

REVIEW

Computational modeling of rhythmic expectations: Perspectives, pitfalls, and prospects

Atser Damsma^{1,2*}, Jonathan Cannon³, Lauren K. Fink³, Keith B. Doelling⁴, Jessica A. Grahn⁵, Henkjan Honing^{6,1}, Thomas Kaplan^{6,7}, Edward W. Large⁸, Fleur L. Bouwer^{9*}

1 Music Cognition Group, Institute for Logic, Language, and Computation, Amsterdam Brain and Cognition, University of Amsterdam, Amsterdam, The Netherlands, **2** Conservatorium van Amsterdam, Amsterdam University of the Arts, Amsterdam, The Netherlands, **3** Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario, Canada, **4** Université Paris Cité, Institut Pasteur, AP-HP, INSERM, CNRS, Fondation Pour l'Audition, Institut de l'Audition, IHU reConnect, Paris, France, **5** Department of Psychology & Centre for Brain and Mind, Western University, London, Ontario, Canada, **6** William Harvey Research Institute, Barts and the London Faculty of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom, **7** Barts Biomedical Research Centre, National Institute of Health and Care Research, Queen Mary University of London, London, United Kingdom, **8** Center for the Ecological Study of Perception and Action, Department of Psychological Sciences and Department of Physics, University of Connecticut, Storrs, Connecticut, United States of America, **9** Cognitive Psychology Unit, Institute of Psychology & Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands

* atser.damsma@ahk.nl (AD); f.l.bouwer@fsw.leidenuniv.nl (FLB)



OPEN ACCESS

Citation: Damsma A, Cannon J, Fink LK, Doelling KB, Grahn JA, Honing H, et al. (2025) Computational modeling of rhythmic expectations: Perspectives, pitfalls, and prospects. *PLoS Comput Biol* 21(12): e1013798. <https://doi.org/10.1371/journal.pcbi.1013798>

Editor: Alejandro Tabas, Basque Center on Cognition Brain and Language, SPAIN

Published: December 10, 2025

Copyright: © 2025 Damsma et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: AD is supported by an Amsterdam Brain and Cognition (ABC) Project Grant (20200529). JC is supported by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2022-05027). LF is supported by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2023-05050). KBD is supported by

Abstract

Rhythmic structure enables precise temporal expectations that are essential to human communication, including speech and music. Computational models have been developed to account for how humans perceive, produce, and learn rhythmic sequences. However, it is unclear how different types of models relate to each other and how they can be evaluated. In this review and perspective, we discuss how three major classes of models—entrainment, probabilistic, and timekeeper models—have been used to study rhythmic expectations. We critically assess each model class in terms of its level of explanation, the rhythmic behaviors it captures, its ability to account for learning and enculturation, and its ability to integrate other features, such as pitch. We show that entrainment, probabilistic, and timekeeper models differ substantially in the aspects of rhythmic expectations they can capture. To move the field forward, we propose that model comparison and integration are crucial. We identify key challenges to this effort, such as the varying nature of the input and output signals and divergent modeling goals. To address these challenges, we arrive at several practical recommendations: to equate input and output signals when comparing models, to consider several model outcomes beyond goodness-of-fit measures in model evaluation, to use model-integration efforts to inform theory building, and to make code and data openly accessible. Ultimately, understanding how models of rhythmic

a French government grant managed by the Agence Nationale de la Recherche under the France 2030 program (ANR-23-IAHU-0003) and the Fondation Pour l'Audition. JAG is supported by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2024-05869). HH is supported by an Open Competition grant from the Dutch Research Council NWO (406.20.CW.002). TK is supported by the National Institute for Health and Care Research Barts Biomedical Research Centre (NIHR203330). EWL is supported by NIH grant 1R01AG078376. FLB is supported by a VENI grant from the Dutch Research Council NWO (VI.Veni.201G.066). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: EWL holds ownership interest in Oscillo Biosciences. All other authors have declared that no competing interests exist.

expectations relate, and how features in these models account for behavioral, neural, and cognitive aspects of rhythmic expectations, will deepen our understanding of a core aspect of human behavior.

1. Introduction

In our dynamic environment, events unfold over time. The way events are structured in time is commonly referred to as *rhythm* [1]. Rhythmic structure allows us to predict event timing, and thereby, to optimize processing [2,3]. Predictable rhythms allowing for precise temporal expectations are found in many natural phenomena and signals, but are especially prominent in language and music. In music, rhythmic expectations have been linked to reward and pleasure [4], social bonding [5], and movement facilitation [6].

Computational modeling is an essential tool in understanding rhythmic behavior and cognition, aiming to describe the mechanisms underlying rhythmic expectations. Computational models formalize theories of human cognition beyond the verbal level to concrete mathematical or programmatic formulations that can be tested against behavioral and neural data [7]. This has important advantages: first, it enables us to directly compare models representing competing theories [7–9]. For example, comparing competing computational models to human data can elucidate whether synchronization relies on estimating concrete durations (*absolute timing*) or on timing relative to a regularity or a previous duration (*relative timing*) [10,11]. It could also shine light on competing theories about the neural networks supporting rhythm perception, such as the Vocal Learning [12] and Gradual Audiomotor Evolution [13] hypotheses. Second, modeling allows for parameter manipulation to explain individual differences meaningfully [14]. Third, in-silico 'lesioning' allows for a systematic examination of how performance is affected by model components, mimicking how brain pathology affects behavior [7].

Different types of computational models capture how listeners perceive and produce rhythm, but it is unclear how these models may be related [15,16] and evaluated [8], hampering our understanding of rhythmic expectations. In this critical review and perspective, we will therefore consider the current state of the field and explicitly address ways to move it forward.

In our discussion, we use the term *rhythm* to refer to any sequence of multiple temporal intervals, including various types of non-isochronous patterns. We focus on models capturing rhythm processing in real time, as it unfolds, excluding models that can only consider a full rhythmic pattern at once [17]. Moreover, we only include models of cognition and brain activity, which formalize some theoretical account [7] of how listeners form rhythmic expectations. This excludes generative AI models [18], which, while arguably successful at processing musical signals, have not yet explicitly contributed to understanding human rhythm cognition. Finally, we focus on models of non-linguistic rhythm. While rhythm is essential to language, linguistic rhythm differs from musical rhythm in several ways. First, whereas periodicity plays an important

role in non-linguistic rhythm, it has been argued that language is only quasi-periodic [19]. Second, the semantic content of language may influence its rhythmicity [19,20]. Third, while we focus on rhythm as continuous timing, research on linguistic rhythm typically emphasizes lexical stress patterns and grouping [19].

Given these constraints, we discuss three broad classes of models (Fig 1A): entrainment, probabilistic, and timekeeper models. These three classes capture the current state-of-the-art in modeling rhythm processing dynamically, as time unfolds [1]. We first introduce each model class and their specific implementations (see Table 1 for a non-exhaustive overview). We then show how these classes differ fundamentally in their level of explanation, the rhythmic representations and learning processes they capture, and their integration of features beyond auditory timing (Fig 1B–1E). Entrainment models take neural mechanisms as their starting point and have been linked to neural markers of periodicity, such as beat and meter; probabilistic models are rooted in cognitive theory and typically capture behavioral markers of symbolic sequences prediction; and timekeeper models emphasize isochronous motor synchronization. Given these discrepancies, we discuss strategies to move the field forward through model comparison and integration and offer practical recommendations for future modeling studies on rhythmic expectations.

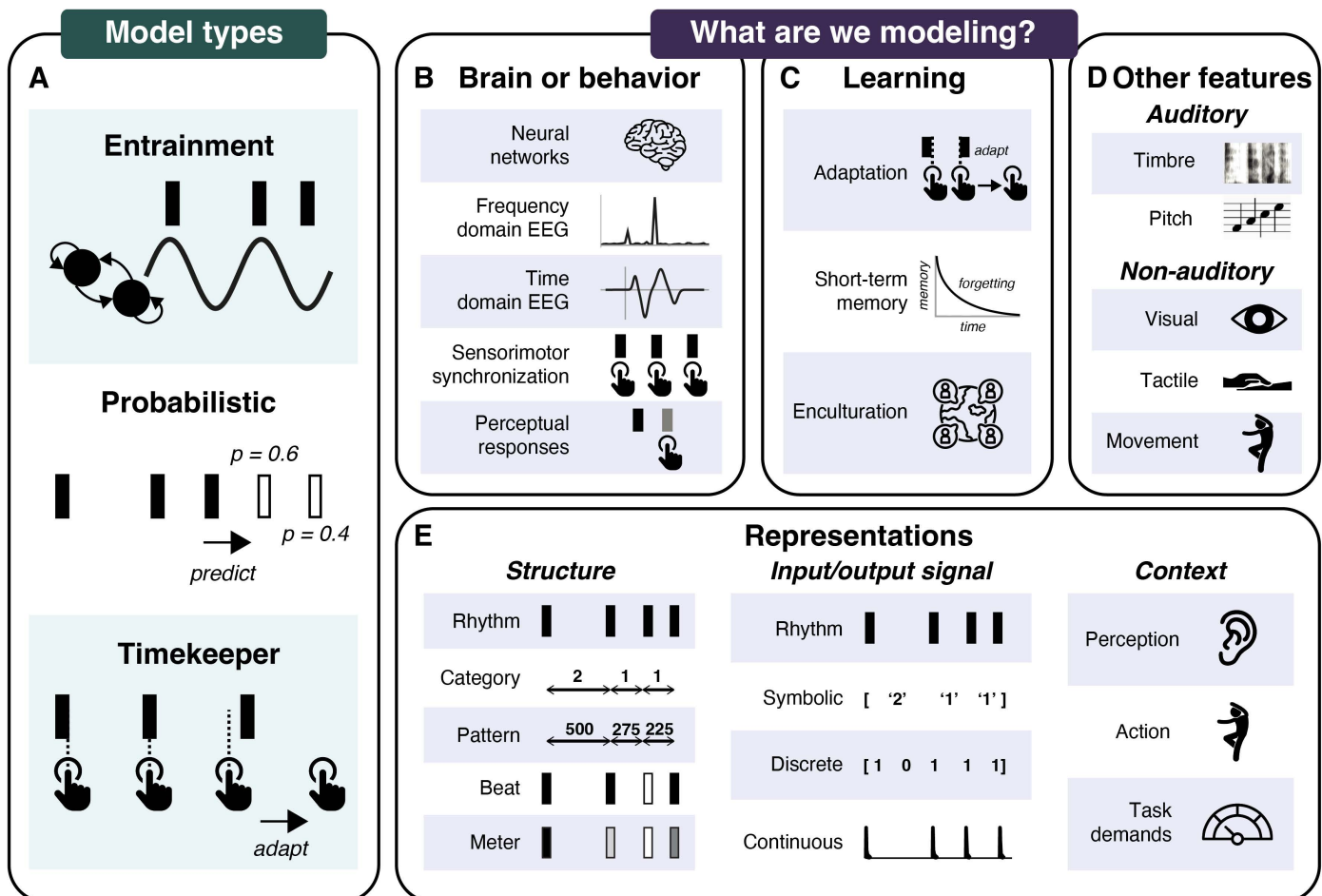


Fig 1. Schematic overview of three classes of models (Section 2), and their targets in terms of modeling (Section 3). Different models (A, Section 2) operate at different levels of explanation (B, Section 3.1.), and target different rhythmic representations (E, Section 3.2.). Models also diverge in their treatment of learning (C, Section 3.3.) and non-temporal or non-auditory features (D, Section 3.4.). Icons were licensed through *Icon Pro* from *The Noun Project* (<https://thenounproject.com/>). Spectrogram in Fig 1D by Aaron Parecki (<https://www.flickr.com/photos/aaronpk/4947807970>) [CC BY 2.0].

<https://doi.org/10.1371/journal.pcbi.1013798.g001>

Table 1. Overview of different models of rhythmic expectations.

Model class	Type of model	Data domain	Phenomenon studied
Entrainment	Bank of linear oscillators	Continuous time	Beat, meter [21–24]; Groove [23]; Sensorimotor synchronization [23,24]
	Multiple connected banks of coupled nonlinear oscillators	Continuous time	Beat, meter [25–28]; (Onset) Pattern [26]; Enculturation [26]; Groove [28]
	Multiple coupled nonlinear oscillators	Continuous time	(Quasi-)Isochrony [29–31]; Beat, meter [32–34]; Interpersonal synchronization [29–31]; Auditory-motor interactions [34]; Expressive timing [33]; Spontaneous motor tempo [31]
	Single nonlinear oscillator	Continuous time	(Quasi-)Isochrony [35–39]; Beat, meter [35,40]; Interpersonal synchronization [38]; Gait [37]
Probabilistic	PPM (prediction by partial matching); Markov chains; Bayesian	Symbolic/categorical	Meter [41,42]; (Onset) Pattern [11,43–46]; Enculturation [41]; Groove [47]
	Phase inference from point process event timing (PIPET)	Continuous time	(Quasi-)Isochrony [48–50]; Beat, meter [48–50]; (Onset) Pattern [51]; Meter [51]; Enculturation [51]; Expressive timing [50]
	Phase inference combined with PPM	Continuous time	Beat, meter, pattern [52]
Timekeeper	Timekeeper with error correction	Continuous value	(Quasi-)Isochrony [53–55]; Interpersonal synchronization [53,55]
	Timekeeper with error correction and anticipation	Continuous value	(Quasi-)Isochrony [56–59]; Interpersonal synchronization [56–58]
Hybrid	Single nonlinear oscillator with probabilistic input for period adjustment	Continuous time	(Quasi-)Isochrony [10]
	Single nonlinear oscillator with added phase/period correction	Continuous time	(Quasi-)Isochrony [60]

Note: This table provides a non-exhaustive overview of models, to illustrate their diversity, and the various domains and rhythmic phenomena studied. Data domains are divided into continuous time (a value outputted at every time point, e.g., a continuous time signal), continuous value (a value outputted on a continuous scale, e.g., the next expected IOI on a continuous scale) or symbolic (a symbolic value as output, e.g., the next expected IOI from a limited set of possible IOIs). Concerning the phenomena studied: we use the term “(quasi-)isochrony” to denote not just isochronous sequences (e.g., metronomes), but also sequences derived from isochrony, with tempo or phase perturbations, or variability introduced. We consider these quasi-isochronous, as they do not explicitly contain higher order structure as found in beat and meter.

<https://doi.org/10.1371/journal.pcbi.1013798.t001>

2. Models capturing rhythmic expectations

2.1. Entrainment models

Human music often contains a high degree of (quasi-)periodicity, which is perceived by listeners as a regular *beat*. Entrainment models capitalize on this property of musical rhythm. Broadly, *entrainment* refers to the phase alignment between two periodic signals. In theories of rhythmic expectations, entrainment involves a (quasi)periodic process in the nervous system—such as low-frequency cortical oscillations—becoming phase-aligned with external rhythmic input [61]. Cognitively, this alignment may cause fluctuations in perceptual performance [2]. Motorically, entrainment may lead to the synchronization of movement with rhythmic input or coordination between multiple individuals [15,29].

Computational models of entrainment typically assume an (endogenous) oscillator which adapts its phase and period to an external regularity, but vary significantly in their properties. First, some models, such as the Beyond-the-Beat model, use linear oscillators, which essentially act as filters [21,24]. Others employ non-linear oscillators, which enable *mode-locking*, with entrainment occurring at frequencies that are related to the eigenfrequency of the oscillator at integer ratios [1]. Second, the flexibility of an oscillator to adapt to frequencies other than its eigenfrequency can be manipulated [40]. Third, whether oscillators are self-sustaining depends on the choice of damping parameters [39,62]. Fourth, the strength of phase coupling between the external stimulus and the oscillator further determines how quickly and robustly entrainment occurs [30,62]. Finally, instead of relying on a single oscillator, many entrainment models—including the Beyond-the-Beat model and gradient frequency neural networks—employ a bank of oscillators reflecting the frequency range of human rhythm perception [21,22,24,27]. The behavior of these models depends on the number of oscillators, their center frequencies, and whether

oscillators are interconnected, with phases either coupled [25] or uncoupled [26]. More complex architectures may include interconnected networks of oscillators, mimicking the interaction between auditory and motor systems [1,63].

2.2. Probabilistic models

Probabilistic models are often situated within the predictive coding framework [64,65]. This framework proposes that the brain constantly generates predictions of sensory input based on a probabilistic mental model of the world. These top-down predictions are compared against the actual input, and prediction errors are used to update the mental model in an approximation of Bayesian inference. In the domain of rhythm perception, probabilistic models aim to uncover the temporal structure that underlies a sequence of sounds (i.e., the mental model), to predict the timing of future events (Fig 1A). Probabilistic models can provide event-by-event measures such as *surprisal* (how unexpected an event is) and *uncertainty* (the precision of the prediction), which have been shown to match human ratings of rhythmic complexity [44,47]. Through Bayesian updating, rhythmic expectations can be revised dynamically in response to new sensory information. Additionally, long-term listening experiences can be encoded as *priors* [41,42,47,51].

Like entrainment models, probabilistic models vary in their computational properties. Models based on IDyOM (Information Dynamics of Music) [42] represent rhythm as a sequence of symbols, with transitional probabilities between them learned using Markov chains [11]. In these models, the *order-bound* parameter determines the maximum length of history (context) the model is allowed to consider when predicting the next symbol. Recently, an alternative was proposed with PIPPET (Phase Inference from Point Process Event Timing), which formalizes rhythm perception as a phase inference problem [48] and functions in continuous time. This model tracks phase within a sequential or cyclical 'expectation template': a predefined [48] or inferred [51] representation of an anticipated (metrical) pattern. The event likelihood precisions determine the sharpness of these temporal expectations. The estimated phase and its uncertainty are adjusted based on incoming events, where the accumulation of temporal uncertainty is governed by the diffusion rate. Interestingly, a cyclical version of the model [49] shows properties that are mathematically equivalent to a damped oscillator, suggesting that probabilistic and entrainment models may align functionally.

2.3. Timekeeper models

Timekeeper models use a linear autoregressive process to model the timing of an event based on the time interval and asynchrony associated with the previous event (Fig 1A) [58]. These models propose an error-correction mechanism to account for timing and motor variability, relying on two separate mechanisms: period and phase correction [53]. The magnitude of these corrections is controlled by the period correction rate and phase adaptation rate, respectively [58]. One implementation, the Adaptation and Anticipation Model (ADAM), also includes an anticipation process [56,58] in which upcoming events are anticipated through linear extrapolation of previous intervals [59]. Anticipation is parameterized by the number of intervals used for extrapolation and the prediction-tracking balance. In addition, the weighting between reactive adaptation and anticipation can be adjusted [56].

More recently, a neuro-mechanistic model was proposed in which a biophysically-based beat generator neuron (BG) synchronizes with a (quasi)isochronous sequence [60]. The period and phase of the BG are adjusted based on a comparison between the number of gamma-frequency cycles between two successive BG spikes and a similar cycle count between stimulus onsets. As such, the BG model can be seen as a neuronal implementation of the error-correction mechanisms that characterize timekeeper models.

3. What are we trying to model? A critical comparison of existing models

3.1. Behavior or brain mechanisms

While the models discussed above are all implemented algorithmically [66], they operate at different explanatory levels (Fig 1B). Probabilistic models mainly account for behavioral or neural markers that indirectly index rhythmic expectations, such as event-related potentials (ERPs) [67,68], without explicitly specifying the underlying neural processes. General predictive processing models

hypothesize that higher-level brain areas provide predictions to lower-level areas via feedback connections, with lower-level areas encoding prediction errors [65]. However, these models often focus on *content* expectations (i.e., predicting “what”, see [69]) instead of *temporal* expectations (i.e., predicting “when”), raising questions about the relevance of this neural model for rhythm prediction. Notably, a recent model based on the Action Simulation for Auditory Prediction (ASAP) hypothesis does attempt to specify the neural networks underlying rhythm processing—including the motor system [70], but this model currently lacks an algorithmic implementation.

In contrast to probabilistic models, entrainment models, like those based on neural resonance theory [1], take the neural level as their starting point. Aspects of behavior emerge from well-studied neural mechanisms (e.g., neuronal oscillation, Hebbian plasticity, synaptic transmission delay, see [1]), instead of being directly implemented. For example, *negative mean asynchrony*—the tendency to tap earlier than note onsets when synchronizing to rhythm [71]—may be parsimoniously explained as a generic effect of neural transmission delays [31,38,72]. Entrainment models of rhythmic expectations have been tested against both neural and behavioral data. They can mimic the selective enhancement of power at beat frequencies in response to non-isochronous rhythm found in human EEG data (*frequency tagging*, see [25]) and motor synchronization [25]. However, frequency tagging as evidence for neural entrainment should be interpreted cautiously, as it may alternatively reflect ERPs to rhythmic stimuli [73,74] or sequential anticipatory neural activity [75].

Timekeeper models have mostly been applied to behavioral sensory-motor synchronization (SMS): the coordination of movement with an external rhythm [71], be it the synchronizing of finger tapping to rhythm, adaptive synchronization behavior between two humans [76], or between humans and “virtual partners” [77]. One fMRI study has associated ADAM’s adaptation and anticipation parameters to specific brain areas, including the cerebellum, basal ganglia, and thalamus [57], linking the behavioral to the neural level.

3.2. Types of representation

3.2.1. Types of structure. In addition to differences in explanatory level, models handle the inherent multiplicity of rhythmic structure (Fig 1E) in different ways. Listeners infer structure from rhythmic sequences by learning predictable *patterns* (i.e., recurring successions of longer and shorter intervals [11]), and by inferring a *beat* (i.e., isochronous pulses) and *meter* (i.e., a hierarchical organization of beats [1]). While these processes may interact [51,78], evidence from behavioral [79] and neural studies [80] suggests they are at least partially dissociable.

Many models exclusively target (quasi-)isochronous sequences. Isochrony processing is often studied by examining reactions to timing perturbations [10,29,30,36,37,39,54,58]. Human responses to perturbations have successfully been modeled using timekeeper models with explicit parameters for phase and period correction [58], and networks of oscillators, like in research examining interpersonal synchrony [30], where the complexity of multiple oscillators was necessary to explain human behavior. Additionally, human temporal fit judgments after quasi-isochronous sequences were modeled using a hybrid adaptive oscillator model, in which a probabilistic aspect was added to a single eigenfrequency entrainment model to account for timing variability in the stimulus [10]. Various classes of models have thus been successful in explaining human responses to (quasi-)isochronous sequences. However, since pattern, beat, and meter may be simultaneously present in isochronous sequences, it is unclear what rhythmic percept these models target. Also, they may not handle complex non-isochronous sequences well, as is apparent for some timekeeper models [40].

Several probabilistic models have explicitly targeted rhythmic patterns [11,43,44,46,51]. Markov-chain models using symbolic representations of absolute inter-onset intervals (IOIs) were used to predict complexity ratings [44] and tapping accuracy [43] in response to rhythmic patterns. Rather than absolute IOIs, other probabilistic models have used representations of ratios between successive intervals [46,51]. While probabilistic models may naturally fit pattern learning, non-isochronous rhythmic patterns can also be learned by an entrainment model using Hebbian plasticity [26]. Here, patterns could be categorized based on oscillator amplitude at different frequencies. However, the input sequences used also contained hierarchical structure represented by loudness differences between sounds, making it unclear whether the output only reflected the temporal features of the pattern.

Hierarchical structure (i.e., meter) has been explicitly targeted by probabilistic models estimating the probability of a sounded event at different metrical positions [41,47]. Here, the meter needs to be known beforehand [47], or the model infers the most likely meter from a set of candidates based on the observed rhythmic pattern [41]. This is similar to PIPPET, which uses cyclical expectation templates [48]—matching the cyclic nature of meter [49]—to predict rhythmic patterns. This type of model has the potential to integrate probabilistic predictions of both rhythmic patterns and meter. However, arriving at such integrated expectations likely requires a bootstrapped learning process, since meter must be inferred from the rhythmic pattern, while the inferred meter is, in turn, used to predict pattern onsets.

Entrainment models often specifically target beat and meter induction. The perceived metrical structure is deduced from the spectral content of the model's output across an entire rhythmic sequence, with the most prominent frequency assumed to represent the perceived beat [24,25,27]. This raises the question how these models deal with non-periodic meters, which are common in many non-Western musical cultures. Whereas the flexibility of the expectation template in PIPPET explicitly allows for such meters [48], they may be a challenge for entrainment models [81].

3.2.2. Types of input and output signals. Models differ in the input they accept and the type of output they generate (Fig 1E). Inputs and outputs can be *symbolic* (e.g., categorical representations of temporal intervals [44] or metrical position [41]), or *discrete* (e.g., a binary decision whether a position on a rhythmic grid is filled with an event or not [45,47]), but such representations cannot capture processing of natural rhythms with continuous timing and tempo variations, which would require infinite symbolic event or grid representations. Alternatively, *continuous values* (e.g., the next IOI to be produced with timekeeper models [54,58]), or *continuous time series* have been used to represent fluctuations in rhythmic expectations [10,21,22,25,27]. However, continuous output signals are often reduced to frequency-domain representations for model evaluation, with the largest spectral peak interpreted as the perceived beat frequency [25,27]. This reduction makes it unclear how the model output translates to real-time behavior (e.g., are a listener's rhythmic expectations indexed by the signal's amplitude, phase, or both?). Models like PIPPET [48] and the Beyond-the-Beat linear oscillator model [21] avoid this reduction by processing continuous input and producing continuous output.

3.2.3. Perception, action, and task demands. Both motor synchrony and perceptual advantages at expected time points have been interpreted as reflecting rhythmic expectations. While synchronization and perceptual rhythmic abilities are indeed correlated, and motor behavior strengthens the perceptual effects of rhythmic expectations, perception and action have also been partially dissociated [82,83]. Indeed, how rhythmic expectations are formed may depend on goals and task demands [84], but these are often not explicitly modeled. Timekeeper models have, to our knowledge, been exclusively designed for and applied to motor behavior. This raises the question whether the error-correction mechanisms in timekeeper models also apply to perceptual processing. Probabilistic models have been applied to both production and perception [41,46]. One variable-order Markov model was shown to explain behavior across three different perceptual and tapping tasks within the same participants [11], suggesting that such models index general underlying rhythmic expectations, which influence a variety of behavioral and perceptual markers. Entrainment models have similarly been applied to both perception [21,22,25,26] and action [23,29,31]. Some of these models explicitly contain separate perceptual and motor networks [25]. To study the relationship between perception and action, and the influence of task demands on rhythmic expectations, one approach would be to enable both an active and a passive mode in such models [83]. One study explicitly modeled task demands using a network of coupled oscillators [34], adjusting the coupling strength between perceptual, attention, and motor oscillators to mimic behavioral data under different conditions. Similarly, audio-motor interactions were explicitly modeled to account for the pleasurable urge to move to musical rhythm [28], integrating aspects of rhythmic perception and action.

In summary, the rhythmic aspect, the type of signal, and the type of task that are modeled vary greatly between models. Table 1 provides a (non-exhaustive) overview of model implementations and their targeted rhythmic aspects and signals.

3.3. Flexibility and learning

Rhythmic expectations are flexibly shaped by both short- and long-term learning processes (Fig 1C). When synchronizing with a rhythm, humans can rapidly adapt their movement to phase perturbations and tempo-changes [35,39], reflecting crucial short-term flexibility. Entrainment models can likewise adapt to changing regularities within a few cycles [31,40], with the magnitude of adaptation determined by the coupling strength between the rhythmic stimulus and the oscillator [32]. In timekeeper models such as ADAM, adaptation is explicitly incorporated, since the next tapping interval is updated based on the current asynchrony, with separate parameters for phase and period correction [58]. A probabilistic implementation of phase and tempo adaptation can be found in PIPPET, where adaptation is based on recent timing information and depends on uncertainty of the predictions [48,49].

While all of these models thus adapt to small timing perturbations, they do not consider them meaningful rhythmic structure. However, deviations in the order of tens of milliseconds (*microtiming* or *expressive timing*) are common in music performance and may convey relevant rhythmic information on their own, clarifying compositional structure and determining whether a performance feels “pushed” or “laid back” [50]. One oscillator model was shown to sustain synchronization when faced with temporal fluctuations, which interestingly could even improve its tracking ability [33]. Recently, PIPPET was shown to discriminate between drum patterns with variable microtiming profiles, and human individual differences in discrimination were explained by the amount of phase-tracking noise in the model [50]. These efforts stress the importance of modeling rhythm in continuous time to account for microtiming.

Besides flexibility, models need to account for short-term memory. The recent probabilistic PPM-Decay model of auditory sequence prediction implemented an exponential memory decay function, in which the weight of events decreases over time [85], to account for recency biases and capacity limitations in human short-term memory. Memory decay indeed improved prediction of chord sequences [85] and tapping behavior [11]. Timekeeper model ADAM similarly includes an anticipatory module which extrapolates the previous (at least three) taps, essentially acting as a short-term memory buffer for prediction [58]. Likewise, in PIPPET [48], phase uncertainty increases over time, reducing the weight of events as they recede into the past. In entrainment models, oscillator decay time can be seen as an implicit short-term memory parameter [8]. Thus, different models can capture how sequence statistics change over time, as in naturalistic rhythm.

Finally, in production or discrimination tasks, cultural familiarity with the presented rhythms can improve accuracy [86], showing long-term learning. In coupled non-linear oscillator models, Hebbian plasticity has been used to simulate perceptual narrowing [26,63]. Through repeated exposure to specific rhythmic structures, adjustment of coupling coefficients between oscillators improves the network’s response to similar structures. In probabilistic models, such as IDyOM, long-term statistical learning has been implemented by estimating the prior probability of certain rhythmic sequences based on a musical corpus approximating the listening experience of individuals [42]. While such probabilistic models of enculturation typically depend on symbolic representations of musical rhythms, continuous probabilistic phase inference has also been used to simulate enculturation based on the statistical distribution of rhythmic sequences in an empirical sample [51]. Parameter optimization within probabilistic generative models has been directly related to Hebbian plasticity [64], suggesting these two experience-dependent learning mechanisms might be functionally related.

Short- and long-term learning processes have also been integrated into a single model. Probabilistic sequence-learning models have used separate models with different rules for online learning and retaining observations, before combining them to make rhythmic predictions [41]. Recently, sequence-learning models and PIPPET were combined, such that the expectation template used by PIPPET for real-time phase inference depends on discretized models of (learned) underlying sequence structure [52]. In coupled oscillator models, Hebbian plasticity is proposed to operate along different timescales: short-term transient plasticity enables neural entrainment to a stimulus, while longer-term stable plasticity accounts for perceptual narrowing to culturally familiar rhythms [63]. However, these models addressing multiple timescales of learning have so far seen limited empirical evaluation and comparison.

3.4. Interaction with non-temporal and non-auditory features

Rhythm cannot exist without the spectral features of the sounds carrying it (Fig 1D). Neural synchronization [87], perceived tempo [88], and beat [89] all depend on pitch and timbral aspects of sound. The Beyond-the-Beat model [24], an entrainment model that takes the continuous sound signal as input, considers such spectral information: the same rhythm performed on different instruments elicits different output, though melodic and harmonic expectations are not modeled explicitly. In contrast, the probabilistic IDyOM model uses linked viewpoints to represent concurrent expectations for pitch and rhythm [42], but spectral information is discretized into symbolic input. Most models of rhythmic expectations, however, are focused exclusively on timing-related information, taking only event onsets as input. Beyond spectral content, non-auditory modalities—such as visual and vestibular information—may interact with auditory rhythm processing [90,91]. As such, the field's focus on signals restricted to timing information is a potential major limitation.

3.5. Interim summary

In summary, different classes of models capture rhythmic expectations in distinct ways. Entrainment models provide a biologically plausible link between brain mechanisms and rhythmic behavior, flexibly mimicking a wide range of rhythmic phenomena from continuous audio input. However, their large parameter space risks making them descriptive rather than explanatory, and analysis mostly focuses on inferring beat and meter from summary statistics instead of real-time model output. Probabilistic models are well-suited for integrating rhythm with non-temporal features, combining long-term enculturation and short-term adaptation, and potentially unifying metrical and pattern-based expectations. Yet, they lack direct correspondence to underlying neural mechanisms and typically rely on symbolic input or pre-defined templates. Finally, timekeeper models provide a compact framework for describing sensorimotor synchronization, but their narrow focus on motor behavior and isochrony leaves unresolved whether they can explain more complex aspects of rhythm perception.

4. Moving rhythm modeling forward

Given the model classes' divergent assumptions and foci outlined above, how can we progress? We propose two key avenues: model comparison and integration (Fig 2).

4.1. Model comparison

4.1.1. Requirements for comparison. **4.1.1.1. Similar input.** Imagine that we want to feed multiple corpora of stimuli through different models to generate sample-wise predictions of rhythmic expectations, to ultimately systematically compare model outputs. We first need to answer two important questions: (1) what type of stimuli do we have (e.g., a symbolic representation, audio), and (2) what type of input is each of our models expecting? While sometimes considered trivial, these questions have important implications for our operationalization of “rhythm” and our interpretation of model results.

While some models operate on raw audio [21,45], many use symbolic representations [41,42,44] or occupy a middle ground where raw audio is converted into a continuous signal representing only onsets [25,48] (Fig 1E). When we compare models operating on symbolic and continuous input, we have to convert continuous data to symbols. For symbols (e.g., A4), the only definable operations are ones of equivalence (e.g., A4! = B4), so we lose spectrotemporal acoustic phenomena, as well as order (e.g., 3 < 5) and scale information (e.g., 4 = 2*2), in the process. The nature of this conversion is therefore crucial. For example, taking inspiration from human categorization [51,78,92], IOIs could be quantized into integer-ratio categories, each with their own symbol. This way, we propose interfacing between real-time and symbolic representations should be cognitively plausible.

4.1.1.2. Similar output. One could propose that we can only directly compare models that target the same aspect of rhythm, at the same level of explanation, with similar output (e.g., “meter in continuous EEG data”). However, regardless

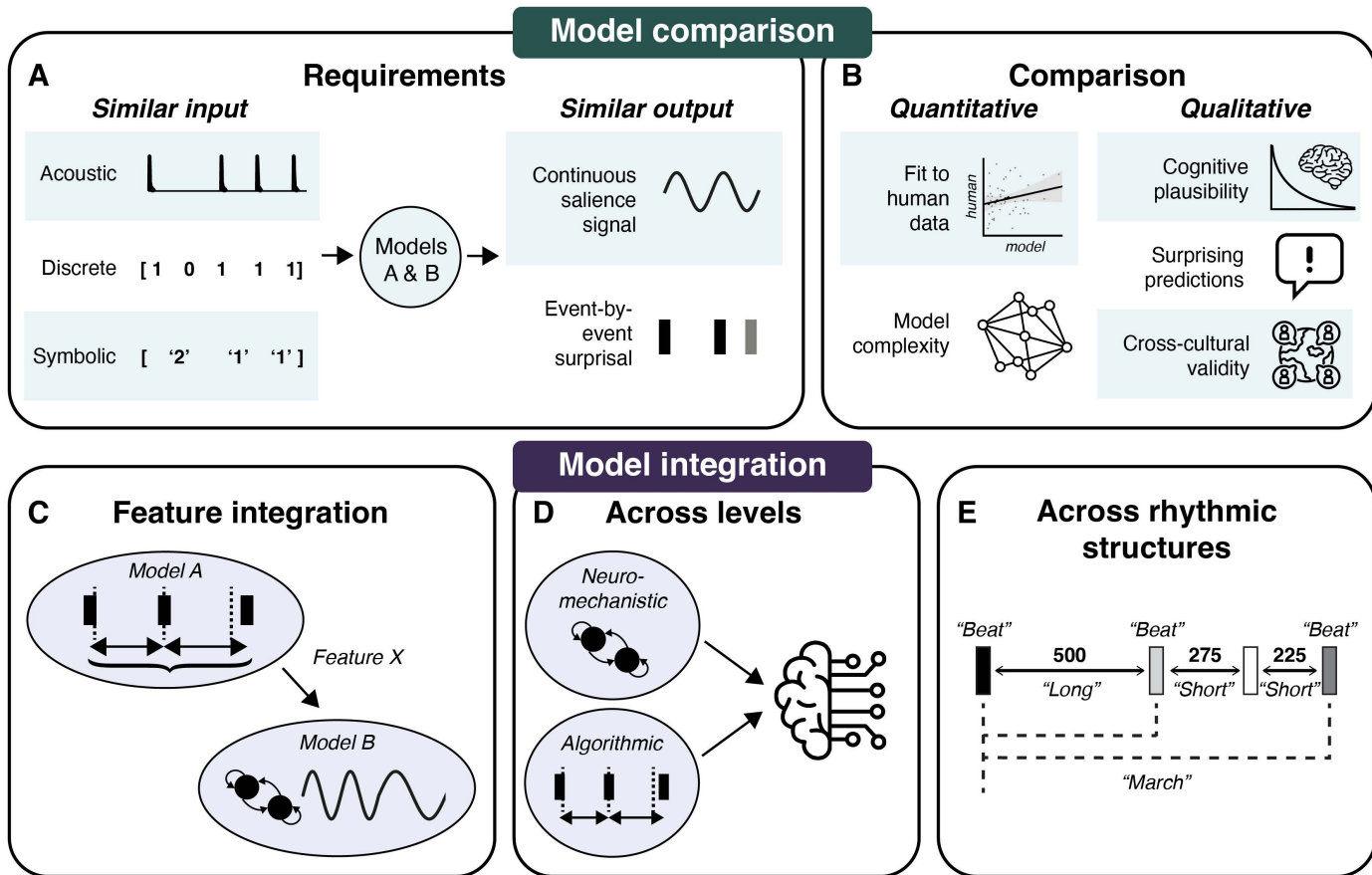


Fig 2. Proposed avenues for moving the field of rhythm modeling forward: model comparison and integration. To compare models (Section 4.1.), input and output signals must be brought into the same space (A). Model comparison may subsequently consider both qualitative and quantitative criteria (B). Ultimately, model integration (Section 4.2.) is a critical step to advance the field, through the integration of features from different classes of models (C), integration of models at different levels of explanation (D), and integration of models that target different rhythmic phenomena (E). Icons were licensed through *Icon Pro* from *The Noun Project* (<https://thenounproject.com>).

<https://doi.org/10.1371/journal.pcbi.1013798.g002>

of how rhythmic expectations are formed, from a functional perspective, all models of rhythmic expectations should arguably capture something similar: the dynamics of expectations over time [8] that affect behavioral performance and neural excitability [3]. The challenge is therefore to translate the specific construct that is modeled (e.g., neural oscillations, metrical category) to the unifying level of continuous rhythmic expectations.

We see two ways to accomplish this: First, a measure of surprise for each separate event may be extracted from different models, and compared to behavioral or neural data [8] (Fig 2A), as was done for melodic expectations [93]. While this is straightforward for (symbolic) probabilistic models (e.g., using *information content* [42] or *expectancy* [48] at each event or time point), for entrainment models, a non-trivial mapping must be defined from their continuous output to a surprise quantity. Some attempts were made by predicting event times from an oscillator's phase [40], but the role of amplitude—an important feature of entrainment models [1]—is unclear. One model converted the continuous output signal of an oscillator to a salience measure at discrete time points that indeed correlated with behavior [22], showing promise for bringing continuous output into the same space as surprisal values from symbolic models. A second approach to unifying output is a conversion from discrete or symbolic output to continuous time. For example, metrical categories or symbolic sequences could be mapped onto continuous-time probability distributions [49,51].

4.1.2. Model selection. **4.1.2.1. Quantitative considerations.** Once input and output type are harmonized, models can be directly compared, typically by computing goodness-of-fit (GOF) statistics for their fit to empirical data (Fig 2B). While the fit between a computational model and empirical data is considered a necessary starting point, it is not the end point of model evaluation [94]. To prevent overfitting, models of least complexity and flexibility may also be preferred, in terms of both functional form and the number of free parameters [95,96]. Targeted, parsimonious models that describe a small subset of rhythmic behavior could be valuable for understanding specific cognitive processes (e.g., tracking isochronous rhythms) and reduce the risk of overfitting to random noise in the data. However, a model that generalizes to different contexts (e.g., more complex rhythmic patterns) without modifications is preferable. Several criteria have been proposed that quantify the trade-off between GOF and complexity (e.g., the Bayesian information criterion; BIC). More advanced Bayesian model selection approaches treat models as random effects, allowing for the possibility that different individuals may be best described by different models [97–99]. At the group level, these methods quantify the probability that each model is more frequent across participants.

While model comparison based on relative fit and complexity is valuable, it does not guarantee that the ‘winning’ model can actually generate the behavioral effects of interest [100]. To allow for *falsification*, an experiment should be designed that can discriminate between competing models [100]. To this end, experimental stimuli could be selected that show maximally opposing predictions for different models [11].

4.1.2.2. Qualitative considerations. Beyond their fit to empirical data and their complexity, computational models may be judged on their cognitive and biological plausibility [95]. While models should not necessarily implement low-level physiological principles, their computations should be consistent with known features of human cognition. A model of rhythmic expectations should, for example, consider limitations and recency effects in human short-term memory, limits in tapping speed [101], limits in tempo changes [102], and categorization of continuous stimuli [78,92]. Models may also aim for consistency with findings from neuroscience, which have demonstrated the importance of motor regions like the basal ganglia, cerebellum, premotor cortex, and supplementary motor area (SMA) for rhythm processing [70]. Mapping high-level algorithmic processes of computational models onto the neural dynamics and proposed functional role of these regions, like beat predictions based on cyclical firing rate trajectories in the SMA that are selectively sequenced by the striatum [70], is a promising way forward [48,103].

A fourth selection criterion is whether a model makes novel, surprising predictions [9], giving it an edge over models that only predict expected outcomes. In the strongest case, a model predicts a phenomenon that is unexpected from known theories, has not yet been observed, is falsifiable, and confirmed empirically [104]. While novel, surprising predictions strongly support a model, known phenomena may also emerge unexpectedly from its principles. Such emergent behavior, not “built in” by design, also enhances model validity (e.g., [48], where time dilation illusions emerge from first principles of Bayesian inference). This criterium raises the question which rhythmic phenomena are particularly surprising (i.e., unexpected without theoretical insight [9]). One example is the *missing pulse* phenomenon, where rhythms lacking spectral power at the beat frequency still evoke a perceptual and neural response of beat, which is surprising under the assumption of a linear response to rhythmic input [25]. Another potentially “surprising” phenomenon is the human bias toward perceiving rhythms as consisting of categorical durations related by low integer ratios (1:1, 1:2, etc.; [51,78,86,92]). This categorization has been linked to interacting brain oscillations at integer-ratio frequencies [1], but may alternatively result from exposure to rhythms produced by humans whose motor behavior is subject to similar constraints [71].

Finally, models may strive for cross-cultural validity, either by demonstrating generalization across different cultures or populations, thereby showing that the model captures universal aspects of human cognition [86], or by providing explicit mechanisms underlying cultural differences, such as by simulating exposure to different musical corpora [41,51]. Similarly, models should account for individual differences, either by generalizing within a population or by capturing individual differences with meaningful variation in model parameters [31,50].

In short, we suggest that models should be judged as more valuable when they (1) fit empirical data best, (2) make clearly delimited predictions, with the least complexity and flexibility (making a model easier to falsify), (3) are cognitively plausible, (4) make unexpected predictions, and (5) explain cognitive universals, enculturated biases, and individual differences.

4.2. Model integration

While comparing models directly is instrumental in scrutinizing and reducing the growing number of rhythm models, a critical next step is to identify how models can be integrated across computational principles and levels of explanation. This could not only lead to a more complete understanding of rhythmic expectations, but would also help assess whether models truly describe distinct cognitive processes or rather describe the same phenomena with different vocabulary (thus aiming to reduce the system to its irreducible parts). Here, we propose three aims for integration: introducing features across models (Fig 2C), integrating models across different levels of explanation (Fig 2D), and integrating across different types of perceived rhythmic structure (Fig 2E).

4.2.1. Adapting features across model types. Using the strengths of disparate models, we can build composite models that more thoroughly reflect human behavior. One example of integrating features across models comes from the extrapolation of tempo changes forward in time as introduced in timekeeper model ADAM [58]. This feature is crucial for predicting naturalistic rhythm, such as performed music that contains tempo shifts and microtiming. Probabilistic models of melodic expectations [85] and rhythm [11] borrowed this idea by making current event probabilities dependent on recent history. Similarly, an oscillator-based model introduced a dynamic tempo variable that modulates oscillator frequency [10], which led to better fits to behavioral data.

Timekeeper and probabilistic models might also borrow ideas from entrainment models. In one non-linear entrainment model, prediction arises from neural transmission delays inherent in the system (*strong anticipation*; see [72]). These delays reflect the system's earlier state, allowing a driven system to predict its driver [38]. This is reminiscent of finger taps preceding sounds during synchronization (i.e., *negative mean asynchrony*, though whether this reflects prediction is debated [71,82]). Characteristics of probabilistic models might be improved by incorporating these concepts and making current states dependent on preceding states via a fixed delay. Linking delayed neural coupling to concepts of adaptation and flexibility may also bridge between low-level brain dynamics and cognitive representations, our second aim in integrating models.

4.2.2. Integrating across different levels of explanation. If cognitive neuroscience aims to link neural processes to cognition and behavior [105], integrating between models at different levels of explanation is particularly important. For example, knowledge can be extrapolated from neurophysiological models to probabilistic predictions at behavioral and cognitive levels. A neural firing rate model that produces self-sustaining oscillations was able to reproduce human psychophysics data showing Bayes-like inference of expected stimulus timing [10]. Similarly, a neural circuit model captured human Bayesian behavior in timing and synchronization tasks [103]. More recently, probabilistic inference of beat phase was shown to be mathematically equivalent to an oscillator model [49]. When a dynamic beat phase estimate and its uncertainty are treated like phase and amplitude (respectively) of a complex variable, the result is an underdamped oscillator with a particular phase-and-amplitude-dependent response to forcing. These examples show the promise of integration between cognitive and neural levels of description.

The behavior-level description of timekeeper models and the neuro-mechanistic description of oscillator models could also potentially be integrated, as both can produce a series of event times showing error-correcting behavior [40,60]. However, certain well-established points of disagreement between the two model classes would have to be overcome, including timekeeper models' assumptions of linear error correction and one-to-one action-event coupling [40]. The latter is related to the focus of timekeeper models on isochronous sequences, whereas entrainment models also capture beat

and meter in non-isochronous rhythm [1,25]. This leads to our final aim in model integration: integrating across different aspects of rhythmic expectations.

4.2.3. Integrating across types of rhythmic expectation. Whether the cognitive mechanisms for rhythmic expectations depend on task demands [84] and rhythmic structure [80] is an open question. Integrating different types of rhythmic expectations (e.g., beat, pattern) into a unifying model forces us to explicitly define their interactions and task-dependence. Here, bringing the output of different models into the same continuous computational space, agnostic of the modeled structure, may be especially helpful. For example, the continuous probability distribution of PIPPET can incorporate both pattern-based and metrical expectations [52].

If two models with explanatory value cannot be integrated, we have identified irreconcilable ideas that must either be evaluated against each other or cautiously accepted to offer independent cognitive mechanisms. Model integration as such not only allows building better and more overarching models, but also furthers cognitive theory about what rhythmic expectations entail.

5. Summary and recommendations

The formal description of rhythmic expectations provided by computational models is crucial to understanding how temporal structure affords benefits in perception, movement, social bonding, and reward associated with musical rhythm [4,5]. Three main classes of computational models (entrainment, probabilistic, and timekeeper models) each have their own focus regarding the level of explanation, type of rhythmic structure, and type of human data they target. Models also deal with learning differently, and non-rhythmic features are rarely accounted for. Often, models of rhythmic expectations are designed to capture specific behavior in response to specific stimuli and work well within this narrow focus.

To reconcile the diverse aims and computational forms of models that arguably target the same underlying concept (i.e., rhythmic expectations), we propose that model comparison and integration are crucial. Model comparison will advance the field by selecting algorithmic descriptions that explain rhythmic behavior best and by forcing us to reduce model flexibility. Model integration can bridge cognitive and neuro-mechanistic explanations, revealing how apparently different rhythmic abilities may result from the same underlying principles. Recent work, such as the PIPPET model, demonstrates how expectations based on rhythmic patterns, beat, and meter can be captured within a unified framework. Alongside PIPPET, models such as the adaptive oscillator model [10] and the neural circuit model [103] show the potential to connect algorithmic accounts (e.g., Bayesian models), low-level neuro-mechanistic processes (e.g., neural oscillations), and high-level theories of auditory-motor coupling in the brain (e.g., the ASAP hypothesis). Such integrative attempts also make it possible to identify truly irreducible abilities, answering pertinent questions about whether multiple mechanisms underlie rhythmic expectations [79,80].

To achieve these goals, we extract several practical recommendations to guide future modeling efforts:

1. *Models to be compared or integrated should handle similar input and produce similar output.* Ideally, models of rhythmic expectations operate on continuous input, or raw audio. Possible conversions of continuous or discrete input to the symbolic domain should consider known cognitive mechanisms involved in rhythm processing, such as categorization. Likewise, model output needs to be brought into the same space, either by extracting a surprise or salience measure for every event in a sequence, or by mapping discrete or symbolic output onto continuous time signals.
2. *Multiple considerations must guide model comparison.* Model evaluation should consider not just GOF, but also model complexity, model scope, cognitive plausibility, unexpectedness of model predictions, and the capacity to capture both cognitive universals and individual differences. Experiments should be designed to maximally differentiate between, and potentially falsify, candidate models.
3. *Model integration must be used to inform theory.* Model integration can not only yield models of greater explanatory power, but also inform theory about irreducible mechanisms of rhythmic expectations.

4. *Open science*. To achieve comparison and integration, models should be publicly available in executable form with clear documentation. Datasets of behavioral and neural data pertaining to rhythmic expectations that may be targeted by models should be made more widely available in open repositories. Frameworks for standardizing and sharing datasets, such as the Cognition and Natural Sensory Processing Initiative (CNSP; <https://cnspworkshop.net/>) and the Timing Database [106], are essential for promoting reproducibility and integration across research groups.

6. Emerging opportunities

Future modeling efforts may broaden the scope of current models by including a wider range of relevant input, like spectral, intensity, and melodic information, and non-auditory information, from visual and tactile modalities. Such efforts may be aided by the availability of rich, multimodal datasets of music and musical behavior [107]. Moreover, while most modeling efforts focus on humans, non-human animals are known to exhibit various aspects of rhythm processing [82]. Cross-species comparisons using the models discussed here may be an especially fruitful way to uncover the mechanisms behind between-species differences in rhythm processing, like differences in beat perception between non-human primates and humans [12,13]. Although entrainment models have been applied to study synchrony in the behavior of non-human animals, such as chorusing in insects, frogs, and birds [108], modeling has rarely been used to directly compare rhythmic behaviors across humans and non-human animals. Notably, one implementation of IDyOM was shown to predict EEG responses to rhythm in rhesus monkeys, similar to humans [67], suggesting that, to some extent, similar processes may underlie rhythmic processing across these species. Extending such modeling approaches to uncover universals and differences across cultures, species, and developmental stages may contribute to the ongoing debate about the evolutionary origins of rhythm processing [5,12], which so far has seen a limited contribution from modeling studies.

Another necessary avenue is to consider deep neural networks [109] and generative AI models [18] for understanding rhythm cognition. Whereas such data-driven models are often designed to optimize metrics of musical quality [110] rather than to understand cognition, there is large merit in using AI to inform theory formation in cognitive psychology [111]. Deep learning models that are trained on large-scale music datasets may implicitly capture statistical regularities, serving as a proxy for human enculturation. Their temporal predictions could be compared with probabilistic models explicitly designed to capture enculturation, such as IDyOM, as well as with human behavioral and neural data [93]. Similarly, the beat synchronization abilities of generative music models [112], models trained on coupled datasets of music and bodily motion such as UniMuMo [113], and deep reinforcement learning models [114], could be compared to human entrainment data. Recent benchmark efforts such as MMAR [115] and MMAU [116] provide large-scale evaluations of audio and music reasoning in multimodal Large Language Models, whereas the MUSE benchmark [117] goes further by directly comparing model performance on tasks like syncopation and meter identification with human perceptual baselines. These comparisons could shed light on whether rhythmic expectations emerge from general-purpose predictive learning or instead require the explicit cognitive mechanisms of probabilistic, entrainment, and timekeeper models. Still, while AI models can serve as theoretical tools, we should be cautious not to assume that human-like performance implies a faithful model of human cognition [111]. Indeed, large AI models often fall short of a key hallmark of cognitive models: interpretability [96].

Author contributions

Conceptualization: Atser Damsma, Thomas Kaplan, Fleur L. Bouwer.

Supervision: Fleur L. Bouwer.

Writing – original draft: Atser Damsma, Jonathan Cannon, Lauren K. Fink, Keith B. Doelling, Jessica A. Grahn, Henkjan Honing, Thomas Kaplan, Edward W. Large, Fleur L. Bouwer.

Writing – review & editing: Atser Damsma, Jonathan Cannon, Lauren K. Fink, Keith B. Doelling, Jessica A. Grahn, Henkjan Honing, Thomas Kaplan, Edward W. Large, Fleur L. Bouwer.

References

1. Large EW, Roman I, Kim JC, Cannon J, Pazdera JK, Trainor LJ, et al. Dynamic models for musical rhythm perception and coordination. *Front Comput Neurosci*. 2023;17:1151895. <https://doi.org/10.3389/fncom.2023.1151895> PMID: [37265781](https://pubmed.ncbi.nlm.nih.gov/37265781/)
2. Large EW, Jones MR. The dynamics of attending: how people track time-varying events. *Psychol Rev*. 1999;106(1):119–59. <https://doi.org/10.1037/0033-295x.106.1.119>
3. Nobre AC, van Ede F. Anticipated moments: temporal structure in attention. *Nat Rev Neurosci*. 2018;19(1):34–48. <https://doi.org/10.1038/nrn.2017.141> PMID: [29213134](https://pubmed.ncbi.nlm.nih.gov/29213134/)
4. Fiveash A, Ferreri L, Bouwer FL, Kösem A, Moghimi S, Ravnani A, et al. Can rhythm-mediated reward boost learning, memory, and social connection? Perspectives for future research. *Neurosci Biobehav Rev*. 2023;149:105153. <https://doi.org/10.1016/j.neubiorev.2023.105153> PMID: [37019245](https://pubmed.ncbi.nlm.nih.gov/37019245/)
5. Savage PE, Loui P, Tarr B, Schachner A, Glowacki L, Mithen S, et al. Music as a coevolved system for social bonding. *Behav Brain Sci*. 2020;44:e59. <https://doi.org/10.1017/S0140525X20000333> PMID: [32814608](https://pubmed.ncbi.nlm.nih.gov/32814608/)
6. Damm L, Varoqui D, De Cock VC, Dalla Bella S, Bardy B. Why do we move to the beat? A multi-scale approach, from physical principles to brain dynamics. *Neurosci Biobehav Rev*. 2020;112:553–84. <https://doi.org/10.1016/j.neubiorev.2019.12.024> PMID: [31846652](https://pubmed.ncbi.nlm.nih.gov/31846652/)
7. Guest O, Martin AE. How computational modeling can force theory building in psychological science. *Perspect Psychol Sci*. 2021;16(4):789–802. <https://doi.org/10.1177/1745691620970585> PMID: [33482070](https://pubmed.ncbi.nlm.nih.gov/33482070/)
8. Fink LK. Computational models of temporal expectations. In: *Future directions of music cognition*; 2021. <https://doi.org/10.18061/fdmc.2021.0041>
9. Honing H. Computational modeling of music cognition: a case study on model selection. *Music Percept*. 2006;23(5):365–76. <https://doi.org/10.1525/mp.2006.23.5.365>
10. Doelling KB, Arnal LH, Assaneo MF. Adaptive oscillators support Bayesian prediction in temporal processing. *PLoS Comput Biol*. 2023;19(11):e1011669. <https://doi.org/10.1371/journal.pcbi.1011669> PMID: [38011225](https://pubmed.ncbi.nlm.nih.gov/38011225/)
11. Bouwer FL, Damsma A, Kaplan TM, Sarvestani MG, Pearce MT. Abstract representations underlie rhythm perception and production: Evidence from a probabilistic model of temporal structure. *Cognition*. 2025;268:106345. <https://doi.org/10.1016/j.cognition.2025.106345> PMID: [41240806](https://pubmed.ncbi.nlm.nih.gov/41240806/)
12. Patel AD. Beat-based dancing to music has evolutionary foundations in advanced vocal learning. *BMC Neurosci*. 2024;25(1):65. <https://doi.org/10.1186/s12868-024-00843-6> PMID: [39506663](https://pubmed.ncbi.nlm.nih.gov/39506663/)
13. Merchant H, Honing H. Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Front Neurosci*. 2014;7:274. <https://doi.org/10.3389/fnins.2013.00274> PMID: [24478618](https://pubmed.ncbi.nlm.nih.gov/24478618/)
14. Lee MD, Webb MR. Modeling individual differences in cognition. *Psychon Bull Rev*. 2005;12(4):605–21. <https://doi.org/10.3758/bf03196751> PMID: [16447375](https://pubmed.ncbi.nlm.nih.gov/16447375/)
15. Demos AP, Palmer C. Musical synchrony, dynamical systems and information processing: merger or redundancy?. *Trends Cogn Sci*. 2023;27(12):1107–8. <https://doi.org/10.1016/j.tics.2023.08.015> PMID: [37739922](https://pubmed.ncbi.nlm.nih.gov/37739922/)
16. Keller PE. Integrating theory and models of musical group interaction. *Trends Cogn Sci*. 2023;27(12):1105–6. <https://doi.org/10.1016/j.tics.2023.07.008> PMID: [37739920](https://pubmed.ncbi.nlm.nih.gov/37739920/)
17. Povel D-J, Essens P. Perception of temporal patterns. *Music Percept* 1985;2(4):411–40. <https://doi.org/10.2307/40285311>
18. Agostinelli A, Denk TI, Borsos Z, Engel J, Verzetti M, Caillon A. MusicLM: generating music from text. 2023. Available from: <http://arxiv.org/abs/2301.11325>
19. Fiveash A, Bedoin N, Gordon RL, Tillmann B. Processing rhythm in speech and music: shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology*. 2021;35(8):771–91. <https://doi.org/10.1037/neu0000766> PMID: [34435803](https://pubmed.ncbi.nlm.nih.gov/34435803/)
20. Ten Oever S, Martin AE. An oscillating computational model can track pseudo-rhythmic speech by using linguistic predictions. *Elife*. 2021;10:e68066. <https://doi.org/10.7554/eLife.68066> PMID: [34338196](https://pubmed.ncbi.nlm.nih.gov/34338196/)
21. Fink LK, Hurley BK, Geng JJ, Janata P. A linear oscillator model predicts dynamic temporal attention and pupillary entrainment to rhythmic patterns. *J Eye Mov Res*. 2018;11(2):10.16910/jemr.11.2.12. <https://doi.org/10.16910/jemr.11.2.12> PMID: [33828695](https://pubmed.ncbi.nlm.nih.gov/33828695/)
22. Hurley BK, Fink LK, Janata P. Mapping the dynamic allocation of temporal attention in musical patterns. *J Exp Psychol Hum Percept Perform*. 2018;44(11):1694–711. <https://doi.org/10.1037/xhp0000563> PMID: [30091636](https://pubmed.ncbi.nlm.nih.gov/30091636/)
23. Janata P, Tomic ST, Haberman JM. Sensorimotor coupling in music and the psychology of the groove. *J Exp Psychol Gen*. 2012;141(1):54–75. <https://doi.org/10.1037/a0024208> PMID: [21767048](https://pubmed.ncbi.nlm.nih.gov/21767048/)
24. Tomic ST, Janata P. Beyond the beat: modeling metric structure in music and performance. *J Acoust Soc Am*. 2008;124(6):4024–41. <https://doi.org/10.1121/1.3006382> PMID: [19206825](https://pubmed.ncbi.nlm.nih.gov/19206825/)
25. Large EW, Herrera JA, Velasco MJ. Neural networks for beat perception in musical rhythm. *Front Syst Neurosci*. 2015;9:159. <https://doi.org/10.3389/fnsys.2015.00159> PMID: [26635549](https://pubmed.ncbi.nlm.nih.gov/26635549/)
26. Tichko P, Large EW. Modeling infants' perceptual narrowing to musical rhythms: neural oscillation and Hebbian plasticity. *Ann N Y Acad Sci*. 2019;1453(1):125–39. <https://doi.org/10.1111/nyas.14050> PMID: [31021447](https://pubmed.ncbi.nlm.nih.gov/31021447/)

27. Tichko P, Kim JC, Large EW. Bouncing the network: a dynamical systems model of auditory-vestibular interactions underlying infants' perception of musical rhythm. *Dev Sci*. 2021;24(5):e13103. <https://doi.org/10.1111/desc.13103> PMID: 33570778
28. Zalta A, Large EW, Schön D, Morillon B. Neural dynamics of predictive timing and motor engagement in music listening. *Sci Adv*. 2024;10(10):eadi2525. <https://doi.org/10.1126/sciadv.adi2525> PMID: 38446888
29. Dotov D, Delasanta L, Cameron DJ, Large EW, Trainor L. Collective dynamics support group drumming, reduce variability, and stabilize tempo drift. *Elife*. 2022;11:e74816. <https://doi.org/10.7554/eLife.74816> PMID: 36317963
30. Heggli OA, Cabral J, Konvalinka I, Vuust P, Kringelbach ML. A Kuramoto model of self-other integration across interpersonal synchronization strategies. *PLoS Comput Biol*. 2019;15(10):e1007422. <https://doi.org/10.1371/journal.pcbi.1007422> PMID: 31618261
31. Roman IR, Roman AS, Kim JC, Large EW. Hebbian learning with elasticity explains how the spontaneous motor tempo affects music performance synchronization. *PLoS Comput Biol*. 2023;19(6):e1011154. <https://doi.org/10.1371/journal.pcbi.1011154> PMID: 37285380
32. Large EW. On synchronizing movements to music. *Hum Mov Sci*. 2000;19(4):527–66. [https://doi.org/10.1016/s0167-9457\(00\)00026-9](https://doi.org/10.1016/s0167-9457(00)00026-9)
33. Large EW, Palmer C. Perceiving temporal regularity in music. *Cogn Sci*. 2002;26(1):1–37. https://doi.org/10.1207/s15516709cog2601_1
34. Zalta A, Petkoski S, Morillon B. Natural rhythms of periodic temporal attention. *Nat Commun*. 2020;11(1):1051. <https://doi.org/10.1038/s41467-020-14888-8> PMID: 32103014
35. Large EW, Fink P, Kelso JAS. Tracking simple and complex sequences. *Psychol Res*. 2002;66(1):3–17. <https://doi.org/10.1007/s004260100069> PMID: 11963276
36. Doelling KB, Assaneo MF, Bevilacqua D, Pesaran B, Poeppel D. An oscillator model better predicts cortical entrainment to music. *Proc Natl Acad Sci U S A*. 2019;116(20):10113–21. <https://doi.org/10.1073/pnas.1816414116> PMID: 31019082
37. Dotov DG, Cochen de Cock V, Geny C, Ihalainen P, Moens B, Leman M, et al. The role of interaction and predictability in the spontaneous entrainment of movement. *J Exp Psychol Gen*. 2019;148(6):1041–57. <https://doi.org/10.1037/xge0000609> PMID: 31180717
38. Roman IR, Washburn A, Large EW, Chafe C, Fujioka T. Delayed feedback embedded in perception-action coordination cycles results in anticipation behavior during synchronized rhythmic action: a dynamical systems approach. *PLoS Comput Biol*. 2019;15: e1007371. <https://doi.org/10.1371/journal.pcbi.1007371>
39. Scheurich R, Pfordresher PQ, Palmer C. Musical training enhances temporal adaptation of auditory-motor synchronization. *Exp Brain Res*. 2020;238(1):81–92. <https://doi.org/10.1007/s00221-019-05692-y> PMID: 31792555
40. Loehr JD, Large EW, Palmer C. Temporal coordination and adaptation to rate change in music performance. *J Exp Psychol Hum Percept Perform*. 2011;37(4):1292–309. <https://doi.org/10.1037/a0023102> PMID: 21553990
41. van der Weij B, Pearce MT, Honing H. A probabilistic model of meter perception: simulating enculturation. *Front Psychol*. 2017;8:824. <https://doi.org/10.3389/fpsyg.2017.00824> PMID: 28588533
42. Pearce MT. Statistical learning and probabilistic prediction in music cognition: Mechanisms of stylistic enculturation. *Ann N Y Acad Sci*. 2018;1423: 378–95. <https://doi.org/10.1111/nyas.13654>
43. Milne AJ, Dean RT, Bulger D. The effects of rhythmic structure on tapping accuracy. *Atten Percept Psychophys*. 2023;85(8):2673–99. <https://doi.org/10.3758/s13414-023-02778-2> PMID: 37817052
44. Sauv e SA, Pearce MT. Information-theoretic modeling of perceived musical complexity. *Music Percept*. 2019;37(2):165–78. <https://doi.org/10.1525/mp.2019.37.2.165>
45. Skeritt-Davis B, Elhilali M. Computational framework for investigating predictive processing in auditory perception. *J Neurosci Methods*. 2021;360:109177. <https://doi.org/10.1016/j.jneumeth.2021.109177> PMID: 33839191
46. Sadakata M, Desain P, Honing H. The bayesian way to relate rhythm perception and production. *Music Percept*. 2006;23(3):269–88. <https://doi.org/10.1525/mp.2006.23.3.269>
47. Senn O. A predictive coding approach to modelling the perceived complexity of popular music drum patterns. *Heliyon*. 2023;9(4):e15199. <https://doi.org/10.1016/j.heliyon.2023.e15199> PMID: 37123947
48. Cannon J. Expectancy-based rhythmic entrainment as continuous Bayesian inference. *PLoS Comput Biol*. 2021;17(6):e1009025. <https://doi.org/10.1371/journal.pcbi.1009025> PMID: 34106918
49. Cannon J, Kaplan T. Inferred representations behave like oscillators in dynamic Bayesian models of beat perception. *J Math Psychol*. 2024;122:102869. <https://doi.org/10.1016/j.jmp.2024.102869>
50. Kaplan T, Jamone L, Pearce M. Probabilistic modelling of microtiming perception. *Cognition*. 2023;239:105532. <https://doi.org/10.1016/j.cognition.2023.105532> PMID: 37442021
51. Kaplan T, Cannon J, Jamone L, Pearce M. Modeling enculturated bias in entrainment to rhythmic patterns. *PLoS Comput Biol*. 2022;18(9):e1010579. <https://doi.org/10.1371/journal.pcbi.1010579> PMID: 36174063
52. Kaplan T. Probabilistic models of rhythmic expectation & synchronisation. Queen Mary University of London; 2024. Available from: <https://qmro.qmul.ac.uk/xmlui/handle/123456789/94727>
53. Repp BH, Keller PE. Sensorimotor synchronization with adaptively timed sequences. *Hum Mov Sci*. 2008;27(3):423–56. <https://doi.org/10.1016/j.humov.2008.02.016> PMID: 18405989

54. Vishne G, Jacoby N, Malinovitch T, Epstein T, Frenkel O, Ahissar M. Slow update of internal representations impedes synchronization in autism. *Nat Commun*. 2021;12(1):5439. <https://doi.org/10.1038/s41467-021-25740-y> PMID: [34521851](#)
55. Jacoby N, Tishby N, Repp BH, Ahissar M, Keller PE. Parameter estimation of linear sensorimotor synchronization models: phase correction, period correction, and ensemble synchronization. *Timing Time Percept*. 2015;3(1–2):52–87. <https://doi.org/10.1163/22134468-00002048>
56. Harry B, Keller PE. Tutorial and simulations with ADAM: an adaptation and anticipation model of sensorimotor synchronization. *Biol Cybern*. 2019;113(4):397–421. <https://doi.org/10.1007/s00422-019-00798-6> PMID: [30963226](#)
57. Harry BB, Margulies DS, Falkiewicz M, Keller PE. Brain networks for temporal adaptation, anticipation, and sensory-motor integration in rhythmic human behavior. *Neuropsychologia*. 2023;183:108524. <https://doi.org/10.1016/j.neuropsychologia.2023.108524> PMID: [36868500](#)
58. van der Steen MCM, Keller PE. The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Front Hum Neurosci*. 2013;7:253. <https://doi.org/10.3389/fnhum.2013.00253> PMID: [23772211](#)
59. van der Steen MCM, Jacoby N, Fairhurst MT, Keller PE. Sensorimotor synchronization with tempo-changing auditory sequences: Modeling temporal adaptation and anticipation. *Brain Res*. 2015;1626:66–87. <https://doi.org/10.1016/j.brainres.2015.01.053> PMID: [25725379](#)
60. Bose A, Byrne Á, Rinzel J. A neuromechanistic model for rhythmic beat generation. *PLoS Comput Biol*. 2019;15(5):e1006450. <https://doi.org/10.1371/journal.pcbi.1006450> PMID: [31071078](#)
61. Obleser J, Kayser C. Neural entrainment and attentional selection in the listening brain. *Trends Cogn Sci*. 2019;23(11):913–26. <https://doi.org/10.1016/j.tics.2019.08.004> PMID: [31606386](#)
62. Doelling KB, Assaneo MF. Neural oscillations are a start toward understanding brain activity rather than the end. *PLoS Biol*. 2021;19(5):e3001234. <https://doi.org/10.1371/journal.pbio.3001234> PMID: [33945528](#)
63. Tichko P, Kim JC, Large EW. A dynamical, radically embodied, and ecological theory of rhythm development. *Front Psychol*. 2022;13:653696. <https://doi.org/10.3389/fpsyg.2022.653696> PMID: [35282203](#)
64. Friston K. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci*. 2005;360(1456):815–36. <https://doi.org/10.1098/rstb.2005.1622> PMID: [15937014](#)
65. Vuust P, Heggli OA, Friston KJ, Kringelbach ML. Music in the brain. *Nat Rev Neurosci*. 2022;23(5):287–305. <https://doi.org/10.1038/s41583-022-00578-5> PMID: [35352057](#)
66. Marr D. *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Walter Freeman and Co.; 1982.
67. Bianco R, Zuk NJ, Bigand F, Quarta E, Grasso S, Arnese F, et al. Neural encoding of musical expectations in a non-human primate. *Curr Biol*. 2024;34(2):444–450.e5. <https://doi.org/10.1016/j.cub.2023.12.019> PMID: [38176416](#)
68. Di Liberto GM, Pelofi C, Bianco R, Patel P, Mehta AD, Herrero JL, et al. Cortical encoding of melodic expectations in human temporal cortex. *Elife*. 2020;9:e51784. <https://doi.org/10.7554/eLife.51784> PMID: [32122465](#)
69. Auksztulewicz R, Schwiedrzik CM, Thesen T, Doyle W, Devinsky O, Nobre AC, et al. Not all predictions are equal: “What” and “When” predictions modulate activity in auditory cortex through different mechanisms. *J Neurosci*. 2018;38(40):8680–93. <https://doi.org/10.1523/JNEUROSCI.0369-18.2018> PMID: [30143578](#)
70. Cannon JJ, Patel AD. How beat perception co-opts motor neurophysiology. *Trends Cogn Sci*. 2021;25(2):137–50. <https://doi.org/10.1016/j.tics.2020.11.002> PMID: [33353800](#)
71. Repp BH, Su Y-H. Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon Bull Rev*. 2013;20(3):403–52. <https://doi.org/10.3758/s13423-012-0371-2> PMID: [23397235](#)
72. Stepp N, Turvey MT. On strong anticipation. *Cogn Syst Res*. 2010;11(2):148–64. <https://doi.org/10.1016/j.cogsys.2009.03.003> PMID: [20191086](#)
73. Novembre G, Iannetti GD. Tagging the musical beat: neural entrainment or event-related potentials?. *Proc Natl Acad Sci U S A*. 2018;115(47):E11002–3. <https://doi.org/10.1073/pnas.1815311115> PMID: [30425178](#)
74. Damsma A, de Roo M, Doelling K, Bazin P-L, Bouwer FL. Tempo-dependent selective enhancement of neural responses at the beat frequency can be mimicked by both an oscillator and an evoked model. *Cereb Cortex*. 2025;35(9):bhaf258. <https://doi.org/10.1093/cercor/bhaf258> PMID: [40996143](#)
75. Hardy NF, Buonomano DV. Neurocomputational models of interval and pattern timing. *Curr Opin Behav Sci*. 2016;8:250–7. <https://doi.org/10.1016/j.cobeha.2016.01.012> PMID: [27790629](#)
76. Konvalinka I, Vuust P, Roepstorff A, Frith CD. Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Q J Exp Psychol (Hove)*. 2010;63(11):2220–30. <https://doi.org/10.1080/17470218.2010.497843> PMID: [20694920](#)
77. Fink LK, Alexander PC, Janata P. The Groove Enhancement Machine (GEM): a multi-person adaptive metronome to manipulate sensorimotor synchronization and subjective enjoyment. *Front Hum Neurosci*. 2022;16:916551. <https://doi.org/10.3389/fnhum.2022.916551> PMID: [35782041](#)
78. Desain P, Honing H. The formation of rhythmic categories and metric priming. *Perception*. 2003;32(3):341–65. <https://doi.org/10.1068/p3370> PMID: [12729384](#)
79. Fiveash A, Bella SD, Bigand E, Gordon RL, Tillmann B. You got rhythm, or more: the multidimensionality of rhythmic abilities. *Atten Percept Psychophys*. 2022;84(4):1370–92. <https://doi.org/10.3758/s13414-022-02487-2> PMID: [35437703](#)

80. Bouwer FL, Fahrenfort JJ, Millard SK, Kloosterman NA, Slagter HA. A Silent Disco: differential effects of beat-based and pattern-based temporal expectations on persistent entrainment of low-frequency neural oscillations. *J Cogn Neurosci*. 2023;35(6):990–1020. https://doi.org/10.1162/jocn_a_01985 PMID: 36951583
81. Rimmele JM, Morillon B, Poeppel D, Arnal LH. Proactive sensing of periodic and aperiodic auditory patterns. *Trends Cogn Sci*. 2018;22(10):870–82. <https://doi.org/10.1016/j.tics.2018.08.003> PMID: 30266147
82. Bouwer FL, Nityananda V, Rouse AA, Ten Cate C. Rhythmic abilities in humans and non-human animals: a review and recommendations from a methodological perspective. *Philos Trans R Soc Lond B Biol Sci*. 2021;376(1835):20200335. <https://doi.org/10.1098/rstb.2020.0335> PMID: 34420380
83. Cannon J. Marching to your own beat. *Music Percept*. 2025;43(1):49–60. <https://doi.org/10.1525/mp.2025.2427447>
84. Shalev N, Nobre AC, van Ede F. Time for what? Breaking down temporal anticipation. *Trends Neurosci*. 2019;42(6):373–4. <https://doi.org/10.1016/j.tins.2019.03.002> PMID: 30885405
85. Harrison PMC, Bianco R, Chait M, Pearce MT. PPM-Decay: a computational model of auditory prediction with memory decay. *PLoS Comput Biol*. 2020;16(11):e1008304. <https://doi.org/10.1371/journal.pcbi.1008304> PMID: 33147209
86. Jacoby N, Polak R, Grahn JA, Cameron DJ, Lee KM, Godoy R, et al. Commonality and variation in mental representations of music revealed by a cross-cultural comparison of rhythm priors in 15 countries. *Nat Hum Behav*. 2024;8(5):846–77. <https://doi.org/10.1038/s41562-023-01800-9> PMID: 38438653
87. Weineck K, Wen OX, Henry MJ. Neural synchronization is strongest to the spectral flux of slow music and depends on familiarity and beat salience. *Elife*. 2022;11:e75515. <https://doi.org/10.7554/eLife.75515> PMID: 36094165
88. Pazdera JK, Trainor LJ. Pitch-induced illusory percepts of time. *Atten Percept Psychophys*. 2025;87(2):545–64. <https://doi.org/10.3758/s13414-024-02982-8> PMID: 39658731
89. Hove MJ, Marie C, Bruce IC, Trainor LJ. Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proc Natl Acad Sci U S A*. 2014;111(28):10383–8. <https://doi.org/10.1073/pnas.1402039111> PMID: 24982142
90. Manning F, Schutz M. “Moving to the beat” improves timing perception. *Psychon Bull Rev*. 2013;20(6):1133–9. <https://doi.org/10.3758/s13423-013-0439-7> PMID: 23670284
91. Varghese L, Mathias SR, Bensussen S, Chou K, Goldberg HR, Sun Y, et al. Bi-directional audiovisual influences on temporal modulation discrimination. *J Acoust Soc Am*. 2017;141(4):2474. <https://doi.org/10.1121/1.4979470> PMID: 28464677
92. Jacoby N, McDermott JH. Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Curr Biol*. 2017;27(3):359–70. <https://doi.org/10.1016/j.cub.2016.12.031> PMID: 28065607
93. Kern P, Heilbron M, de Lange FP, Spaak E. Cortical activity during naturalistic music listening reflects short-range predictions based on long-term experience. *Elife*. 2022;11:e80935. <https://doi.org/10.7554/eLife.80935> PMID: 36562532
94. Pitt MA, Myung IJ. When a good fit can be bad. *Trends Cogn Sci*. 2002;6(10):421–5. [https://doi.org/10.1016/s1364-6613\(02\)01964-2](https://doi.org/10.1016/s1364-6613(02)01964-2) PMID: 12413575
95. Pitt MA, Myung IJ, Zhang S. Toward a method of selecting among computational models of cognition. *Psychol Rev*. 2002;109(3):472–91. <https://doi.org/10.1037/0033-295x.109.3.472> PMID: 12088241
96. Wilson RC, Collins AG. Ten simple rules for the computational modeling of behavioral data. *Elife*. 2019;8:e49547. <https://doi.org/10.7554/eLife.49547> PMID: 31769410
97. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian model selection for group studies. *Neuroimage*. 2009;46(4):1004–17. <https://doi.org/10.1016/j.neuroimage.2009.03.025> PMID: 19306932
98. Rigoux L, Stephan KE, Friston KJ, Daunizeau J. Bayesian model selection for group studies - revisited. *Neuroimage*. 2014;84:971–85. <https://doi.org/10.1016/j.neuroimage.2013.08.065> PMID: 24018303
99. Piray P, Dezfouli A, Heskes T, Frank MJ, Daw ND. Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies. *PLoS Comput Biol*. 2019;15(6):e1007043. <https://doi.org/10.1371/journal.pcbi.1007043> PMID: 31211783
100. Palminteri S, Wyart V, Koehlin E. The importance of falsification in computational cognitive modeling. *Trends Cogn Sci*. 2017;21(6):425–33. <https://doi.org/10.1016/j.tics.2017.03.011> PMID: 28476348
101. Bulger D, Milne AJ, Dean RT. A point-process model of tapping along to difficult rhythms. *J Math Psychol*. 2022;111:102724. <https://doi.org/10.1016/j.jmp.2022.102724>
102. Honing H. The final ritard: on music, motion, and kinematic models. *Comput Music J*. 2003;27(3):66–72. <https://doi.org/10.1162/014892603322482538>
103. Egger SW, Le NM, Jazayeri M. A neural circuit model for human sensorimotor timing. *Nat Commun*. 2020;11(1):3933. <https://doi.org/10.1038/s41467-020-16999-8> PMID: 32770038
104. Popper K. *Conjectures and refutations*. London: Routledge and Keagan Paul; 1963.
105. Krakauer JW, Ghazanfar AA, Gomez-Marin A, Maclver MA, Poeppel D. Neuroscience needs behavior: correcting a reductionist bias. *Neuron*. 2017;93(3):480–90. <https://doi.org/10.1016/j.neuron.2016.12.041> PMID: 28182904

106. Aydođan T, Karřılar H, Duyan YA, Akdođan B, Baccarani A, Brochard R, et al. The timing database: an open-access, live repository for interval timing studies. *Behav Res Methods*. 2024;56(1):290–300. <https://doi.org/10.3758/s13428-022-02050-9> PMID: 36595180
107. Gotham M, Bemman B, Vatolkin I. Towards an 'Everything Corpus': a framework and guidelines for the curation of more comprehensive multi-modal music data. *Trans Int Soc Music Inform Retrieval*. 2025;8(1):70–92. <https://doi.org/10.5334/tismir.228>
108. Greenfield MD, Aihara I, Amichay G, Anichini M, Nityananda V. Rhythm interaction in animal groups: selective attention in communication networks. *Philos Trans R Soc Lond B Biol Sci*. 2021;376(1835):20200338. <https://doi.org/10.1098/rstb.2020.0338> PMID: 34420386
109. Davies M, Böck S, Fuentes M. Tempo, beat and downbeat estimation. 2021. [cited 2023 Oct 1]. <https://tempobeatdownbeat.github.io/tutorial/intro.html>
110. Lerch A, Arthur C, Bryan-Kinns N, Ford C, Sun Q, Vinay A. Survey on the evaluation of generative models in music. *ACM Comput Surv*. 2025;58(4):1–36. <https://doi.org/10.1145/3769106>
111. van Rooij I, Guest O, Adolfi F, de Haan R, Kolokolova A, Rich P. Reclaiming AI as a theoretical tool for cognitive science. *Comput Brain Behav*. 2024;7(4):616–36. <https://doi.org/10.1007/s42113-024-00217-5>
112. Copet J, Kreuk F, Gat I, Remez T, Kant D, Synnaeve G. Simple and controllable music generation. *Adv Neural Inform Process Syst*. 2023;36:47704–20.
113. Yang H, Su K, Zhang Y, Chen J, Qian K, Liu G, et al. UniMuMo: unified text, music, and motion generation. *AAAI*. 2025;39(24):25615–23. <https://doi.org/10.1609/aaai.v39i24.34752>
114. Ommi Y, Yousefjadi M, Cannon J. Learning sensorimotor synchronization through reinforcement in neural networks. *SSRN*; 2025. <https://doi.org/10.2139/ssrn.5233815>
115. Ma Z, Ma Y, Zhu Y, Yang C, Chao YW, Xu R. MMAR: a challenging benchmark for deep reasoning in speech, audio, music, and their mix; 2025. <https://doi.org/10.48550/arXiv.2505.13032>
116. Sakshi S, Tyagi U, Kumar S, Seth A, Selvakumar R, Nieto O. MMAU: a massive multi-task audio understanding and reasoning benchmark; 2024. <https://doi.org/10.48550/arXiv.2410.19168>
117. Carone BJ, Roman IR, Ripollés P. The MUSE benchmark: probing music perception and auditory relational reasoning in audio LLMs; 2025. <https://doi.org/10.48550/arXiv.2510.19055>