Now you hear it: A predictive coding model for understanding rhythmic incongruity

Vuust, Peter; Witek, Maria; Dietz, Martin J.; Kringelbach, Morten L

DOI:
10.1111/nyas.13622

Now You Hear It: A predictive coding model for understanding rhythmic incongruity

Peter Vuust[a]
Martin Dietz[b]
Maria Witek[a]
Morten L. Kringelbach[a,c]

a) Center for Music in the Brain, Dept. of Clinical Medicine, Aarhus University & The Royal Academy of Music Aarhus/Aalborg, Denmark
b) Center for Functionally Integrative Neuroscience, Dept. of Clinical Medicine, Aarhus University, Denmark
c) Department of Psychiatry, University of Oxford, Oxford OX3 7JX, United Kingdom

Corresponding author:

Peter Vuust
Professor, MSc., PhD
Director, Center for Music in the Brain (MIB)
Prof. Neuroscience, Dept. of Clinical Medicine, Aarhus University
Prof. of Music, Royal Academy of Music, Aarhus/Aalborg, Denmark

Nørrebrogade 44, 8000 Aarhus C, Denmark
petervuust@gmail.com
ph: (+45) 78 46 16 17, (+45) 72 26 74 66, mobile: (+45) 27 11 94 71

Short title:
Now You Hear It

Key Words: Rhythm; Music; Predictive Coding; Brain;

Abstract
Rhythmic incongruity in the form of syncopation is a prominent feature of many contemporary musical styles. Syncopations afford incongruity between rhythmic patterns and the meter – giving rise to mental models of differently accented isochronous beats. Syncopations occur either in isolation or as part of rhythmic patterns, so-called *grooves*. Based on the predictive coding framework, we here discuss how brain processing of rhythm can be seen as a special case of predictive coding. We present a simple, yet powerful model for how the brain processes rhythmic incongruity: *the model for predictive coding of rhythmic incongruity (PCRI)*. Our model proposes that a given rhythm's *syncopation* and its metrical uncertainty (*precision*) is at the heart of how the brain models rhythm and meter based on priors, predictions and prediction error. Our minimal model can explain prominent features of brain processing of syncopation: why isolated syncopations lead to stronger prediction error in the brains of musicians as evidenced by larger ERPs to rhythmic incongruity, and why we all experience a stronger urge to move to grooves with a medium level of syncopation compared to low and high levels of syncopation.

A brief introduction to predictive coding

Prediction is increasingly viewed as a fundamental principle of brain processing that determines perception, action, and learning. Emerging predictive coding theories[1-6] have offered novel explanations for how specialized brain networks can identify and categorize causes of its sensory inputs, integrate information with other networks, and adapt to new stimuli. Here, for simplicity, we will use the term predictive coding (PC) as synonymous with Karl Friston's hierarchical predictive coding framework [5]. Briefly, PC proposes that perception, action and learning is a recursive Bayesian process by which the brain attempts to minimize the prediction error between lower-level sensory input and the brain's top-down predictions. An excellent summary of the recent advances was given by Andy Clark[5].

Under a Bayesian formulation of predictive coding in the brain, perception corresponds to inverting a generative model of the things in the world that cause our sensations. These causes are hidden in the sense that things in the world can only be observed through noisy sensory input that evolves over time. Computationally, this model inversion could be achieved in continuous time by minimizing a free-energy bound on the surprise $\mathcal{F} > -\ln p(\tilde{s}|m)$ about sensory input $\tilde{s}$ given the brain's model m of the world. The free energy $\mathcal{F}$ is a function of sensory input $\tilde{s}$ and a probability density $q(\vartheta)$ that parameterises its hidden causes and their states.

Free energy

$$\mathcal{F} = \mathrm{E}_q[\ln q(\vartheta|\mu) - \ln p(\tilde{s}, \vartheta|m)]$$

Minimizing the free energy F corresponds to maximizing the evidence $\ln p(\tilde{s}|m)$ for the brain's model of the world (Friston 2010). In predictive coding, top-down connections provide lower levels with predictions in the form of prior expectations about states of the world, whereas bottom-up connections carry prediction errors that update posterior expectations in higher levels to provide better predictions. This leads to the following hierarchical equations for how top-down predictions $g(\mu^{(i)})$ given by posterior expectations $\mu^{(i)}$ at higher levels and bottom-up prediction errors $\epsilon^{(i)} = \mu^{(i-1)} - g(\mu^{(i)})$ from lower levels evolve when exposed to changes in stimuli $\tilde{s}$

Predictions

$$\dot{\mu}^{(i)} = \frac{\partial g(\mu^{(i)})}{\partial \mu^{(i)}} \cdot \xi^{(i)} - \xi^{(i+1)}$$

Precision-weighted prediction errors

$$\xi^{(i)} = \pi^{(i)}(\mu^{(i-1)} - g(\mu^{(i)}))$$

where the dot notation $(\dot{\cdot})$ denotes the time derivative and π is the precision assigned to the prediction errors. The i index is used to refer to a relative hierarchical level. Both higher-level

predictions and lower-level prediction errors are weighted by their precision. The precision is the inverse of the variance and encodes the confidence about sensory inputs in lower areas, relative to the confidence with which states in the world that cause sensory inputs can be predicted in higher areas.

The predictive coding of rhythmic incongruity (PCRI) model

In terms of music, the ideas behind predictive coding in the brain are remarkably similar to theories addressing the role of prediction in music perception and appreciation[7-10]. Predictive processes constitute central mechanisms in the perception and cognition of music. They are prerequisites for joint music making[11-13] and are essential for musical tension and surprise[10] as well as for intra-musical meaning[9]. For this reason, several authors have proposed music as an ideal domain for testing and further developing predictive coding theories[14-16], informing our understanding of brain mechanisms in general, and perhaps even helping us to understand the fundamental prediction principles of the brain.

Recent behavioral and neuroscientific experiments have shown how brain processing of rhythm can be seen as a special case of predictive coding[17]. This prompted us to develop a simple, yet powerful model explaining how the brain processes rhythmic incongruity, the PCRI model (Figure 1). The present account will focus on two phenomenologically distinct, yet structurally related types of rhythmic incongruity: the occasional appearance of a surprising beat followed by a surprising rest (*syncopation),* and repeated syncopated patterns (*groove*). However, the model may possibly be further extended to micro-timing and maybe even to the relationship between tonal center and melody or harmony.

Figure 1 shows a schematic of our proposed model. Specifically, we propose that the brain's perception of syncopation is determined not only by the prediction errors that follow from rhythmic incongruity, but also by how these are weighted by their relative precision. This means that the expected precision encodes the confidence with which we extract the meter from a particular rhythm. By assigning *more or less* precision or confidence to the ensuing prediction errors, the brain perceives the rhythm as *more or less* syncopated, because these prediction errors are given more salience. In other words, the prediction errors that matter are those that we assign a greater precision or confidence. This means that the prediction errors that matter for perceptual synthesis have to be violations that are 'predictably unpredicted'. If there are too many violations, prediction errors will be attenuated because the 'predicted precision' is itself too low, and there is a high degree of uncertainty about the meter. In the case of *syncopation*, we can obtain an estimate of both the precision and the prediction error. The syncopation in a given musical rhythm can be calculated directly from the musical score, demonstrated e.g. by Longuet-Higgens & Lee's formulation[18] or Witek and colleagues' adjusted formulation[19]. The *precision* (metrical uncertainty) can be behaviorally estimated by measuring participants' sensorimotor synchronization to the beat using finger-tapping paradigms[20, 21], motion capture[22, 23], or by neurophysiological measures[24-26].

In the following, we will show how predictive coding in general and the PCRI model in particular can help to explain experimental observations concerning musical rhythm and meter. In particular, we shall demonstrate that PCRI accounts well for the observed U-shaped relationship between syncopation and experience of groove and for the effect of expertise on brain processing of syncopations, where the prediction errors that matter for the perception of syncopation are violations that are 'predictably unpredicted' under the brain's model of the meter.

Predictive Coding of Rhythm

Traditionally, music theory holds that rhythmic events are perceived as groupings of temporal events against the backdrop of an implied reference structure, namely the meter. The meter is a hierarchical framework consisting of evenly spaced and differentially accented beats, providing to each metric position a timing and a metrical weight. The metrical weights are thought to linearly correspond to the strength of the expectation towards events occurring at these time points[27]. In other words, the more metrically salient a position is in the hierarchy, the stronger the expectation that events will occur at this metrical position.

Under predictive coding, the rhythm is the acoustical input to our ears, whereas the meter is the brain's posterior expectations that constitute its predictive model. The rhythm can be more or less in accordance with the meter, creating stronger or weaker prediction error between auditory input and predictive model. Brochard et al. [24] provided strong evidence for the existence of metric expectations in the simplest possible experimental setting, when they showed that listening to a series of entirely regular and unaccented metronome beats causes the brain to automatically register the beats as alternating in salience (a 2/4 or a 4/4 meter). In predictive coding terms, the brain is interpreting the neutral input, in this case un-accented metronomic beats, according to its own predictive framework (*the meter*[28, 29]).

Important for our PCRI model, the way we experience the rhythmic content in music is heavily dependent on how precise our model of the meter is. In music, the rhythms will usually be much less predictable than simple metronomic beats. They will in different ways engage the brain, creating prediction error that can challenge the metrical model, even to the point where a different meter may be as good or better at explaining the auditory input [30, 31]. The latter is the case for polyrhythm, where two rhythms indicating two competing meters are played simultaneously, creating tension between the rhythmic events and the meter[30]. When listening to music, the brain is constantly trying to assess the plausibility of competing models or hypotheses (i.e. meters), given its musical input. The ensuing prediction errors are evoked by the actual music (bottom-up) on one hand and depend on the expectations of the interpreting brain (top-down) on the other. Importantly, brain processing and structure underlying musical expectation are shaped by culture, personal listening history, musical training, and biology [32-37].

The central PC claim that the brain uses Bayesian inference when choosing a plausible metrical model for a given rhythmical input was recently supported experimentally. Using a finger-tapping paradigm, Elliot and colleagues provided evidence suggesting that humans exploit a Bayesian inference process to control movement timing, when facing microtemporal differences[38]. They presented two metronomes of equal tempo, but differing in phase and temporal regularity to participants, and asked participants to synchronize their tapping with the experienced beat. When participants chose to integrate the two timing cues into a single-event estimate, modeling the behavior as a Bayesian inference process provided a better description of the data than other plausible models. This is consistent with the PC claim that the brain uses Bayesian inference when choosing a mental model for interpreting noisy sensory data. Note, though, that such a behavioral finding is not sufficient to conclude that the brain processes are also governed by Bayesian inference.

Syncopations and PCRI

Our PCRI model targets the frequently investigated example of prediction error arising from a rhythm-meter discrepancy: *syncopation*[19, 27, 39, 40]. Syncopation occurs when onsets occur on metrically weak accents and subsequent rests or tied notes occur on metrically strong accents. Such expectations can be conceptualized in Bayesian terms[41, 42]: By assigning relative probabilities to all notes and rests of a pattern, based on prior information about statistical frequencies and a hierarchical model of meter, a syncopation's perceptual effect is a consequence of its predictability within the context of music as a whole. Importantly, for a syncopation to obtain its characteristic effect, it must be experienced as contradicting the meter, yet not so strongly that the experience of the meter falls apart. Syncopations can also be thought of as phase-shifts, where the rhythmic onset, rather than occurring in phase with its metric reference point, has a negative lag and occurs before it. Hence, syncopations will influence the two terms on the right side of the upper equation in Figure 1. On one hand, they create a prediction error between the sensory input and the prediction. On the other hand, they may unsettle the precision of our meter perception and thus the precision-weighted prediction error.

Using the PCRI model to understand isolated syncopations in musicians

Vuust and colleagues were the first to note that neural responses to isolated syncopations occurring in continuous rhythmic streams are consistent with the predictive coding framework in that they have properties similar to electrophysiological error signals and their subsequent evaluation [43]. They performed magneto-encephalography (MEG) while musicians and non-musicians were listening without attending to isolated syncopations occurring pseudo-randomly in musical drum rhythm excerpts. These syncopations elicited two prominent ERPs, the magnetic counterpart of the mismatch negativity (MMNm) and the P3am. The mismatch negativity appears to have the properties of an error signal arising from superficial cortical layers as posited by PC. It is elicited to violations of auditory expectancy and has been found in response to pattern deviations determined by physical parameters, such as frequency[44], intensity[45], spatial localization[46] and duration[47], but also to patterns with more abstract properties[48, 49].

Musical experts are known to have larger amplitude and latency of the MMN[50, 51]. Accordingly, in the study by Vuust and colleagues, rhythmic expert musicians were observed to have larger MMN amplitudes compared to rhythmically unskilled participants. We know from a large corpus of tapping literature (for reviews, see [20, 21]) and neurophysiological data [24] that musically trained individuals have more precise meter perception than non-musicians. Because of that, the larger error response observed in the brains of rhythmically skilled musicians is consistent with the PCRI model. Here, the precision-weighted prediction error is the difference between the prediction and the sensory input, multiplied by the precision of the prediction (Figure 1). Hence, even though the discrepancy between the rhythm and the meter as measured in the stimulus was the same for both musician and non-musician participants, the experienced prediction error is, according to the model, weighted differently. This is consistent with the larger ERPs to the rhythmic incongruity in the musicians.

The neural processing of these isolated syncopations seems to involve the attentional network. The MMNm, originating primarily in the auditory cortices, was followed by a P3am that was localized to a larger network tying together components from auditory cortex with parietal and frontal brain,

consistent with the typical localization of the P3a to frontal[52, 53], auditory[54, 55] and temporo-parietal[56, 57] sources (for a review see[58]). Research on P3a demonstrates that it represents a network with both task-specific and general elements[58]. One likely explanation is that the P3a reflects a network involving both the modality/task specific areas evoking the error signal and higher regions that can evaluate it[46]. This is exactly what would be expected for error signals in response to a rhythmic incongruity in a predictive coding framework, suggesting that the P3am reflects a neural network that acts on the error signal of the MMN. The MMN and the P3a are generally believed to reflect different stages of processing subserving an attention switching mechanism[59, 60]. Whereas the MMN is thought to be the first stage in involuntary attention capture[61], the P3a most likely reflects the actual switch of attention[62]. The P3a response may indicate that attention should be designated to the metric violation as a means of providing a better estimate. In terms of predictive coding, there is an intimate relationship between attention and precision. Prediction errors that are afforded greater precision are effectively boosted, such that they have a greater influence on higher level expectations and consequent predictions. Crucially, the brain has to predict both the content of the sensorium and its precision. Simulations of predictive coding using e.g. the Posner paradigm suggest that late (endogenous) responses, such as the P300, may reflect a revision of beliefs about the precision or predictability of sensory streams[63]. This suggests that early (i.e. mismatch negativity) violation responses correspond to a precision-weighted prediction error, while later (i.e., P300) responses reflect belief updates about precision per se – that underwrite a redeployment of attentional gain [46].

Music in general encompasses such incongruities that direct the listener's attention towards salient parts of the music. Vuust et al.[43] found larger MMNm and P3am in experts suggesting that both the competence of the listener (top-down) and strength of the musical violation (bottom-up) determine whether attention is attracted to the stimulus. Participants tested in this study were all improvising musicians who need to be able to respond swiftly to such incongruities[30], and it might be that they have developed a more precise metrical model. Interestingly, the debriefing of the participants indicated that they did not consciously distinguish the different types of metric displacement nor displayed any aesthetic appreciation of the stimuli, despite clear brain processing differences between groups.

Groove

While syncopations occurring in isolation seem predominantly to engage attention switching brain mechanisms, syncopations that occur regularly within the rhythmic texture of music may have a quite different purpose and effect on the nervous system – one which makes us want to move and which feels pleasurable. With many contemporary styles of popular music, especially music with African-American influences, the sensation of groove is an important affective response. Groove is characterized by a pleasurable drive towards body movement in response to rhythmically entraining elements in the music[64-68]. In groove-directed music, such as jazz, soul, funk, hip-hop, electronic dance music and reggae, the tempo of the music is mostly kept constant (Figure 3). On one hand, this eliminates the possibility of expressing emotions through tempo alterations, as is more common in classical music[69, 70]. On the other hand, a stable tempo makes it easier to create tension between the rhythm and the meter. In most groove-directed music, listeners expect the tempo to remain largely unchanged within a certain piece. They may therefore evaluate any rhythmic incongruity as a potential syncopation. This opens up a fine-grained grid of possible rhythmic layers for the musicians to play with. In these styles of music, the rhythm section usually consisting of drums/percussion, bass and guitar/keyboard will often play repeated syncopated rhythmic patterns

(*grooves*), keeping the amount of syncopation relatively constant in the different sections of the pieces[71].

The pleasurable sensation of wanting to move to highly repetitive syncopated rhythms was investigated by Witek and colleagues in a series of studies[22, 68, 72]. Using a battery of 50 groove-based drum patterns, they asked participants in an online survey to rate the patterns on a 7-point Likert scale as to how much they wanted to move and the pleasure felt. The degree of syncopation in the stimuli was calculated using Witek et al's index of syncopation, which adds instrumental weights to the model proposed by Longuet-Higgens & Lee[73] to adjust for the polyphonic character of drum patterns. Briefly, a pattern's overall degree of syncopation is calculated by taking the sum of differences in metrical weights between the notes and rest that constitute the syncopations, adjusting for instrumental weights. The 50 drum patterns included 34 transcribed from real funk tracks, whereas the remaining patterns were constructed specifically for the experiment aiming for a continuum from weakly syncopated to strongly syncopated rhythm patterns. Witek and colleagues found an inverted U-shaped relationship between degree of syncopation and the groove ratings, suggesting that the sensation of groove is strongest at intermediate levels of discrepancy between the rhythmic (sensory) input and the metrical predictive framework. The inverted U-shape[74] has earlier been hypothesized to reflect the relationship between music complexity and liking[75-77] and perceptual complexity and arousal in art more broadly[78], although empirical studies have shown that this function largely depends on the musical style in question[77]. In a subsequent study Witek and colleagues[79] used motion-capture to record free movements in hand and torso while participants listened to a subset of 15 of the drum patterns mentioned above, categorized into three levels of syncopation; low, medium and high. For low and medium levels of syncopation, participants synchronized their movements to the meter, whereas for high levels of syncopation they synchronized very poorly.

How to understand the inverted U-shape of groove in terms of the PCRI model

The notion of an inverted U-shaped relationship between syncopation and the pleasurable drive to move is congruent with the notion of precision-weighted prediction error as formulated in the PCRI model, shown in the bottom panel of Figure 1. The regularly organized rhythms with lower levels of syncopation feed forward only little prediction error. For the highest levels of syncopation the meter becomes obscured, leading to less precision in the predictive model. Here, it is difficult for the brain to detect the signal in the noise. In contrast, what the system experiences as precision-weighted prediction error is highest at intermediate levels of syncopation for which both objective prediction error and the precision of the prediction are moderate (Figure 1). According to PC, the brain can minimize prediction error through action. By moving the body in a way that changes the bottom-up proprioceptive and sensory input and thus resampling the evidence[80], the error signal will self-suppress. In the context of groove, we feel the urge to move our bodies to the metrical beat in order to – at least at an unconscious level – strengthen the metric model and suppress or attenuate the precision of prediction errors.

Importantly, this reasoning is dependent on a linearly decreasing relationship between meter perception and syncopation, as schematically shown in the middle panel of Figure 1. This relation is partly supported by the decrease in synchronization in response to increase in syncopation found in Witek et al's motion capture study[22]. Here, the results suggested a broken metrical model for the

highest levels of syncopation. Hence, 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 the intermediate levels of syncopation, we may experience a strong drive towards reinforcing the meter by moving in time with the beat. We may here elect to ignore violations by attenuating or suppressing their sensory precision. This account rests upon the formulation of sensory attenuation through the attenuation of precision that accompanies the consequences of action. In other words, in active inference formulations of predictive coding, 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). Psychologically, this corresponds to ignoring the consequences of action to selectively discount evidence against our predictions of sensory input[81]. Future studies should aim at testing this hypothesis comparing e.g. dancing and non-dancing participants' perception of or memory for syncopation. Paradoxically, though, moving to the beat and hence reinforcing the meter allows for more precise predictions, which would reinforce the prediction error from subsequent syncopations.

The urge to move to music seems to be related to auditory-motor coupling as described in a number of neuroscientific studies. These studies show activity in brain networks linking auditory and sensory-motor areas of the brain to the perception of musical rhythm[82, 83]. Furthermore, electrophysiological data shows that even for rhythms in which the meter is not acoustically accented, the fundamental frequencies of the meter still dominate the signal[84]. Recently, Large and colleagues showed that participants' degree of synchronization with increasingly syncopated rhythms could successfully be explained by a neuronal network model encompassing a hierarchy of only two levels; one corresponding to the sensory system modelled with a simple Hopf bifurcation, the other corresponding to the motor system tuned to operate near a double limit cycle bifurcation[85].

Despite being consistent with the PCRI model's rhythmic and metric levels, Large's model does not explicitly incorporate the behaviorally reported pleasure aspect of groove. Prediction and expectation are frequently linked to emotion and pleasure in music scholarship[86, 87], but there is still no empirical evidence for why medium levels of prediction error in music are the most pleasurable. Kringelbach and Berridge[88] suggested that the brain rewards prediction error since it leads to learning and thereby maximizes future prediction. Another perspective on the paradoxical attractiveness of prediction errors is that they play a central role in active inference formulations of predictive processing. In this instance, prediction errors portend an opportunity to resolve uncertainty and minimize prediction errors in the future[88]. Formally, this has been cast in terms of salience or epistemic affordance. This fits comfortably with the opportunity provided by predictably unpredictable music. Rewarding actions minimize the brain's free energy or maximize epistemic value, thus building a more generalizable and accurate model of the world. In Bayesian terms, this translates into an optimization of the evidence for our models, or succinctly, self-evidencing[89].

Though it is important not to confuse reward prediction error with predictive coding, a likely candidate for mediating the effect of musical reward is the neurotransmitter dopamine in the mesolimbic pathway, as suggested by Gebauer et al[90]. Research in rodents[91, 92] has shown dopamine release to both expected and unexpected stimuli, suggesting that the complex interaction between dopamine release and predictions leads to adaptive learning in the short and long term. A still

unresolved question is whether the relationship between syncopation in groove and pleasure is modulated by the dopamine system.

Generalizability of the PCRI model

For simplicity we have restricted our PCRI model to target rhythmic syncopations. Importantly, we have argued that it explains prominent features of the brain processing of syncopation. As evidenced by musicians' larger ERPs to rhythmic incongruity, isolated syncopations seem to lead to larger precision-weighted prediction errors than in non-musicians[93]. As shown in subjective rating studies, listeners experience a stronger urge to move to grooves with medium levels of syncopation, compared to low and high levels[68]. Here, it is important to note that the rhythm/meter dichotomy, a *schematic prediction* that is culturally learned from early childhood[94], is only part of the predictive processes related to groove. While listening to a musical groove, the brain also forms *short term predictions*, through drum/bass patterns which are repeated over and over again, and is influenced by *veridical expectation*, i.e. knowledge about the time course of a specific musical piece after repeated listening. These expectations are thought to be processed by different brain networks [87]. For simplicity, the PCRI model does not at present consider these veridical predictive processes, but they could potentially be incorporated into future versions of the model.

Another example of metric incongruity in groove is microtiming[95] – the small temporal discrepancies between the meter and the rhythmic events as played or phrased by musicians and music producers. However, the contribution of microtiming to the pleasurable experience of wanting to move (groove) remains unclear. On one hand, it is clear that these systematic incongruities between the actual realization of the rhythms and the meter in well-played groove music exist[96]. On the other hand, their contribution to the sensation of groove has been difficult to determine in a laboratory context. In a recent study[97], Davies and colleagues found, contrary to common belief, that systematic increase in microtiming led to decreased groove ratings except for a prototypical jazz pattern. For this pattern the groove ratings were largely unaffected, an effect that was more pronounced in an expert listener group than for untrained listeners. The general decrease in groove ratings for larger magnitudes of microtiming is consistent with PCRI. According to PCRI, microtiming increase would supposedly lead to a decrease in metrical certainty whereas the syncopation in the rhythms used in that study remained the same. However, as Davies et al's study indicates, this effect can be overwritten by musical expertise. For expert jazz listeners, microtiming differences are stylistically expected; hence they may not lower the precision. This is consistent with our earlier discussion of isolated syncopations, where musicians' larger ERPs compared to non-musicians could be explained by the fact that musicians have more stable metrical representations than non-musicians.

Conclusions and caveats

In the present paper, we have presented a simple model for understanding how the brain processes rhythmic incongruity, namely *the model for predictive coding of rhythmic incongruity*. The model proposes that the explainable prediction error processed by the brain depends on a combination of syncopation and the uncertainty of the meter perception. While this model can effectively explain important phenomenological aspects of rhythmic incongruity, including expertise-related

differences in brain processing of *isolated syncopations* and the *inverted U-shaped relationship* between the experience of wanting to move and amount of syncopation, we still lack evidence regarding the pleasure component of the sensation of groove. We may of course speculate that our affective evaluation of a rhythm's relative 'grooviness' also depends on a combination of the actual prediction error and the uncertainty of our rhythmic prediction. In this regard, the PCRI model formulations emphasize that rewarding actions are those that minimize the brain's free energy, thus building a stronger and more accurate model of the world. But it does not consider prediction error as such to be positive or negative. Future studies should aim to clarify the relationship between rhythmic incongruity and the resolution of uncertainty (i.e., salience and epistemic affordance), and to determine the role of this relationship in making meaningful and enjoyable musical experiences.

Figure legends

Figure 1: Proposed model of the predictive coding of rhythmic incongruity (PCRI). The figure provides a schematic illustration of the variables related to increasing syncopation of musical grooves (dotted lines). A) Under predictive coding, the precision-weighted prediction error is given by the difference between the sensory stream $\tilde{s}$ and the brain's predictions $g(\mu)$ timed with the precision π. The *i* index is used to refer to a relative hierarchical level in the brain. For grooves the syncopations result in a prediction error $\epsilon = \tilde{s} - g(\mu)$, which can be calculated directly from the score by using e.g. Witek et al's modification[73] of Longuet-Higgens & Lee's formulation. B) By assigning *more or less* precision or confidence to the ensuing prediction errors, the brain perceives the grooves as *more or less* groovy. C) We propose that the observed U-shaped relationship between syncopation and grooviness[98] can be explained by the PCRI model as a function of the level of syncopation and precision or confidence assigned to the ensuing prediction errors. D) The formulas for describing the relationship.

Figure 2. Syncopation and meter. Syncopation (a) is as a mismatch between the auditory input (the rhythm) and the meter (the brain's predictive model – *s* and *w* denote strong and weak beats respectively), which creates prediction error between lower-level sensory areas and higher-level areas (b) leading to perception, action (in the form of *wanting to move*), emotion and learning.

Figure 3: Different musical styles have different relationships between rhythm and meter. The figure is a schematic illustration of the stylistic differences in the use of syncopation and tension between rhythm and meter. A) In classical music, the tempo is (often) flexible allowing for expressive timing. B) In jazz music the tempo is kept relatively constant, but the rhythm section will constantly vary the degree of tension between the rhythms and the meter using single (*) or multiple syncopations and polyrhythms (**). C) In groove-based music the tempo is ideally kept completely constant throughout a piece of music. Here the rhythm section will often play a groove in which the amount of syncopation in different sections of the piece is kept constant. Vocalists or soloists might vary the use of syncopations, however.

References

1. Friston, K. 2005. A theory of cortical responses. *Philos.Trans.R.Soc.Lond B Biol.Sci*. **360**: 815-836.
2. Bar, M. 2009. Predictions: a universal principle in the operation of the human brain. *Philos T R Soc B*. **364**: 1181-1182.
3. Rao, R.P. & D.H. Ballard. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat.Neurosci*. **2**: 79-87.
4. llinas, R.R. 2001. "Prediction is the Ultimate Function of the Brain". In *I of the Vortex*. R.R. Llinas, Ed.: 21-52. Massachusetts: The MIT Press.
5. Clark, A. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci*. **36**: 181-204.
6. Friston, K. 2010. The free-energy principle: a unified brain theory? *Nat Rev Neurosci*. **11**: 127-138.
7. Meyer, L. 1956. *Emotion and Meaning in Music*. Chicago: University of Chicago Press.
8. Huron, D. 2006. *Sweet Anticipation*. Cambridge, Massachusetts, London: The MIT Press.
9. Vuust, P. & M.L. Kringelbach. 2010. The pleasure of making meaning: evidence from the neuroscience of music. *ISR.Interdisciplinary science review*. **35**: 166-182.
10. Rohrmeier, M.A. & S. Koelsch. 2012. Predictive information processing in music cognition. A critical review. *International journal of psychophysiology*. **83**: 164-175.
11. Keller, P.E., G. Knoblich & B.H. Repp. 2007. Pianists duet better when they play with themselves: on the possible role of action simulation in synchronization. *Conscious Cogn*. **16**: 102-111.
12. Gebauer, L., M. Witek, N. Hansen, *et al.* 2016. Oxytocin improves synchronisation in leader-follower interaction. *Scientific reports*. **6**.
13. Konvalinka, I., P. Vuust, A. Roepstorff, *et al.* 2010. Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology*. **63**: 2220-30
14. Vuust, P., L. Ostergaard, K.J. Pallesen, *et al.* 2009. Predictive coding of music. *Cortex*. **45**: 80-92.
15. Friston, K. & D.A. Friston. 2013. "A Free Energy Formulation of Music Generation and Perception: Helmholtz Revisited". In *Sound - Perception - Performance*, Vol. 1. R. Bader, Ed.: 43-69. Switzerland: Springer International Publishing.
16. Schaefer, R.S., K. Overy & P. Nelson. 2013. Affect and non-uniform characteristics of predictive processing in musical behaviour. *Behav Brain Sci*. **36**: 2.
17. Vuust, P. & M.A. Witek. 2014. Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Front Psychol*. **5**: 1111.
18. Longuet-Higgins, H.C. & C.S. Lee. 1984. The rhythmic interpretation of monophonic music. *Music Perception*. **1**: 18.
19. Witek, M.A.G., E.F. Clarke, M.L. Kringelbach, *et al.* in press. Effects of polyphonic context, instrumentation and metric location on syncopation in music. *Music Perception*. **2**: 201-217
20. Repp, B.H. 2005. Sensorimotor synchronization: a review of the tapping literature. *Psychon Bull Rev*. **12**: 969-992.
21. Repp, B.H. & Y.H. Su. 2013. Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin and Review*. **20**: 403-452.

22. Witek, M.A., T. Popescu, E.F. Clarke, *et al.* 2017. Syncopation affects free body-movement in musical groove. *Experimental brain research*. **235**: 995-1005.
23. Burger, B., J. London, M.R. Thompson, *et al.* 2017. Synchronization to metrical levels in music depends on low-frequency spectral components and tempo. *Psychol Res*.E-pub ahead of print.
24. Brochard, R., D. Abecasis, D. Potter, *et al.* 2003. The "ticktock" of our internal clock: direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*. **14**: 362-366.
25. Snyder, J.S. & E.W. Large. 2005. Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Brain Res Cogn Brain Res*. **24**: 117-126.
26. Jongsma, M.L., P. Desain & H. Honing. 2004. Rhythmic context influences the auditory evoked potentials of musicians and nonmusicians. *Biol.Psychol*. **66**: 129-152.
27. Longuet-Higgins, H.C. & C. Lee. 1984. The rhythmic interpretation of monophonic music. *Music Perception*. **1**: 424-440.
28. Benjamin, W.E. 1984. A Theory of Musical Meter. *Music Perception*. **1**: 355-413.
29. Palmer, C. & C.L. Krumhansl. 1990. Mental representations for musical meter. *Journal of Experimental Psychological Human Perception and Performance*. **16**: 728-741.
30. Vuust, P. 2000. *Polyrhythm and Metre in Modern Jazz - a Study of the Miles Davis' Quintet of the 1960'ies (Danish)*. Aarhus: RAMA.
31. Pressing, J., J. Summers & J. Magill. 1996. Cognitive multiplicity in polyrhythmic pattern performance. *Journal of Experimental Psychology-Human Perception and Performance*. **22**: 1127-1148.
32. Zatorre, R.J. & I. Peretz. 2001. *The biological foundations of music*. New York: New York Academy of Sciences.
33. Bangert, M. & G. Schlaug. 2006. Specialization of the specialized in features of external human brain morphology. *Eur.J.Neurosci*. **24**: 1832-1834.
34. Chakravarty, M.M. & P. Vuust. 2009. Musical morphology. *Annals of the New York Academy of Sciences*. **1169**: 79-83.
35. Bigand, E. & B. Poulin-Charronnat. 2006. Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition*. **100**: 100-130.
36. Tillmann, B., J.J. Bharucha & E. Bigand. 2000. Implicit learning of tonality: a self-organizing approach. *Psychol Rev*. **107**: 885-913.
37. Krumhansl, C.L., P. Toivanen, T. Eerola, *et al.* 2000. Cross-cultural music cognition: cognitive methodology applied to North Sami yoiks. *Cognition*. **76**: 13-58.
38. Elliott, M.T., W. A.M. & W. A.E. 2014. Moving in time: Bayesian causal inference explains movement coordination to auditory beats. *Proceedings of the Royal Society B*. **281**.
39. Fitch, W.T. & A.J. Rosenfeld. 2007. Perception and production of syncopated rhythms. *Music Perception*. **25**: 43-58.
40. Ladinig, O., H. Honing, G. Haden, *et al.* 2009. Probing attentive and preattentive emergent meter in adult listeners without extensive musical training. *Music Perception*. **26**: 377-386.
41. Temperley, D. 2007. *Music and probability*. MIT Press. Campridge, MA.
42. Temperley, D. 2010. Modeling common-practice rhythm. *Music Perception*. **27**: 355-376.
43. Vuust, P., L. Ostergaard, K.J. Pallesen, *et al.* 2009. Predictive coding of music--brain responses to rhythmic incongruity. *Cortex*. **45**: 80-92.
44. Sams, M., P. Paavilainen, K. Alho, *et al.* 1985. Auditory frequency discrimination and event-related potentials. *Electroencephalogr. Clin. Neurophysiol.* **62**: 437-448.

45. Näätänen, R., P. Paaviliainen, K. Alho, *et al.* 1987. The mismatch negativity to intensity changes in an auditory stimulus sequence. *Electroencephalogr. Clin. Neurophysiol.* **40**: 125-131.
46. Dietz, M.J., K.J. Friston, J.B. Mattingley, *et al.* 2014. Effective connectivity reveals right-hemisphere dominance in audiospatial perception: implications for models of spatial neglect. *J Neurosci*. **34**: 5003-5011.
47. Paavilainen, P., M.-L. Karlsson, K. Reinikainen, *et al.* 1989. Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalogr. Clin. Neurophysiol.* **73**: 129-141.
48. Paavilainen, P., J. Simola, M. Jaramillo, *et al.* 2001. Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology*. **38**: 359-365.
49. Van Zuijen, T.L., E. Sussman, I. Winkler, *et al.* 2004. Grouping of sequential sounds-an event-related potential study comparing musicians and nonmusicians. *J. Cogn. Neurosci.* **16**: 331-338.
50. Vuust, P., E. Brattico, M. Seppanen, *et al.* 2012. Practiced musical style shapes auditory skills. *Annals of the New York Academy of Sciences*. **1252**: 139-146.
51. Tervaniemi, M. 2009. Musicians - same or different? *Ann N Y Acad Sci*. **1169**: 151-156.
52. Daffner, K.R., M.M. Mesulam, P.J. Holcomb, *et al.* 2000. Disruption of attention to novel events after frontal lobe injury in humans. *J.Neurol.Neurosurg.Psychiatry*. **68**: 18-24.
53. Schröger, E., M.H. Giard & C. Wolff. 2000. Auditory distraction: event-related potential and behavioral indices. *Clinical Neurophysiology*. **111**: 1450-1460.
54. Alho, K., I. Winkler, C. Escera, *et al.* 1998. Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. *Psychophysiology*. **35**: 211-224.
55. Opitz, B., A. Mecklinger, A.D. Friederici, *et al.* 1999. The functional neuroanatomy of novelty processing: integrating ERP and fMRI results. *Cerebral Cortex*. **9**: 379-391.
56. Downar, J., A.P. Crawley, D.J. Mikulis, *et al.* 2000. A multimodal cortical network for the detection of changes in the sensory environment. *Nat.Neurosci.* **3**: 277-283.
57. Knight, R.T., D. Scabini, D.L. Woods, *et al.* 1989. Contributions of temporal-parietal junction to the human auditory P3. *Brain Res.* **502**: 109-116.
58. Friedman, D., Y.M. Cycowicz & H. Gaeta. 2001. The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci.Biobehav.Rev.* **25**: 355-373.
59. Knight, R.T. & D. Scabini. 1998. Anatomic bases of event-related potentials and their relationship to novelty detection in humans. *J.Clin.Neurophysiol.* **15**: 3-13.
60. Woods, D.L. 1990. "The physiological basis of selective attention: Implications of event-related potential studies". In: *Event-related brain potentials: Basic issues and applications*. J. W. Rohrbaugh, R. Parasuraman, & R. Johnson, Jr. (Eds.), 178-209. New York: Oxford University Press.
61. Schröger, E. 1996. A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *Journal of Cognitive Neuroscience*. **8**: 527-539.
62. Escera, C., K. Alho, E. Schröger, *et al.* 2000. Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiol.Neurootol.* **5**: 151-166.
63. Feldman, H. & K.J. Friston. 2010. Attention, uncertainty, and free-energy. *Frontiers in human neuroscience*. **4**: 215.
64. Madison, G. 2006. Experiencing groove induced by music: Consistency and phenomenology. *Music Perception*. **24**: 201-208.

65. Madison, G., F. Gouyon, F. Ullén, *et al.* 2011. 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.
66. Janata, P., S.T. Tomic & J.M. Haberman. 2012. Sensorimotor coupling in music and the psychology of the groove. *J. Exp. Psychol. Gen.* **141**: 54-75.
67. Stupacher, J., M.J. Hove, G. Novembre, *et al.* 2013. Musical groove modulates motor cortex excitability: A TMS investigation. *Brain Cogn.* **82**: 127-136.
68. Witek, M.A.G., E.F. Clarke, M. Wallentin, *et al.* 2014. Syncopation, body-movement and pleasure in groove music. *PloS one*. **9**: e94446.
69. Clarke, E.F. 1989. The perception of expressive timing in music. *Psychological Research*. **51**: 2-9.
70. Repp, B.H. 1999. Detecting deviations from metronomic timing in music: effects of perceptual structure on the mental timekeeper. *Percept.Psychophys.* **61**: 529-548.
71. Danielsen, A. 2010. *Musical rhythm in the age of digital reproduction*.Ashgate: Farnham.
72. Witek, M.A. 2016. Filling in: syncopation, pleasure and distributed embodiment in groove. *Music Analysis*.**36:** 138–160
73. Witek, A.G.C., E.F.;Kringelbach,M.L.; Vuust, P. 2014. Effects of Polyphonic Context and Instrumentation on Syncopation in Music. *Music Perception*, **32:** 201 - 217 .
74. Wundt, W. 1874. *Grundzuge der physiologischen psychologie*. Leipzig: Englemann.
75. North, A.C. & D.J. Hargreaves. 1995. Subjective complexity, familiarity, and liking for popular music. *Psychomusicology*. **14**: 77-93.
76. North, A.C. & D.J. Hargreaves. 1997. "Experimental aesthetics and everyday music listening". In *The social psychology of music*. D.J. Hargreaves & A.C. North, Eds. Oxford: Oxford University Press.
77. Orr, M.G. & S. Ohlsson. 2005. Relationship between complexity and liking as a function of expertise. *Music Perception*. **22**: 583-611.
78. Berlyne, D.E. 1971. *Aesthetics and psychobiology*. East Norwalk, CT: Appleton-Century-Crofts..
79. Witek, M.A., E.F. Clarke, M. Wallentin, *et al.* 2015. Correction: Syncopation, Body-Movement and Pleasure in Groove Music. *PloS one*. **10**.
80. Friston, K. 2003. Learning and inference in the brain. *Neural Netw.* **16**: 1325-1352.
81. Brown, H., R.A. Adams, I. Parees, *et al.* 2013. Active inference, sensory attenuation and illusions. *Cogn Process*. **14**: 411-427.
82. Grahn, J.A. & M. Brett. 2007. Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*. **19**: 893-906.
83. Chen, J.L., V.B. Penhune & R.J. Zatorre. 2008. Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *J Cogn Neurosci*. **20**: 226-239.
84. Nozaradan, S., I. Peretz, M. Missal, *et al.* 2011. Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*. **31**: 10234-10240.
85. Velasco, M.J. & E.W. Large. 2011. Pulse detection in syncopated rhythms using neural oscillators. *pulse*. **1**: 3-4.
86. Meyer, L.B. 1956. *Emotion and meaning in music*. Chicago and London: University of Chicago Press..
87. Huron, D. 2006. *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: The MIT Press.

88. Kringelbach, M.L. & K.C. Berridge. 2009. Towards a functional neuroanatomy of pleasure and happiness. *Trends Cogn Sci.* **13**: 479-487.
89. Hohwy, J. 2016. The Self-Evidencing Brain. *Noûs*. **50**: 259-285.
90. Gebauer, L., M.L. Kringelbach & P. Vuust. 2012. Ever-changing cycles of musical pleasure: The role of dopamine and anticipation. *Psychomusicology*. **22**: 152-167.
91. Schultz, W. 2007. Behavioral dopamine signals. *Trends Neurosci.* **30**: 203-210.
92. Schultz, W., K. Preuschoff, C. Camerer, *et al.* 2008. Explicit neural signals reflecting reward uncertainty. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **363**: 3801-3811.
93. Vuust, P., K.J. Pallesen, C. Bailey, *et al.* 2005. To musicians, the message is in the meter. *Neuroimage*. **24**: 560-564.
94. Phillips-Silver, J. & L.J. Trainor. 2005. Feeling the beat: Movement influences infant rhythm perception. *Science*. **308**: 1430-1430.
95. Iyer, V. 2002. Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception: An Interdisciplinary Journal*. **19**: 387-414.
96. Danielsen, A. 2006. *Presence and Pleasure. The funk grooves of James Brown and Parliament*. Middletown, Connecticut: Wesleyan University Press.
97. Davies, M., G. Madison, P. Silva, *et al.* 2013. The effect of microtiming deviations on the perception of groove in short rhythms. *Music Perception: An Interdisciplinary Journal*. **30**: 497-510.
98. Witek, M.A., E.F. Clarke, M. Wallentin, *et al.* 2014. Syncopation, body-movement and pleasure in groove music. *PLoS One*. **9**: e94446.