: four bioinspired principles to integrate bottom–up and adaptive top–down attention tested with a simple camera-arm robot. *IEEE Transactions on Autonomous Mental Development* **7**, 3–25 (2015).

24. Wurtz, R. H., McAlonan, K., Cavanaugh, J. & Berman, R. A. Thalamic pathways for active vision. *Trends in cognitive sciences* **15**, 177–184 (2011).

25. Davison, A. J. & Murray, D. W. Simultaneous localization and map-building using active vision. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 865–880 (2002).

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

27. Pearce, M. T. & Wiggins, G. A. Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science* **4**, 625–652 (2012).

28. Ritter, W., Sussman, E., Deacon, D., Cowan, N. & Vaughan, H. G. Two cognitive systems simultaneously prepared for opposite events. *Psychophysiology* **36**, 835–838 (1999).

29. Sussman, E., Winkler, I. & Schröger, E. Top-down control over involuntary attention switching in the auditory modality. *Psychonomic Bulletin & Review* **10**, 630–637 (2003).

30. Rinne, T., Antila, S. & Winkler, I. Mismatch negativity is unaffected by top-down predictive information. *NeuroReport* **12**, 2209–2213 (2001).

31. Winkler, I. & Czigler, I. Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology* **83**, 132–143 (2012).

32. Garrido, M. I. *et al.* The functional anatomy of the MMN: a DCM study of the roving paradigm. *Neuroimage* **42**, 936–944 (2008).

33. Garrido, M. I., Kilner, J. M., Kiebel, S. J. & Friston, K. J. Evoked brain responses are generated by feedback loops. *Proceedings of the National Academy of Sciences* **104**, 20961–20966 (2007).

34. Garrido, M. I., Kilner, J. M., Stephan, K. E. & Friston, K. J. The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology* **120**, 453–463 (2009).

35. Donchin, E. & Coles, M. G. H. Is the P300 component a manifestation of context updating? *Behavioral & Brain Sciences* **11**, 357–374 (1988).

36. Frens, M. A. & Donchin, O. Forward models and state estimation in compensatory eye movements. *Frontiers in Cellular Neuroscience* **3**, 13 (2009).

37. Feldman, H. & Friston, K. J. Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience* **4** (2010).

38. Tavano, A., Widmann, A., Bendixen, A., Trujillo-Barreto, N. & Schröger, E. Temporal regularity facilitates higher-order sensory predictions in fast auditory sequences. *European Journal of Neuroscience* **39**, 308–318 (2014).

39. Brown, H., Adams, R. A., Parees, I., Edwards, M. & Friston, K. Active inference, sensory attenuation and illusions. *Cognitive Processing* **14**, 411–427 (2013).

40. Zeller, D., Litvak, V., Friston, K. J. & Classen, J. Sensory processing and the rubber hand illusion – an evoked potentials study. *Journal of Cognitive Neuroscience* **27**, 573–582 (2015).

41. Rohrmeier, M. & Rebuschat, P. Implicit learning and acquisition of music. *Topics in Cognitive Science* **4**, 525–553 (2012).

42. Huron, D. B. *Sweet anticipation: Music and the psychology of expectation*. The MIT Press (2006).

43. Gebauer, L., Kringelbach, M. L. & Vuust, P. Predictive coding links perception, action and learning to emotions in music. *Physics of Life Reviews* (2015).

44. Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A. & McIntosh, A. R. Predictions and the brain: how musical sounds become rewarding. *Trends in Cognitive Sciences* **19**, 86–91 (2015).

45. Vuvan, D. T., Zendel, B. R. & Peretz, I. Random feedback makes listeners tone-deaf. *Scientific Reports* **8**, 7283 (2018).

46. Fiveash, A., Thompson, W. F., Badcock, N. A. & McArthur, G. Syntactic processing in music and language: Effects of interrupting auditory streams with alternating timbres. *International Journal of Psychophysiology* **129**, 31–40 (2018).

47. Lagrois, M.-É., Peretz, I. & Zendel, B. R. Neurophysiological and behavioral differences between older and younger adults when processing violations of tonal structure in music. *Frontiers in Neuroscience* **12**, 54 (2018).

48. Sun, L., Liu, F., Zhou, L. & Jiang, C. Musical training modulates the early but not the late stage of rhythmic syntactic processing. *Psychophysiology* **55**, e12983 (2018).

49. Zhang, J., Zhou, X., Chang, R. & Yang, Y. Effects of global and local contexts on chord processing: An erp study. *Neuropsychologia* **109**, 149–154 (2018).

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

51. Zendel, B. R., Lagrois, M.-É., Robitaille, N. & Peretz, I. Attending to pitch information inhibits processing of pitch information: the curious case of amusia. *Journal of Neuroscience* **35**, 3815–3824 (2015).

52. Kim, C. H. *et al.* Melody effects on ERANm elicited by harmonic irregularity in musical syntax. *Brain research* **1560**, 36–45 (2014).

53. Brattico, E., Tupala, T., Glerean, E. & Tervaniemi, M. Modulated neural processing of Western harmony in folk musicians. *Psychophysiology* **50**, 653–663 (2013).

54. Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A. & Bhattacharya, J. Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage* **50**, 302–313 (2010).

55. Guo, S. & Koelsch, S. The effects of supervised learning on event-related potential correlates of music-syntactic processing. *Brain Research* **1626**, 232–246 (2015).

56. Guo, S. & Koelsch, S. Effects of veridical expectations on syntax processing in music: Event-related potential evidence. *Scientific Reports* **6** (2016).

57. FitzGerald, T. H., Moran, R. J., Friston, K. J. & Dolan, R. J. Precision and neuronal dynamics in the human posterior parietal cortex during evidence accumulation. *Neuroimage* **107**, 219–228 (2015).
58. Vossel, S., Mathys, C., Stephan, K. E. & Friston, K. J. Cortical coupling reflects Bayesian belief updating in the deployment of spatial attention. *Journal of Neuroscience* **35**, 11532–11542 (2015).
59. Friston, K. Hierarchical models in the brain. *PLoS Computational Biology* **4**, e1000211 (2008).
60. Auksztulewicz, R. & Friston, K. Attentional enhancement of auditory mismatch responses: a DCM/MEG study. *Cerebral Cortex* **25**, 4273–4283 (2015).
61. Lehne, M. & Koelsch, S. Toward a general psychological model of tension and suspense. *Frontiers in Psychology* **6** (2015).
62. Gregory, R. L. Perceptions as hypotheses. *Phil. Trans. R. Soc. Lond. B* **290**, 181–197 (1980).
63. Schmidhuber, J. Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. *Connection Science* **18**, 173–187 (2006).
64. Still, S. & Precup, D. An information-theoretic approach to curiosity-driven reinforcement learning. *Theory in Biosciences* **131**, 139–148 (2012).
65. Mirza, M. B., Adams, R. A., Mathys, C. D. & Friston, K. J. Scene construction, visual foraging, and active inference. *Frontiers in Computational Neuroscience* **10**, 56 (2016).
66. Metzinger, T. K. The myth of cognitive agency: subpersonal thinking as a cyclically recurring loss of mental autonomy. *Frontiers in Psychology* **4**, 931 (2013).

67. Rizzolatti, G., Riggio, L., Dascola, I. & Umiltá, C. Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia* **25**, 31–40 (1987).
68. Pearson, J. M., Watson, K. K. & Platt, M. L. Decision making: the neuroethological turn. *Neuron* **82**, 950–965 (2014).
69. Kanai, R., Komura, Y., Shipp, S. & Friston, K. Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philosophical Transactions of the Royal Society B* **370**, 20140169 (2015).
70. Shipp, S., Adams, R. A. & Friston, K. J. Reflections on agranular architecture: predictive coding in the motor cortex. *Trends in neurosciences* **36**, 706–716 (2013).
71. Chanes, L. & Barrett, L. F. Redefining the role of limbic areas in cortical processing. *Trends in Cognitive Sciences* **20**, 96–106 (2016).
72. Farb, N. *et al.* Interoception, contemplative practice, and health. *Frontiers in Psychology* **6**, 763 (2015).
73. Hohwy, J. *The predictive mind*. Oxford University Press (2013).
74. Seth, A. K. Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences* **17**, 565–573 (2013).
75. Oudeyer, P.-Y. & Kaplan, F. What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics* **1**, 6 (2009).
76. Schmidhuber, J. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development* **2**, 230–247 (2010).

77. Nelson, J. D., McKenzie, C. R., Cottrell, G. W. & Sejnowski, T. J. Experience matters: Information acquisition optimizes probability gain. *Psychological Science* **21**, 960–969 (2010).

78. McBurney, P. & Parsons, S. Representing epistemic uncertainty by means of dialectical argumentation. *Annals of Mathematics and Artificial Intelligence* **32**, 125–169 (2001).

79. Rohrmeier, M. & Graepel, T. Comparing feature-based models of harmony. *Proceedings of the 9th International Symposium on Computer Music Modelling and Retrieval* 357–370 (2012).

80. Raphael, C. & Stoddard, J. Functional harmonic analysis using probabilistic models. *Computer Music Journal* **28**, 45–52 (2004).

81. Rohrmeier, M., Zuidema, W., Wiggins, G. A. & Scharff, C. Principles of structure building in music, language and animal song. *Philosophical Transactions of the Royal Society B* **370**, 20140097 (2015).

82. Itti, L. & Baldi, P. Bayesian surprise attracts human attention. *Vision Research* **49**, 1295–1306 (2009).

83. Barto, A., Mirolli, M. & Baldassarre, G. Novelty or surprise? *Frontiers in Psychology* **4**, 907 (2013).

84. Jaynes, E. T. Information theory and statistical mechanics. *Physical Review* **106**, 620 (1957).

85. Jones, D. S. *Elementary information theory*. Oxford: Clarendon Press (1979).

86. Vuust, P., Witek, M. & Kringelbach, M. Now you hear it: A predictive coding model for understanding rhythmic incongruity. *Annals of New York Academy of Sciences* **1423**, 19–29 (2018).

87. Friston, K. Learning and inference in the brain. *Neural Networks* **16**, 1325–1352 (2003).

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

89. Witek, M. A. *et al.* Syncopation affects free body-movement in musical groove. *Experimental Brain Research* **235**, 995–1005 (2017).

90. Ainley, V., Tajadura-Jiménez, A., Fotopoulou, A. & Tsakiris, M. Looking into myself: Changes in interoceptive sensitivity during mirror self-observation. *Psychophysiology* **49**, 1672–1676 (2012).

91. Seth, A. K. & Friston, K. J. Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B* **371**, 20160007 (2016).

FIGURE LEGENDS

Figure 1: Hierarchical predictive coding in the auditory system. These schematics describe the hierarchical message passing implicit in predictive coding based on deep generative models. In this scheme, sensory input is conveyed to sensory (e.g., auditory) cortex via ascending prediction errors (e.g., from the medial geniculate). Posterior expectations, encoded by the activity of deep pyramidal cells, are driven by ascending prediction errors (red arrows). These cells then provide descending predictions (black arrows) that inform prediction errors at the lower level. At the same time, they are subject to lateral interactions that mediate (empirical) priors. Crucially, prediction errors are modulated by *predictions of their precision* (blue lines with filled blue circles). This means we have two sets of ascending and descending counter streams: the first dealing with predictions of (first-order) content and the second dealing with (second-order) context; namely, the precision of first-order prediction errors. Heuristically, expectations about precision release posterior expectations from constraints in the vicinity of an inferred attribute or trajectory – and allow them to respond more sensitively to ascending input (illustrated by the thick red arrow in the left panel). The key point here is that prediction errors compete for influence over pyramidal cells representing stimulus features (i.e., expectations). If a representation (here, the black triangle in the middle) is released from top-down constraints, it is disinhibited and becomes more sensitive to ascending prediction error. For a more detailed description of the implicit belief updating and accompanying neuronal dynamics, see [69].

Figure 2. Influence of predictive processes on music-syntactic processing. (a) Beginning of Mozart's 31st Symphony (in *D* major), and a passage from the development section (bar 131). The final note of the first phrase (the highest note in the upper panel) is the final note of a

scale (a *D* major scale); therefore, give the preceding scale context, listeners familiar with tonal music have a strong expectation for this particular note (*1st-order prediction*), and they can be relatively certain that this prediction is correct (*2nd-order prediction*). The final note of the lower panel violates these predictions: Given the context of a major scale (here *A* major) the final note is an out-of-key note (*b flat*). (b) When participants listen to musical sequences similar to those shown in (a), irregular, thus unexpected, harmonies (as composed by a composer) elicit an early right anterior negativity (ERAN), even though participants are presented repeatedly with the same sequences, and are told whether the sequence will be regular or irregular (red line: electric brain responses to irregular harmonies; blue line: brain responses to regular harmonies; the black line indicates the difference wave: regular subtracted from irregular harmonies). The isopotential maps in the lower panel of (b) show the frontal scalp distribution of the ERAN, and that the ERAN was followed a P3a with frontal preponderance and a P3b / late positive component (LPC) with parietal preponderance (isopotential maps indicate difference potentials: regular subtracted from irregular harmonies). (c) Across repeated presentations, both P3a and P3b amplitudes decline (reflecting that participants learned to predict the irregular chords). This is in contrast to the ERAN amplitude which did not change systematically. Data shown in (b) and (c) are pooled data from 20 non-musicians and 20 amateur musicians, figure modified from REF. [55].

Figure 3. Predictive processes influence latency of early music-syntactic processing. Musicians were presented repeatedly with two different chord sequences (upper right panel), ending on a regular harmony (a *tonic* chord, upper sequence) or an irregular harmony (a *double dominant*, lower sequence). In different blocks, participants were either informed or not informed about the sequence ending (i.e., about whether the final chord would be regular or

irregular). Independent of whether participants were not informed (a) or informed (b) about the upcoming chords, the irregular sequence endings elicited an ERAN that did not significantly differ in amplitude between conditions (not informed, informed). P1, N1, and P2 ERPs elicited by regular chords were not affected by the fact that participants were or were not informed. The scalp distribution of the ERAN is shown in the upper scalp maps in panels (a) and (b). By contrast, irregular chords elicited a P3a only when participants were not informed and a P3b only when participants were informed about the sequence endings. The scalp distribution of P3a and P3b is shown in the lower scalp maps in panels (a), and (b) respectively. Note that when participants were told about the upcoming ending, the ERAN was maximal at around between 160-200 ms, i.e. around 10 ms shorter than when participants were not informed (170-210 ms). A similar latency difference (in the absence of amplitude change) was also observed in an independent group of non-musicians (see bottom right panel). Thus, top-down predictions did not modify the ERAN amplitude, but they accelerated bottom-up processing of expectancy violations, or error signals. Data were obtained from 20 musicians and 20 non-musicians. Figure modified from REF. [56].

Figure 4. Decreasing influence of top-down predictions on ascending prediction errors. The image illustrates the precision-weighted decrease of top-down (descending) predictions towards lower levels of the processing hierarchy. Empirically, this is reflected, e.g., in the observation that music-syntactic irregularities strongly influence the P3b, less so the P3a, and only the latency (but not the amplitude) of the ERAN.

Box Figure I. Proposed model of predictive coding of syncopation.