Journal Articles

Department of Biomechanics

8-23-2021

Cross-frequency coupling explains the preference for simple ratios in rhythmic behaviour and the relative stability across non-synchronous patterns

Dobromir Dotov

Laurel Trainor

Cross-frequency coupling explains the preference for simple ratios in rhythmic behaviour and the relative stability across non-synchronous patterns

Dobromir Dotov[1,2] and Laurel J. Trainor[1,2,3]

[1]LIVELab, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S4K1
[2]Psychology, Neuroscience and Behaviour, McMaster University, Ontario, Canada
[3]Rotman Research Institute, Toronto, Canada
DD, 0000-0002-5543-360X; LJT, 0000-0003-3397-2079

Rhythms are important for understanding coordinated behaviours in ecological systems. The repetitive nature of rhythms affords prediction, planning of movements and coordination of processes within and between individuals. A major challenge is to understand complex forms of coordination when they differ from complete synchronization. By expressing phase as ratio of a cycle, we adapted levels of the Farey tree as a metric of complexity mapped to the range between in-phase and anti-phase synchronization. In a bimanual tapping task, this revealed an increase of variability with ratio complexity, a range of hidden and unstable yet measurable modes, and a rank-frequency scaling law across these modes. We use the phase-attractive circle map to pro- pose an interpretation of these findings in terms of hierarchical cross-frequency coupling (CFC). We also consider the tendency for small-integer attractors in the single-hand repeated tapping of three-interval rhythms reported in the literature. The phase-attractive circle map has wider basins of attractions for such ratios. This work motivates the question whether CFC intrinsic to neural dynamics implements low-level priors for timing and coordination and thus becomes involved in phenomena as diverse as attractor states in bimanual coordination and the cross-cultural tendency for musical rhythms to have simple interval ratios.

This article is part of the theme issue 'Synchrony and rhythm interaction: from the brain to behavioural ecology'.

Author for correspondence:

Dobromir Dotov e-mail: dotovd@mcmaster.ca

Introduction

Rhythms are ubiquitous in biological systems, likely because their repetitive nature enables anticipation and prediction. A major challenge is how to co- ordinate between different rhythms both within an individual or between individuals. The ecological importance of coordination is revealed by the fact that either synchronous or asynchronous chorusing among conspecifics can result from the relative contribution of cooperative and competitive constraints [1,2]. In research, the range of coordination considered is restricted oftentimes to the extreme limits of complete synchrony (coincidence of periodic events) or a total lack of synchrony (random). Yet, there are also forms of interaction in chorusing animals consisting of simple ratios or even patterns that are non- isochronous yet coordinated [3]. Rhythms are also central to human action and human auditory-based communication systems, namely music and language. Within individuals, coordination of rhythms is also essential. In the action domain, it is necessary to coordinate rhythmic movements across effectors in order to locomote. Such behaviour has been studied extensively using bimanual coordination. Typically, only two points of stability are considered: in-phase (synchronous, 0° phase) and anti-phase (180° phase). The latter can be thought of as encompassing a 1 : 2 ratio between the cycle duration (tap to tap period within a hand) and the duration between all taps from both hands. We are interested in the grey area between these two stable modes because there have been fewer comprehensive attempts to quantify the stability of actions across a range of possible phases and identify possible attractor states.

Here we use a theoretically motivated measure of ratio (or phase) complexity with the hypothesis that complexity will be negatively associated with the stability of bimanual tap- ping beyond the classic two cases. We explore whether a cross-frequency coupled oscillator model can capture these complexity constraints. We then generalize the modelling approach by applying it to a different task: when asked to synchronize with one hand to a repeating

three-interval rhythm, people's tapping behaviour is attracted to rhythms where the three intervals stand in small integer relations [4].

In the auditory domain, the importance of small integer ratios has been recognized across diverse domains from the role of harmonic relations in pitch, consonance/dissonance and harmony perception, [5,6] to beat perception [7] and the prevalence around the world of rhythms with small-integer ratios [4]. Musical rhythmic behaviour in humans is a natural domain for observing refined complex forms of coordination and perception that are neither random nor coincidence synchrony. A large class of rhythms exhibits nesting, subdivision, and a discrete set of inter-beat intervals that are related as simple integer ratios, such as $1:3$, and are associated with perceptual and coordination benefits [8,9]. Speech also exhibits a rhythm-like hierarchical organization of temporally coordinated prosodic units [10], although with a higher variance or noise than in the typical musical case.

Despite the ubiquity of simple ratios in musical rhythms, a theoretically motivated metric of complexity has yet to be directly applied and a number of questions remain. For example, setting participants' tapping on various initial rhythmic patterns and then allowing their intervals to evolve freely over successive iterations has consistently revealed two contrasting results: there is a tendency towards simple integer ratios but it is not perfect and there appear to be additional factors that bias the attractive states in one way or another [11,12]. Taking this task to different human populations with different musical traditions has shown that not all simple ratios are represented equally and that, while there are common tendencies, culture-specific rhythmic priors can also be identified [4]. Additionally, developmental work shows that young infants easily encode rhythms containing various ratios, but by one year of age their perception has narrowed to become better at those ratios in their environment and worse at those ratios not in their environment [13–15].

We defined complexity as the level of the Farey tree on which a given ratio is found[1] (figure 1c). This is the definition used typically for the complexity of multifrequency coordination patterns. Multifrequency synchronization tasks such as tapping, limb oscillation, or gait-breathing coordination consistently show that low integer ratios are easier to learn and perform [16–19], and may even confer stronger neural entrainment as measured by electroencephalography, at least when tested over a very small set of rhythms [20]. A

minimal mathematical model of cross-frequency coupling (CFC) known as the circle map suggests that the benefit of simple ratios is that oscillators with such frequencies synchronize more easily [17,18]. What is novel here is that by expressing phase as a ratio we can apply the same ideas to a wider range of coordination tasks.

The Farey sequence of order *N* is the set of all fractions in lowest terms between 0 and 1 whose denominators do not exceed *N*. For example, the Farey sequence of order 3 is {0/1, 1/3, 1/2, 2/3, 1/1}. The Stern-Brocot tree is a mathematical graph in which the vertices correspond uniquely to all the positive rational numbers expressed in their lowest terms. Often referred to as Farey tree, the Stern-Brocot can be limited between 0 and 1 in which case it contains all the Farey sequences. It applies a rule to the fractions from one level to obtain the fractions of the next level. The tree-like procedure begins with 0 : 1 and 1 : 1 and each next level contains ratios of higher integers, figure 1*c*. It is easy to see how the notion of small and large integer ratios relates to multifrequency coordination and CFC. This notion can be related to phase too if phase is expressed as a proportion of the cycle and thus as a *p* : *q* ratio (figure 1*b*).

Here we address the theoretical hypothesis that the phase of bimanual coordination obeys the constraints of a covert CFC dynamic even if the behaviour is same-frequency. We conducted an experiment to quantify the stability of bimanual tapping, where each hand tapped once per cycle. The instructed phase relations between the hands, cued with auditory tones at the beginning of each trial, could be expressed as ratios of the tapping cycle and thus described as more or less complex. This metric makes predictions about the stability of a dense set of phases, not just in- and anti-phase. Farey trees often arise when dealing with CFC oscillator models. Interestingly, there is growing evidence for the role of cross-frequency hierarchical coupling in organizing neural dynamics [21]. Here we explored whether a CFC oscillator model could capture the phase complexity dynamics we measured behaviourally in our bimanual tapping task. We then extended this for previously reported data [4] on small integer ratio attractors where one hand is tapping a repeating three-interval rhythms.

Experiment 1: phase-synchronization- continuation tapping study
Method
Participants

Twelve participants (age range 18–21 years; 5 male, 7 female), undergraduate students at the local psychology department, completed the task in exchange for extra course credit. All participants self-reported being right-hand dominant or ambidextrous. All grew up listening to Western music and some had musical backgrounds from Asian or Middle- Eastern origin. The research was approved by the McMaster University Research Ethics Board (MREB#2164).

Apparatus

Tapping times were collected with a digital MIDI piano and a computer. The computer, running a custom patch in Pure Data (Pd), synthesized and played through a pair of closed- back circumaural headphones (Shure SRH440) two distinct pairs of sound samples: one pair for the stimulus and one pair for the participant's key presses. The stimulus consisted of two pulse-like drum sounds with sharp attack, 20 ms decay, and centre frequencies at 770 (left) and 990 (right) Hz. They were played separately through the left and right headphone channels. Participants' taps triggered synthetic- sounding pure tones with sharp attack and 50 ms decay time in lieu of traditional piano sounds. The left index finger tapped the C6 piano key (1046 Hz) and the right index finger the adjacent D6 (1174 Hz). The same Pd patch also recorded the timing of piano key presses as the participant tapped along to the stimulus.

Stimuli

Both the leading and secondary tones were played once a cycle, but separated by a phase-offset (figure 1*a,b*). For each trial, the phase relation between the secondary and leading cue was selected pseudo-randomly from a set of 12 phases spanning 0° (in-phase) to 180° (anti-phase). These phases corresponded to integer ratios taken from different levels of the Farey tree (see electronic supplementary material, table and figure 1*c*). The secondary cue was silenced during the continuation stage. The cycle period was set to 1000 ms (60 bpm) in the syn-chronization stage. During the continuation stage, the period decreased in steps of $0.1n\text{-}e^{.05n}$ ms (*n* was the cycle number), reaching 284 ms (210 bpm) at the end.
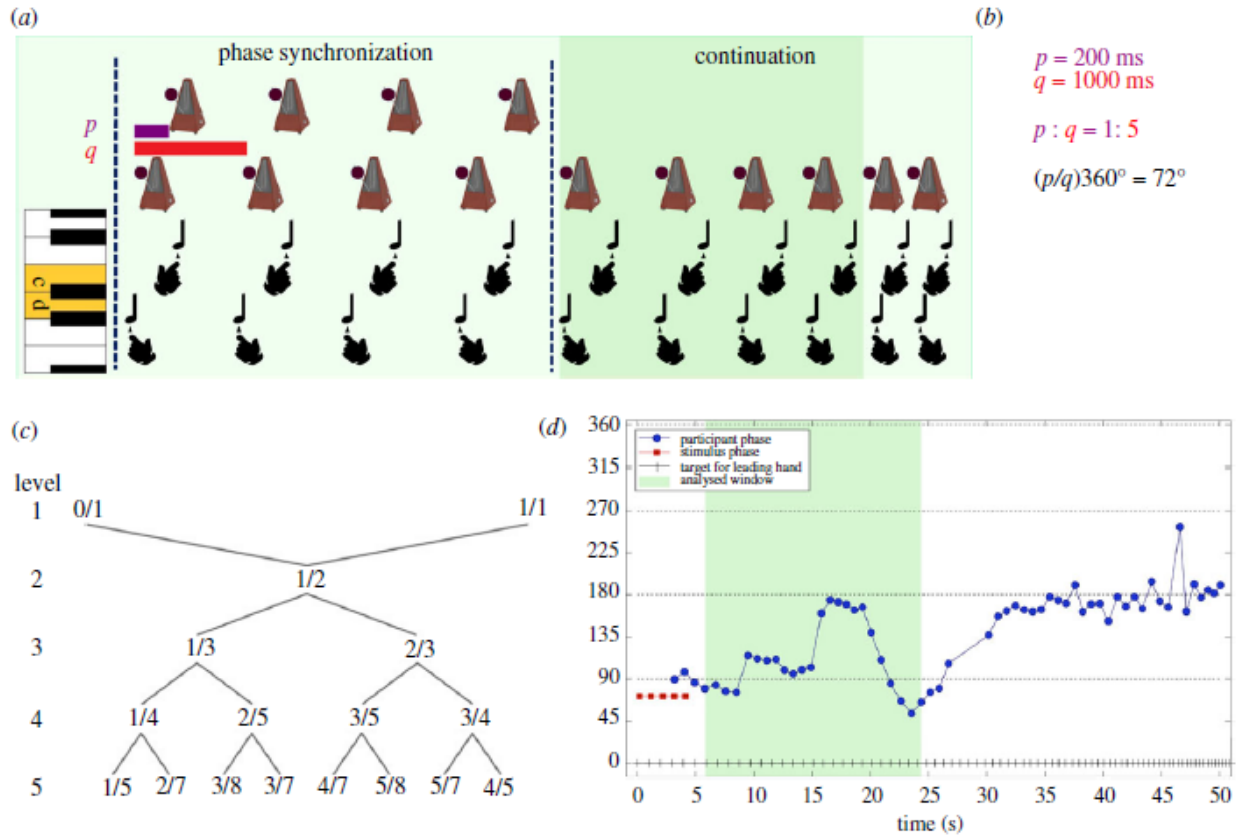
Figure 1. Bimanual phase-synchronization-continuation task (PSCT). (*a*) Participants produced a two-tap pattern with their index fingers on an electronic piano by synchronizing with an auditory stimulus. The stimulus in the synchronization stage consisted of left and right target cues played at the same period but separated by an offset (phase). The instruction was to synchronize the left and right taps with the left and right auditory cues, respectively. To fix the period and phase of tapping during the synchronization stage, the target cues were played on the leading (dominant, right) side and the following (secondary) cues on the non- dominant side. During the continuation stage, only dominant-side cues were played to fix the period, allowing participants to vary the phase between hands. The stimulus began speeding up during the continuation stage. (*b*) An example stimulus pattern expressed equivalently as time delays, *p* : *q* ratio, or phase. The *p* : *q* ratio is the time interval from the leading to the secondary cue divided by the interval between two leading cues. The ratio stands for phase normalized to unity. (*c*) Stimulus phase in each trial was taken pseudo-randomly from a set generated with a Farey tree. A subset up to level 5 is shown here for brevity; for a full list, see the electronic supplementary material, table. Our main hypothesis was that bimanual tapping performance is associated with the tree level of the instructed phase. (*d*) A sample trial. Phase is defined in terms of the secondary tap relative to the leading tap. We measured accuracy, variability, and entropy in the first 20 s of the continuation stage (shaded area). (Online version in colour.)

Task

The bimanual phase-synchronization-continuation task (PSCT) consisted of tapping in two stages: synchronization, where the tempo and phase were cued by the stimulus, and continuation, where only the tempo was cued (figure 1). During the synchronization stage, participants produced a succession of taps with the index finger of their dominant hand in synchrony

with the leading tones in the ipsilateral channel of their headphones and, at a given phase-offset, they produced a succession of taps with the index finger of their other hand in synchrony with the secondary tones played in the respective ipsilateral channel (in this study, right happened to be the dominant side in all cases). They continued this rhythmic tapping for the 50 s duration of the trial. After the initial synchronization stage (5 s), there was a continuation stage during which the secondary tones stopped but the participant continued to tap both parts (figure 1a). The leading cue pacing the dominant (right) side tap remained audible and slowly increased in tempo while the cue pacing the secondary tap was silenced. The instruction was to synchronize left and right taps with the corresponding left and right stimulus sounds during the synchronization stage of the trial and to try to maintain the pattern during the continuation stage while following the increase in tempo. Controlling the tempo during the continuation stage prevented participants from speeding up or slowing down freely to reach a more comfortable tapping rate which would have resulted in individual variation in the trade-off between tempo and accuracy. We included the gradual increase of tempo in the continuation stage to destabilize the tapping pattern. In retrospect, speeding up was unnecessary because participants found it difficult to maintain the phase even at the fixed tempo (figure 1d).

Procedure

Upon arrival, participants read and signed an informed con- sent sheet, performed a series of practice trials until both they and the experimenter felt confident that the task was well understood, and then performed 24 trials, or two for each stimulus phase pattern. After that, participants completed a demographic questionnaire.

Pre-processing and measures

The recorded tapping times were converted to a time series of phases by linear interpolation,

$$\theta_n = \frac{T_{\text{follow},n} - T_{\text{lead},m-1}}{T_{\text{lead},m} - T_{\text{lead},m-1}} 360,$$

where $T_{\text{follow},n}$ is a tap time of the secondary hand found between two successive taps of the leading hand, $T_{\text{lead},m-1}$ and $T_{\text{lead},m}$. (Separate indices m and n were needed because the left and right taps were not always necessarily aligned; on rare occasions the leading hand

could miss a tap. The corresponding phases were discarded.) Participants tended to exhibit large variability and converge to one of the two basic attractor phases as the instructed tempo kept increasing, namely in- or anti-phase (0° or 180°). For this reason, only the first 20 s of continuation were analysed (figure 1*d*).

As is frequently done when analysing phase, we applied circular statistics to characterize how well participants maintained the instructed phase in the continuation part of the trial. For a measure of *accuracy*, we took the difference between tapping phase and instructed phase, then the circular mean of that, then the absolute. For *variability*, we used two closely related measures, circular standard deviation and phase consistency, which is the mean resultant vector length of phases projected in the complex plane.

Entropy of phases was also computed as a measure alternative to the variance because there was a tendency for multimodal distributions to emerge even within the same trial. We used $H = \sum_{i=1}^{n} \rho(\varphi_i) \log \rho(\varphi_i)$ in the domain from 0° to 180° with a fixed set of bins. The tendency for multimodality was confirmed by using Gaussian mixture modelling (gmm from the mclust package for R) and the Bayesian information criterion to determine the best number of Gaussians. This was applied separately per participant and stimulus phase, see electronic supplementary material, figures S2 and S3. Furthermore, all phases from all trials and participants were pooled together to obtain a probability distribution and determine whether some phases were more attractive than others in the sense of being visited more frequently. Next, we fitted a Zipf-Mandelbrot scaling law $P = c(k + q)^{-s}$ to their rank-frequency distribution. This was motivated by the multiplicative branching character of the Farey tree and by the observation of such scaling in musical rhythms [22]. Pooling rhythmic intervals across various musical genres reveals the kind of scaling relation between their rank and frequency of occurrence that is usually associated with the distribution of words in languages [22]. The histogram was computed with a dense set of bins aligned to Farey tree ratios and then the coefficients were fitted against the log-transformed ranks and frequencies using nonlinear least-squares optimization. This procedure is available as part of the modelling repository, see §4a.
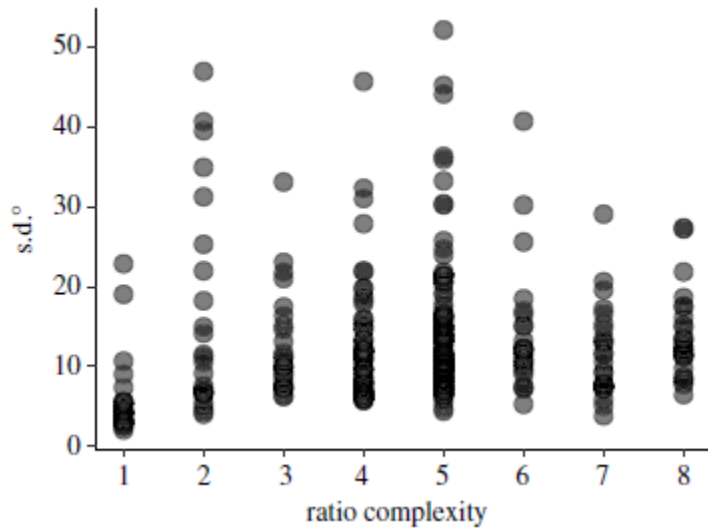
Figure 2. Variability in the PSCT. See electronic supplementary material, figure S1 for accuracy and consistency.

Statistical analysis

Separate linear mixed-effects models were fitted for each dependent variable, namely accuracy, consistency, variability and entropy of the produced phase (lme4 in R). The model-decision procedure involved a minimal specification including an intercept and a random effect for the intercept, accounting for individual overall level of performance, and incrementally including as predictors stimulus ratio complexity and phase. The coefficients of the significant fixed effects are reported in the text, with significance determined using the Satterthwaite method.

Results

Accuracy and variability

A linear increasing trend for accuracy (deviation) with respect to ratio complexity was observed, as expected, while the association with phase was more complicated (see electronic supplementary material, figure S1a,d). The linear model confirmed the effect of the continuous predictor ratio complexity level ($\beta = 2.439$, s.e. $= 0.989$, $t = 2.47$, $p < 0.05$), or an increase of 2.439 degrees per complexity level, while phase, also a continuous predictor, was not significant ($t < 1$). Phase consistency, after applying the logit transformation [23], exhibited a decreasing trend with complexity level ($\beta = -0.206$, s.e. $= 0.029$, $t = -7.22$, $p < 0.001$), as expected, and in addition to that with phase ($\beta = -2.962$, s.e. $= 0.385$, $t = -7.70$, $p < 0.001$; see electronic supplementary material, figure S1b,e). Standard deviation exhibited an increasing trend with

complexity level ($\beta = 0.753$, s.e. $= 0.219$, $t = 3.44$, $p < 0.001$), as expected, as well as with phase ($\beta = 17.132$, s.e. $= 2.948$, $t = 5.81$, $p < 0.0001$; figure 2 and electronic supplementary material, figure S1f). Entropy exhibited a similar pattern of effects for complexity level ($\beta = 0.107$, s.e. $= 0.013$, $t = 8.13$, $p < 0.001$) and phase ($\beta = 1.236$, s.e. $= 0.179$, $t = 6.92$, $p < 0.001$).
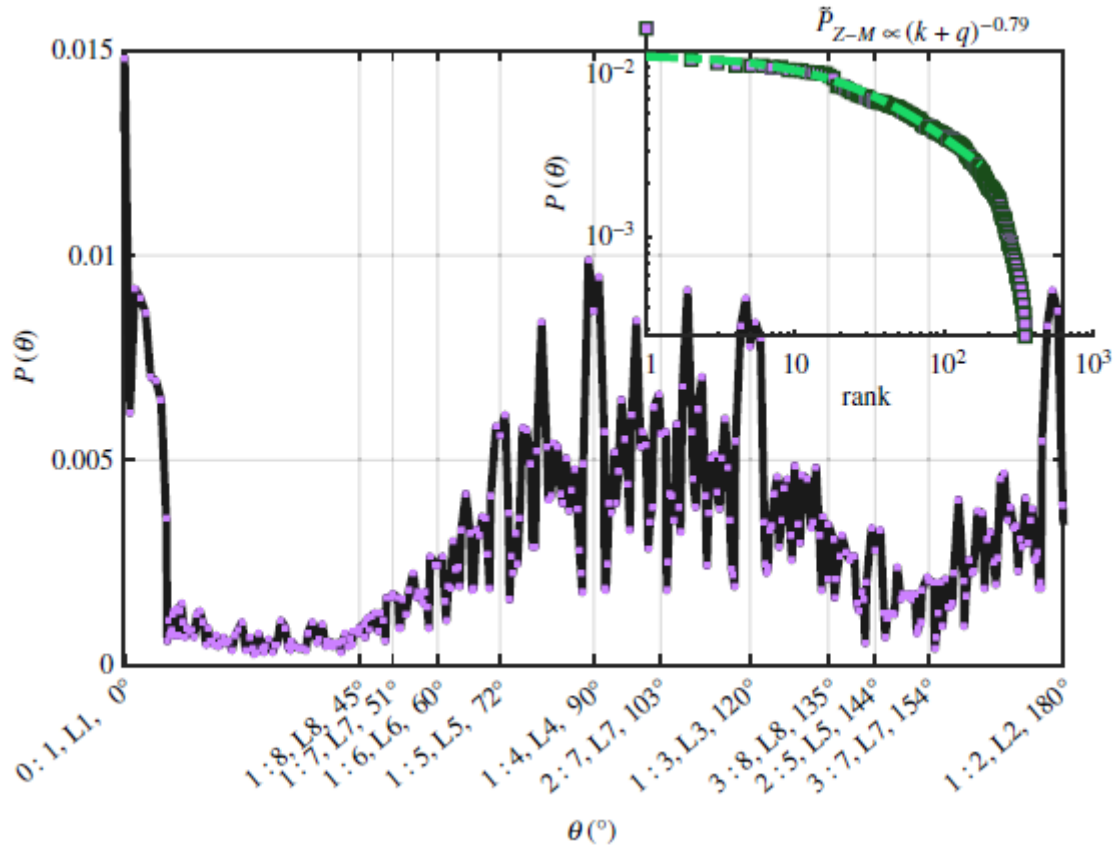

Figure 3. The distribution of tapping phases exhibits scaling. The histogram in the main panel is over data pooled from all trials and participants. The log-log inset shows the same histogram but with bins re-ordered according to their complexity rank; the fitted Zipf-Mandelbrot law is shown with a dashed line.

Rank-frequency scaling

We took the frequency distribution of all tapping phases pooled across trials and participants. Figure 3 suggests that tapping was attracted more strongly to phases described by simple ratios. Converting this histogram to a logarithmic rank-frequency plot revealed a very good fit with a Zipf-Mandelbrot scaling law $P = c(k \, \text{þ} \, q)^{-s}$ ($c = 0.17$, $q = 34.92$, $s = -0.79$, $R^2 = 0.986$).

Experiment 2: perception of phase relations in an auditory deviant detection task

In two studies, we examined the auditory perception of the phase-offset stimuli from Experiment 1. An auditory yes-no task tested detection of a small time deviation, with the hypothesis being that the detection would depend on the Farey level of the stimulus.

Methods
Participants

Undergraduate students from the local psychology department (experiment 2A: $N = 58$, age range 18–37 years, median = 19; 34/24 female/male; 2B: $N = 62$, age range 17–21 years, median = 18; 50/12 female/male) completed the task online on their personal computers at undisclosed locations in exchange for extra course credit. The research was approved by the McMaster Research Ethics Board (MREB#2164).

Apparatus

The experiment was administered remotely and was fully browser-based, implemented with jsPsych plugins [24], and hosted on a commercial server (pavlovia.org). Using designated yes/no keyboard buttons, participants indicated if the last tone in a short sound stimulus deviated from its expected temporal location.

Stimuli

The stimuli had the same structure as in the synchronization stage of Experiment 1 (figure 1a–c) but with only four repetitions. The phases (listed in electronic supplementary material, table) were taken from a denser set in Experiment 2A, 24 phases up to level 12 of the Farey tree, and a sparser set in 2B, 10 phases but three trials of each. The last sound of the stimulus pattern, the last secondary tone, was subject to deviation $\tau = \pm\{100, 70, 40, 20, 10, 0\}$ ms in A, $\tau = \pm\{200, 120, 80, 40, 20, 0\}$ ms in B. Crossing all phases and deviants resulted in 264 different stimuli in A, 110 in B. The left side was leading, unlike in Experiment 1. For a more pleasant sound over repetitive observation, tones with sharp wood- block sounds were used (90 ms duration, 5 ms attack, 85 ms linear decay), peak power at 880/968 Hz on left/right, respectively, mixed with a noisy waveform with a broad power spectrum to make the task more difficult.

Procedure

Participants visited a series of web pages in order: the department's recruitment site, an information and consent sheet, an anonymized demographics questionnaire, and the auditory task, where they also saw demonstration and practice trials, and a headphone usage detection task [25], followed by the task of interest.

Measures

Performance was evaluated in terms of reaction times and the threshold and sensitivity at 50% probability of detection, separately for each phase and positive and negative deviations. This is an exploratory study with an important limitation: few samples were collected at each stimulus level in order to sample a wide range of phase relations and deviations within a reasonable amount of time. In Study A, it was impossible to fit psychometric curves for each participant. Instead, we calculated per cent correct responses across participants. We eliminated trials with stimulus phase lower than 90° because of erratic responses.

Results

The results from Experiment 2A and 2B partially agreed with our hypothesis but were somewhat inconsistent. In Experiment 2A, simple linear regressions did not find an association between detection threshold and ratio complexity ($\beta_0 = 81.77$, $\beta_1 = 1.13$, $R^2 = 0.06$, $F_{1,24} = 1.57$, $p = 0.22$). However, sensitivity magnitude exhibited a negative association with complexity ($\beta_0 = 0.043$, $\beta_1 = -0.001$, $R^2 = 0.19$, $F_{1,24} = 5.43$, $p < 0.05$; electronic supplementary material, figure S4a), in agreement with our hypothesis. Linear mixed-effects models for the reaction times did not show an effect of complexity level ($t < 1$). For the hit rates, a logistic mixed-effects model did not find an association with complexity level ($z < 1$).

In Experiment 2B, sensitivity and threshold parameters were obtained per individual. Linear mixed-effects models indicated an association between threshold and complexity level ($\beta = 1.954$, s.e. $= 0.374$, $t = 5.229$, $p < 0.001$), in agreement with our hypothesis, and a trend for an effect of complexity level on sensitivity ($\beta = -0.0088$, s.e. $= 0.0048$, $t = -1.832$, $p = 0.067$; electronic supplementary material, figure S4b). Reaction times tended to increase with level, but the effect was only marginally significant ($\beta = 3.427$, s.e. $= 1.674$, $t = 1.94$, $p = 0.052$). For the hit rates, a logistic mixed-effects model showed that correct detections were negatively associated with

complexity level ($\beta = -0.0217$, s.e. $= 0.0045$, $z = -4.851$, $p < 0.001$).

Cross-frequency coupling model for bimanual phase tapping

We used a minimal mathematical model of cross-frequency synchronization, the phase-attractive circle map [18].

$$\theta_{n+1} = \theta_n + \Omega - \frac{K}{2\pi}(1 + A\cos 2\pi\theta_n)\sin 2\pi\theta_n$$
$$+ \sigma\Gamma \,(\text{modulo } 1).$$

The so-called bare winding number Ω equals p rotations of a slow oscillator per q rotations of a fast oscillator. Coupling strengths K or A different from zero mean that every fast cycle 'kicks' the slow one according to the given nonlinear coupling terms. Coupling leads the system to mode-lock on periodic behaviour with a so-called dressed winding number W that can be understood as an effective rate of rotation. Increasing K or A widens the range of initial conditions Ω for which the system mode-locks to low-integer ratios W. The width of mode-locking regions, known as Arnol'd tongues, is inversely proportional to the level of the Farey tree of the given W, also expressed as a $p{:}q$ ratio. This has helped explain how the stability of multifrequency synchronization is linked to simplicity of the required integer ratio in motor behaviour [17,18]. To endow the simulations with some of the variability characteristic of human performance, we added a zero-mean Gaussian noise source Γ with standard deviation σ. Note that strict mode-locking is unlikely to be observed in the presence of noise but for the simplest ratios. This does not prevent the winding rates from being weakly attractive in their vicinity (see electronic supplementary material, figure S5). This also agrees with our bimanual tapping data, where the phase drifted widely, as confirmed by figure 1d and the large s.d. in figure 2, but also tended to slow down and be observed more frequently in the vicinity of some ratios.

The relevance of this model to the present behavioural task is contingent on two important assumptions. First, the increment of the circle map, the local[2] winding number w_n $=[\theta_{n+1} - \theta_n]$ mod 1, and not phase *per se*, is what stands for the control process timing the secondary tap after the leading tap. This is consistent with recent findings that in synchronization tasks the timing of taps is associated with the speed and/or amplitude of neural rotational dynamics in relevant areas of the monkey brain [26,27]. Second, a motor inhibition prevents some repeated rotations (from the faster oscillator) from being converted to behaviour.

Consider that in the circle map one oscillator can rotate a number of times relative to the other. Presumably, a slow oscillator times the beginning of a tapping cycle and a fast oscillator times the waiting time for the following tap. Yet, in a bimanual tapping task each hand taps only once per cycle. This implies that an additional gating process is needed to block secondary taps without interrupting the covert timing mechanism. Without needing to speculate about possible neural mechanisms for such inhibition, it suffices to say that with training, macaque monkeys learn to engage in covert timing while suppressing their movements [28,29]. Together, these are strong assumptions and as such can be theoretically productive. We propose a hierarchy with dynamics on different time scales instead of two parallel commensurate dynamics such as in modern versions of the well-known Haken-Kelso-Bunz model [30,31].

Methods

The model system was iterated 30 steps and the local winding rate w_n, not the phase, was kept for analysis. On each trial, the parameter Ω was either a ratio like the ones in the behavioural task or a random initial condition. A total of 480 trials were collected per model configuration. The parameter space was explored by varying the coupling strength, $K = \{0.0, 0.1, \ldots, 1.2\}$, and the second coupling term, $A = \{0.0, 0.1, \ldots, 1.\}$, while the standard deviation of the Gaussian noise was kept fixed, $\sigma = 0.02$. The code for this model and the respective figures (figure 4; electronic supplementary material, figures S5 and S8) is available from an online repository.[3]

Results

We tested whether the pattern of results seen in the bimanual tapping phase task could be reproduced. First, as predicted, we found that variability increased with the ratio complexity of the driving frequency Ω (figure 4). A second aspect of the tapping data that was reproduced qualitatively was the clustering of winding rates around low-integer ratios and the scaling property of this distribution expressed as a log-log rank-frequency plot, $s = -0.51$ (see electronic supplementary material, figure S5). The model exhibited multimodal behaviour (see electronic supplementary material, figures S6 and S7). Future work could address whether multimodal behaviour in the tapping task can be understood in terms of how

coexisting attractive states in the model depend on coupling and noise parameters. Importantly, large parts of the parameter space in the range $0 < A < 1$ and $0 < K < 1$ agreed, qualitatively at least, with the key properties of the tapping study, namely, increasing variability with stimulus level and scaling (see electronic supplementary material, figure S8).
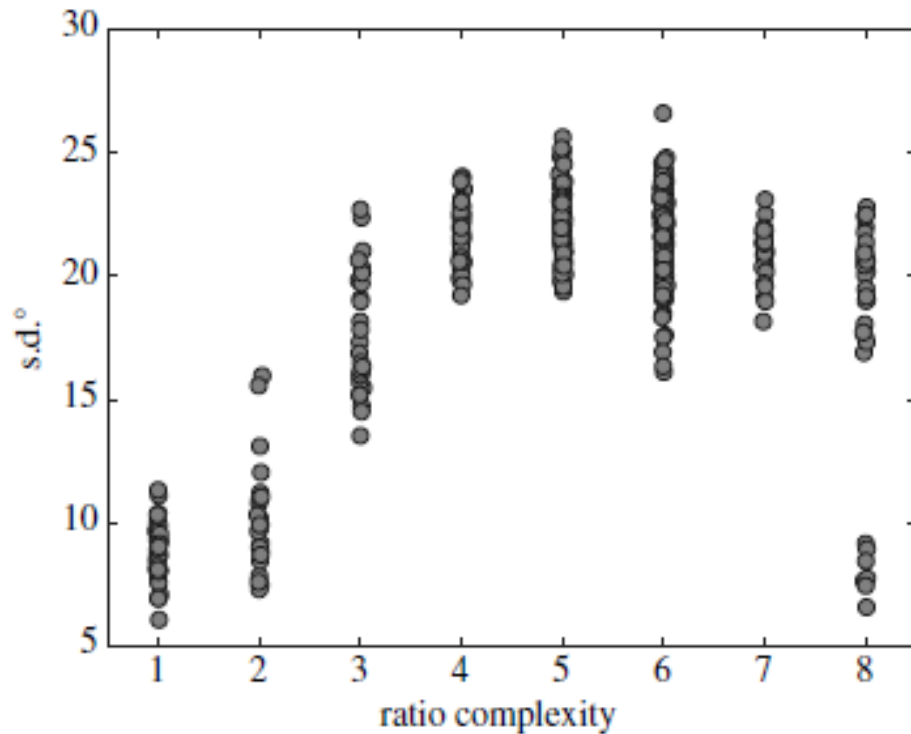

Figure 4. The relation between variability of winding rate in the phase- attractive circle map and complexity level of its driving frequency, the ratio $\Omega = p : q$, for a selected parameter configuration, $K = 0.6$, $A = 0.5$, $\sigma = 0.02$. Each trial, circular standard deviation was used to analyse the increments $w_n = [\theta_{n+1} - \theta_n]$ mod 1, after converting them to the necessary scale by multiplying by 360. Note that the multimodality at higher complexity levels is not of the same nature as in figure 3 but is related to the fact that very different initial conditions can fall on the same tree level when approximating random stimuli to ratios.

Cross-frequency coupling model for iterated rhythms

A theoretical approach is stronger if it generalizes easily beyond the specific task it was developed for originally. We considered the iterated single-hand tapping of rhythms because it also exhibits a tendency towards simple ratios. Historically, iterated tapping of rhythms and bimanual coordination of oscillatory movements have been addressed separately, oftentimes using different explanatory frame- works. The task considered here consists of tapping a three- interval rhythm repetitively. On each trial, the rhythm is given initially by an

auditory stimulus with three successive tones, consisting of intervals that may or may not obey certain musical intuitions. Tapping this rhythm for a number of repetitions after the stimulus has been silenced tends to converge closely but not perfectly to time intervals described as simple ratios [11]. An addition to this task consists of the rhythm iteration procedure: on each trial, participants initially tap to a rhythm with intervals randomly chosen. Then they tap the stimulus that is a copy of their most recently produced pattern and this is repeated several times. The results of this task show that people's tapping tends to converge on simple-ratio patterns, but it also reveals differences across cultures that can be thought of as culturally specific priors [4].

Figure 5. Projection of a three-dimensional histogram of convergent relative winding rates of three independent circle maps, treated here as intervals of a rhythm. Sides of the triangle, each on the scale from zero to one, represent proportion relative to the full rhythm which is the concatenation of the three circle maps. The heat map is for probability density. Crosses mark locations where the corresponding rates reduce to simple ratios. The driving frequencies of each of the three circle maps were random in the first trial of a simulation but at the beginning of each new block they were updated from the last effective winding numbers from the previous block, eventually converging on simple integer ratios. (Online version in colour.)

Methods

In this modelling experiment, we assumed that three independent oscillatory processes specified by equation (4.1) were activated in a chain to produce three intervals. Other parameters such as number of taps and iterations were set after those found in

behavioural experiments reported else- where [4]. The trials consisted of five blocks of simulated tapping, each block consisting of ten repetitions of the rhythm. The first block started with a set of three random rotation numbers from a non-uniform distribution in the range $\Omega_n \in (0,3]$, $n = \{1,2,3\}$, and a mode[4] at one. The effective winding numbers W_n of the independent oscillators at the end of a block were carried over as Ω_n in the next block. We simulated trials, $N = 10,000$, using a hand-picked set of parameters, $K = 1.1$, $A = 0.8$, $\sigma = 0.005$, that exhibited reason- able variability as well as similarity to empirical results. The code for this model and figure 5 is available from an online repository[5].

Results

Figure 5 is a projection of the three-dimensional histogram of the convergent winding rates of each of the three indepen- dent oscillators. The heat map indicates regions of high concentration and, for reference, low-ratio locations are marked. It shows that, despite the random initial conditions, this process tends to converge towards simple integer ratios. Here all low-integer combinations tend to be represented. This stands in contrast to cross-cultural studies where different human cultures show bias for some combinations of simple integer ratios over others, consistent with a mixture of innate constraints and experience [4].

We found that the coupling strengths K and A needed to be higher and noise lower than in the bimanual simulation study. It is known that past the critical $K > 1$ the circle map exhibits different properties. Among others, its regions of stability (the Arnol'd tongues) overlap, meaning that with added noise convergence to the lowest levels of the tree should be faster but jumps back and forth are possible too. More work is needed to study in detail different dynamical regimes of CFC and how they relate to various behavioural phenomena.

Discussion

The present work stresses the importance of complex and unstable forms of coordination beyond in-phase and anti- phase synchronization. A bimanual tapping study explored the hypothesis that additional coordination modes exist in between the stable in-phase and anti-phase movements typically reported in the literature. We expressed phase as a proportion of the cycle and then used the levels of the Farey tree to define a metric of ratio complexity, drawing a prediction about the relative stability of tapping. We found evidence

linking ratio complexity to the accuracy and variability of tapping at unusual phases. Albeit not stable modes, these phases impact performance, and can be thought of as so-called ghost attractors. Interestingly, the auditory perceptual study provided convergent but weak evidence. This suggests that the phenomena discussed here are strongly dependent on involvement of the motor system, although it is also possible that the unusual testing circumstances resulted in poorer experimental control.

A second kind of evidence came from the distribution of tapping phases which, pooled across trials and participants, exhibited a hierarchical property in the form of a rank-frequency scaling law. We then confirmed that these phenomena are reproduced qualitatively by the phase-attractive circle map, a CFC approach. As an added feature, the same model- ling approach could capture previously reported privileged simple ratio relations in rhythmic three-interval tapping pat- terns. This convergence suggests that small integer ratios are easier to perform because CFC oscillators are more stable at such ratios, contributing to our understanding of phenomena as diverse as attractor states in coordination and universal tendencies in musical behaviour.

CFC provides an alternative path to addressing the rich repertoire of coordination patterns exhibited by humans and other animals. Speculatively, the plausibility of the phase-attractive circle map needs to be discussed in light of related phenomena in neural dynamics. There is growing evidence that cross-frequency hierarchical coupling is intrinsic to, and may help organize, neural dynamics in individual neurons [32] as well as populations of neurons [21], and is linked to diverse scenarios involving many different cognitive and motor functions [33,34], including rhythm and musical meter [8,35]. In this context, it is worth raising the question whether intrinsic neural CFC implements a low-level prior for rhythmic behaviour in general.

Our claim is not that CFC fully explains both low-level sensorimotor coordination processes and musical rhythms, only that it provides a so-called enabling constraint [36]. In this sense, intrinsic neural dynamics understood as an enabling constraint act as predispositions that facilitate and bias the emergence of rhythmic behaviour but do not completely determine it. Other sources of constraint include the musical experience and enculturation that begins very early in development as additive or subtractive processes of perceptual narrowing [13]. The heterogeneities that are observed between different musical cultures [4] likely reflect

these effects of experience and enculturation. To complicate the picture further, some musical traditions include rhythms that are complex in ways that cannot be framed as a hierarchy of integer ratios [37], suggesting that either the current model is too simplistic or that experience can be very powerful.

The present work led to another interesting observation, that a scaling law described the distribution of coordination modes and was matched qualitatively by the winding numbers of the phase-attractive circle map. This could be a natural consequence of the hierarchical branching nature of the Farey tree [38]. This observation also makes sense in light of so-called universal scaling laws in musical and temporal behaviour of humans and other animals [22,39,40]. It is beyond the scope of this paper to propose a focused explanation in relation to scaling phenomena in general. It suffices to say that such phenomena appear when complex systems are poised near a critical point of instability, and balance opposite constraints such as for stability and diversity.

thank Till Frank, Pablo Padilla and Claire Michaels for encouraging earlier versions of these ideas.

Endnotes

[1]An alternative, yet closely, related definition is possible in terms of the Farey sequence rather than the tree (see electronic supplementary material, table), in which case the level corresponds to the ratio denominator.

[2]This is different from the usual definition of the so-called dressed winding number W; $\lim_{n \to \infty} (1/n)\theta_n$, the mode-locked rate at which phase increments converge to in the long limit.

[3]https://gitlab.com/dodo_bird/rstb_2021.

[4]Using a uniform distribution resulted in the same overall pattern but weaker tendency towards integer ratios.

[5]https://gitlab.com/dodo_bird/rstb_2021.

References

1. Greenfield MD, Marin-Cudraz T, Party V. 2017 Evolution of synchronies in insect choruses. *Biol. J. Linn. Soc.* 122, 487–504. (doi:10.1093/biolinnean/blx096)

2. Ravignani A, Verga L, Greenfield MD. 2019 Interactive rhythms across species: the evolutionary biology of animal chorusing and turn-taking. *Ann. N. Y. Acad. Sci.* 1453, 12–21. (doi:10.1111/nyas.14230)

3. Ravignani A, Bowling D, Fitch WT. 2014 Chorusing, synchrony and the evolutionary functions of rhythm. *Front. Psychol.* 5, 1–15. (doi:10.3389/fpsyg.2014.01118)

4. Jacoby N, McDermott JH. 2017 Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Curr. Biol.* 27, 359–370. (doi:10.1016/j.cub.2016.12.031)

5. Terhardt E. 1974 Pitch, consonance, and harmony. *J. Acoust. Soc. Am.* 55, 1061. (doi:10.1121/1.1914648)

6. Trainor LJ. 2015 The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Phil. Trans. R. Soc. B* 370, 20140089. (doi:10.1098/rstb.2014.0089)

7. Grahn JA. 2012 Neural mechanisms of rhythm perception: current findings and future perspectives. *Top. Cogn. Sci.* 4, 585–606. (doi:10.1111/j.1756- 8765.2012.01213.x)

8. Ravignani A, Thompson B, Lumaca M, Grube M. 2018 Why do durations in musical rhythms

conform to small integer ratios? *Front. Comput. Neurosci.* 12, 86. (doi:10.3389/FNCOM.2018.00086)

9. Loehr JD, Palmer C. 2009 Subdividing the beat: auditory and motor contributions to synchronization. *Music Percept.* 26, 415–425. (doi:10.1525/mp.2009.26.5.415)

10. Cummins F, Port R. 1998 Rhythmic constraints on stress timing in English. *J. Phon.* 26, 145–171. (doi:10.1006/jpho.1998.0070)

11. Repp BH, London J, Keller PE. 2013 Systematic distortions in musicians' reproduction of cyclic three-interval rhythms. *Music Percept.* 30, 291–305. (doi:10.1525/MP.2012.30.3.291)

12. Repp BH, London J, Keller PE. 2012 Distortions in reproduction of two-interval rhythms: when the 'Attractor Ratio' is not exactly 1:2. *Music Percept.* 30, 205–223. (doi:10.1525/mp.2012.30.2.205)

13. Hannon EE, Trainor LJ. 2007 Music acquisition: effects of enculturation and formal training on development. *Trends Cogn. Sci.* 11, 466–472. (doi:10.1016/j.tics.2007.08.008)

14. Trainor LJ, Marsh-Rollo S. 2019 Rhythm, meter, and timing: the heartbeat of musical development. In *The Oxford handbook of music and the brain* (eds MH Thaut, DA Hodges), pp. 591–622. Oxford, UK: Oxford University Press.

15. Hannon EE, Trehub SE. 2005 Metrical categories in infancy and adulthood. *Psychol. Sci.* 16, 48–55. (doi:10.1111/j.0956-7976.2005.00779.x)

16. Bouvet CJ, Varlet M, Dalla Bella S, Keller PE, Zelic G, Bardy BG. 2019 Preferred frequency ratios for spontaneous auditory-motor synchronization: dynamical stability and hysteresis. *Acta Psychol.* 196, 33–41. (doi:10.1016/j.actpsy.2019.03.007)

17. Peper CE, Beek PJ, van Wieringen PCW. 1995 Multifrequency coordination in bimanual tapping: asymmetrical coupling and signs of supercriticality. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1117–1138. (doi:10.1037//0096-1523.21.5.1117)

18. deGuzman GC, Kelso JAS. 1991 Multifrequency behavioral patterns and the phase attractive circle map. *Biol. Cybern.* 64, 485–495. (doi:10.1007/ BF00202613)

19. Hoffmann CP, Bardy BG. 2015 Dynamics of the locomotor–respiratory coupling at different frequencies. *Exp. Brain Res.* 233, 1551–1561. (doi:10.1007/s00221-015-4229-5)

20. Mathias B, Zamm A, Gianferrara PG, Ross B, Palmer C. 2020 Rhythm complexity modulates behavioral and neural dynamics during auditory–motor synchronization. *J. Cogn. Neurosci.* 32, 1864–1880. (doi:10.1162/jocn_a_01601)

21. Lakatos P, Gross J, Thut G. 2019 A new unifying account of the roles of neuronal

entrainment. *Curr. Biol.* 29, R890–R905. (doi:10.1016/J.CUB.2019. 07.075)

22. Mehr SA *et al.* 2019 Universality and diversity in human song. *Science* 366, aax0868. (doi:10.1126/ science.aax0868)

23. Sowiński J, Dalla Bella S. 2013 Poor synchronization to the beat may result from deficient auditory- motor mapping. *Neuropsychologia* 51, 1952–1963. (doi:10.1016/j.neuropsychologia.2013.06.027)

24. de Leeuw JR. 2015 jsPsych: a JavaScript library for creating behavioral experiments in a web browser. *Behav. Res. Methods* 47, 1–12. (doi:10.3758/ s13428-014-0458-y)

25. Woods KJP, Siegel MH, Traer J, McDermott JH. 2017 Headphone screening to facilitate web-based auditory experiments. *Atten. Percep. Psychophys.* 79, 2064–2072. (doi:10.3758/s13414-017-1361-2)

26. Gámez J, Mendoza G, Prado L, Betancourt A, Merchant H. 2019 The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biol.* 17, e3000054. (doi:10.1371/ journal.pbio.3000054)

27. Egger SW, Remington ED, Chang CJ, Jazayeri M. 2019 Internal models of sensorimotor integration regulate cortical dynamics. *Nat. Neurosci.* 22, 1871– 1882. (doi:10.1038/s41593-019-0500-6)

28. Donnet S, Bartolo R, Fernandes JM, Cunha JPS, Prado L, Merchant H. 2014 Monkeys time their pauses of movement and not their movement- kinematics during a synchronization-continuation rhythmic task. *J. Neurophysiol.* 111, 2138–2149. (doi:10.1152/jn.00802.2013)

29. Gámez J, Yc K, Ayala YA, Dotov D, Prado L, Merchant H. 2018 Predictive rhythmic tapping to isochronous and tempo changing metronomes in the nonhuman primate. *Ann. N. Y. Acad. Sci.* 1423, 396–414. (doi:10.1111/nyas.13671)

30. Avitabile D, Słowiński P, Bardy B, Tsaneva- Atanasova K. 2016 Beyond in-phase and anti-phase coordination in a model of joint action. *Biol. Cybern.* 110, 201–216. (doi:10.1007/s00422-016-0691-9)

31. Słowiński P, Tsaneva-Atanasova K, Krauskopf B. 2016 Effects of time-delay in a model of intra- and inter-personal motor coordination. *Eur. Phys. J. Spec. Top.* 225, 2591–2600. (doi:10.1140/epjst/e2015-50327-6)

32. Zhang Y, Bose A, Nadim F. 2009 The influence of the A-current on the dynamics of an oscillator- follower inhibitory network. *SIAM J. Appl. Dyn. Syst.* 8, 1564–1590.

(doi:10.1137/090760994)

33. Hyafil A, Giraud AL, Fontolan L, Gutkin B. 2015 Neural cross-frequency coupling: connecting architectures, mechanisms, and functions. *Trends Neurosci.* 38, 725–740. (doi:10.1016/j.tins. 2015.09.001)

34. Damm L, Varoqui D, De Cock VC, Dalla Bella S, Bardy B. 2020 Why do we move to the beat? A multi-scale approach, from physical principles to brain dynamics. *Neurosci. Biobehav. Rev.* 112, 553–584. (doi:10.1016/j.neubiorev.2019.12.024)

35. Large EW. 2008 Resonating to musical rhythm: theory and experiment. In *Psychology of time* (ed. S Grondin), pp. 189–231. Bingley, UK: Emerald Group Publishing.

36. Baggs E, Raja V, Anderson ML. 2020 Extended skill learning. *Front. Psychol.* 11, 1956. (doi:10.3389/ fpsyg.2020.01956)

37. London J. 1995 Some examples of complex meters and their implications for models of metric perception. *Music Percept.* 13, 59–77. (doi:10.2307/40285685)

38. Annand CT, Fleming SM, Holden JG. 2021 Farey trees explain sequential effects in choice response time. *Front. Physiol.* 12, 611145. (doi:10.3389/fphys. 2021.611145)

39. Kello CT, Dalla Bella S, Médé B, Balasubramaniam R. 2017 Hierarchical temporal structure in music, speech and animal vocalizations: jazz is like a conversation, humpbacks sing like hermit thrushes. *J. R. Soc. Interface* 14, 20170231. (doi:10.1098/rsif. 2017.0231)

40. Beltrán del Río M, Cocho G, Naumis GG. 2008 Universality in the tail of musical note rank distribution. *Phys. A Stat. Mech. Appl.* 387, 31. (doi:10.1016/j.physa.2008.05.031)