

Attention to rhythms: sensory-specific constrained sampling of temporal regularities

Arnaud Zalta^{1,2}, **Spase Petkoski**¹, **Benjamin Morillon**^{1*}

¹ Aix Marseille Univ, Inserm, INS, Inst Neurosci Syst, Marseille, France

² Aix Marseille Univ, APHM, INSERM, Inst Neurosci Syst, Service de Pharmacologie Clinique et Pharmacovigilance, Marseille, France

Corresponding author: Dr. Benjamin Morillon

Aix Marseille Univ, Inserm, INS, Inst Neurosci Syst, Marseille, France

Faculté de Médecine, 27, Boulevard Jean Moulin, 13005 Marseille

E-mail: bnmorillon@gmail.com

Abstract (142 words)

That attention is a fundamentally rhythmic process has recently received abundant empirical evidence. The essence of temporal attention, however, is to flexibly focus in time. Whether this function is hampered by an underlying rhythmic mechanism is unknown. In six interrelated experiments, we behaviourally quantify the sampling capacities of periodic temporal attention during auditory or visual perception. We reveal the presence of limited attentional capacities, with an optimal sampling rate of ~ 1.4 Hz in audition and ~ 0.7 Hz in vision. Investigating the motor contribution to temporal attention, we show that it scales with motor rhythmic precision, maximal at ~ 1.7 Hz. Critically, the motor modulation is beneficial to auditory but detrimental to visual temporal attention. These results are captured by a computational model of coupled oscillators, that reveals the underlying structural constraints governing the temporal alignment between motor and attention fluctuations.

Introduction

Adapting our behaviour according to external stimuli requires extraction of relevant sensory information over time ¹. This ability relies on the capacity to flexibly adapt and adjust our temporal attention to the natural dynamics of the environment. For instance, sailing on tumultuous seas, following the flow of an animated speaker or listening to drums in an ebullient jazz band requires specific tuning of our temporal attention capacities. But in some cases, this ability fails. A succession of events appearing too fast, or two stimuli too close in time, are situations typically difficult to attend to. Multiple types of temporal structures are capable of guiding temporal attention ¹, such as isochronous ² or heterochronous streams of events ³, symbolic cues ⁴, or hazard functions ⁵. Paradigms that involve isochronous perceptual streams in the auditory and/or visual modality are consistently designed with rhythms in the 1-2 Hz frequency range ^{2,6-27}, which, incidentally, corresponds to the natural musical beat ²⁸. Strikingly, the propensity to flexibly focus in time and its limits – in other words, the sampling capacities of temporal attention – have never been investigated.

Contrary to the continuous flow of perceptual events, actions are coordinated and rhythmic. For instance, walking is intrinsically rhythmic and operates at ~2 Hz ²⁹. Spontaneous motor rhythmic behaviours such as finger tapping also operate at a preferred tempo of ~1.5-2 Hz and motor tapping has an optimal temporal precision within the range of 0.8-2.5 Hz ³⁰⁻³². Moreover, delta (0.5-4 Hz) neural oscillations shape the dynamics of motor behaviour and motor neural processes ³³. For instance, during production of complex motor behaviours such as speech, the coordination of articulatory movements is encoded in kinematic trajectories characterized by damped oscillatory dynamics ³⁴. Even during non-periodic motor behaviours, such as reaching, motor trajectories are encoded in neural dynamical patterns that oscillate around 1-2 Hz ³⁵⁻³⁷. And crucially, the motor cortex exhibits resting-state dynamics at the delta rate ^{38,39}. Thus, delta oscillations are an intrinsic rhythm of the motor system visible in the dynamics of most basic motor acts.

In line with the active sensing framework, perception involves motor sampling routines like sniffing and whisking in rodents or visual search in primates ⁴⁰⁻⁴². Attention is an essential

component of this process, its influence helping to impose the motor sampling pattern on the relevant sensory stream^{13,41}. Accordingly, previous studies showed that overtly moving during an auditory attention task improves perceptual performance^{13,14}. Importantly, these experiments were also performed at 1.5 Hz, which both corresponds to the rhythm classically used to investigate periodic temporal attention and to the natural rate of rhythmic movements. In virtue of the fundamental relationship between motor and active/attentive sensory processes, one could hypothesize that the sampling capacities of periodic temporal attention derive from those of the motor system and are thus limited around 1.5 Hz. Alternatively, one could hypothesize that temporal attention is not rate-restricted but that it is the motor benefit to temporal attention that is restricted around this rate. Finally, the rhythmic sampling rate of visual sustained attention was recently shown to be restricted around 4-8 Hz^{43,44} which suggests that sensory-specific temporal constraints could also shape the sampling capacities of temporal attention.

To investigate these issues, we developed a paradigm to behaviourally quantify the sampling capacities of periodic temporal attention during auditory and visual perception. The quality of temporal attention was estimated for different rhythms, ranging from ~0.5 to ~4 Hz. In each modality, we first investigated temporal attention during passive perception – i.e. without overt motor involvement – and then quantified in another set of experiments the motor contribution to temporal attention. Through six interrelated behavioural experiments, we reveal the existence of a limited sampling capacity of temporal attention, which is moreover sensory-specific. Besides, we highlight that the motor contribution to temporal attention is also sensory-specific and derives from the compatibility of temporal dynamics underlying motor and sensory-specific attentional processes. Finally, we show that our results are reproduced by a simple model involving three coupled oscillators. While the optimal sampling rate of temporal attention is directly reflected in the natural frequency of the attentional oscillator, the quality of the motor modulation crucially depends on the time-delay in the coupling between the stimulus and the motor oscillator.

Results

Investigating the rhythmic sampling capacities of temporal attention. The tasks of this study are all based on the same paradigm. Sequences of stimuli were presented on each trial, for 2 to ~10 s. Three reference stimuli defining the isochronous beat of the sequence preceded a mixture of on- and off-beat stimuli. Participants performed a beat discrimination task at the end of each trial, by deciding whether the last stimulus of the sequence, a deviant, was on or off beat (Fig. 1a and Fig. 3a). While on-beat stimuli were providing the beat, off-beat stimuli had a distracting influence. This interleaved delivery of sensory events forced participants to track the beat throughout the entire duration of the sequence. This protocol thus ensured that their attentional focus was temporally modulated over an extended time period. The density of distractors (i.e. number of distractors per beat) was adjusted for each participant prior to the experiment to reach threshold performance for a 2 Hz beat (see [Methods](#)). The beat frequency varied from ~0.5 Hz to 3.8 Hz across conditions, to span most of the range of discernible beats ^{45,46}.

Auditory periodic temporal attention optimally operates at ~1.2-1.5 Hz. In a first passive auditory experiment (exp. 1) we used pure tone stimuli. Eight conditions were investigated with isochronous beats of 0.6, 0.7, 1, 1.3, 1.7, 2.2, 2.9 and 3.8 Hz (Fig. 1a; see [Methods](#)). The average difficulty level (density of distractors) was around 1 ($M = 1.03$; $SD = .70$; Fig. 1b). The comparison of conditions revealed significant fluctuations in performance (% correct responses) across beats (repeated-measures ANOVA: $F(7,196) = 17.6$, $p < .001$; Fig. 1c). They moreover had an inverse U-shape profile, which could be properly approximated with a third-degree polynomial function ($R^2(\beta) = .86$, $p = .002$; see [Methods](#)), whose local maximum estimates the beat at which performance is optimal. Estimates of the *optimal beat* measured with individual fits revealed that auditory temporal attention has an optimal rhythmic sampling frequency of ~1.2 Hz ($M = 1.23$ Hz; $SD = .53$ Hz; Fig. 1d). Data from one participant could not be correctly fitted and were excluded from this analysis. These results suggest that during auditory perception temporal attention presents an optimal sampling rate, around 1.2 Hz.

In this experiment, tones lasted 10 % of the beat period (see [Methods](#)). Our results could thus be due to the existence of either an optimal beat or an optimal stimulus duration, during

auditory temporal attention. In a subsequent control experiment, we thus orthogonalized beats and stimulus duration, by fixing across conditions tone length to 22.5 ms (Fig. S1). We replicated all findings of experiment 1 (repeated-measures ANOVA: $F(7,98) = 8.9$, $p < .001$; difficulty level: $M = .95$; $SD = .63$; 3rd order fit quality: $R^2(8) = .86$, $p = .002$; individual estimates of optimal beat: $M = 1.32$ Hz; $SD = .48$ Hz). This indicates that fluctuations of performance across conditions are due to the existence of an optimal beat at which auditory temporal attention operates.

Motor contribution to auditory temporal attention. In a second experiment (exp. 2), we investigated whether motor activity helps to synchronize temporal fluctuations of attention with the timing of events in a task-relevant stream. Participants carried out two sessions. A ‘passive’ one where they performed the task while staying completely still for the duration of the trial (as in exp. 1), and a ‘tracking’ one, where they hit the reference beat in phase with their index finger on a noiseless pad. Therefore, the absence or presence of overt movement was the single difference between the two sessions. While it is not possible to control for the covert involvement of motor and/or premotor structures during temporal attention tasks, comparing passive and tracking sessions allowed us to quantify the influence of overt (relative to covert) motor activity on the precision of temporal attention.

We observed significant fluctuations in performance across beats (repeated-measures ANOVA, condition: $F(7,133) = 15.9$, $p < .001$; Fig. 2a). The comparison of passive and tracking sessions revealed a significant difference in categorization performance (session: $F(1,19) = 7.5$, $p = .013$), which was moreover beat-selective (interaction: $F(7,133) = 2.8$, $p = .023$). Post-hoc t-tests indicated that overt motor tracking significantly increased performance only when participants performed the task between 1.3 and 2.2 Hz (paired t-tests, 1.3 Hz: $p = .013$, $t = 2.75$; 1.7 Hz: $p = .009$, $t = 2.93$; 2.2 Hz: $p = .002$, $t = 3.59$; all other beats: $t < 1.95$, $p > .05$). The inverse U-shape profile of performance could be properly approximated with a 3rd order fit for both sessions (passive: $R^2(8) = .73$, $p = .011$; tracking: $R^2(8) = .91$, $p = .001$). The optimal beat estimated with individual fits was around 1.5 Hz in both sessions (passive: $M = 1.47$ Hz; $SD = .59$ Hz; tracking: $M = 1.47$ Hz; $SD = .45$ Hz; passive vs. tracking: paired welsh t-test: $t(38) = .02$, $p = .99$; Fig. 2d). To evaluate the likelihood that this absence of difference across sessions corresponds to a genuine

absence of difference, we computed the corresponding Bayes factor (see [Methods](#)). We obtained a Bayes factor of 0.22 for this null effect, indicating that the “null” hypothesis (no difference of optimal beat frequency between sessions) is more likely than the alternative (significant difference of optimal beat frequency). These results confirm previous findings showing that overt motor activity optimizes auditory temporal attention ^{13,14} and further reveal that this benefit is rate-restricted and maximal around 1.5 Hz.

Similar optimal rates for motor tapping and auditory temporal attention. To further investigate the nature of the interaction between motor activity and auditory attention, in a third study we asked participants to perform a standard tapping experiment (BASTAA ⁴⁷; see [Methods](#)). In the absence of any sensory cue, they naturally tapped on average at ~1.7 Hz (spontaneous tapping frequency: $M = 1.67$ Hz; $SD = 0.74$ Hz; Fig. 2b), which confirms previous studies ^{29,32,48}. We also instructed participants to tap as slow and as fast as possible and found that the range of producible taps (slow: $M = .60$ Hz; $SD = .30$ Hz; fast $M = 4.69$ Hz; $SD = 1.13$ Hz) was similar to the range of discernible beats, i.e. ~0.5-4 Hz ^{45,46}.

Furthermore, we analysed the tapping precision of participants across conditions during the tracking session of experiment 2. Participants tended to tap too fast during perception of the slowest beats, too slowly for the fastest ones and tapped at the appropriate pace during presentation of a ~1.7 Hz beat (Fig. S2a). The coefficient of variation (CV; i.e. relative standard deviation) of tapping across conditions confirmed that the quality of tapping differed across conditions (repeated-measures ANOVA: $F(7,126) = 5.93$, $p = .001$; Fig. 2c). It had a U-shape profile across conditions which could be properly approximated with a 3rd order fit ($R^2(8) = .95$, $p < .001$). Strikingly, individual estimates of the beat associated to an optimal tapping rhythmicity ($M = 1.35$ Hz; $SD = .55$ Hz; Fig. 2d) were overall similar to the optimal beat of auditory temporal attention in both passive and tracking sessions (paired welsh t-tests, tapping vs. passive: $t(38) = .69$, $p = .5$, Bayes factor = .32; tapping vs. tracking: $t(38) = .77$, $p = .45$, Bayes factor = .33). Thus, the optimal frequency of rhythmic movements in the absence or presence of synchronous periodic auditory cues is ~1.5 Hz, which is also similar to the optimal frequency of auditory temporal attention (in both the absence and presence of concomitant rhythmic movements).

The quality of motor tracking positively impacts auditory performance accuracy. To further explore these results, we investigated whether the quality of motor tapping influenced the quality of auditory temporal attention on a trial-by-trial basis. First, we compared the CV of tapping for correct and incorrect trials of the tracking session (Fig. S2b). We observed an absence of difference between trials in which participants responded correctly or incorrectly (repeated-measures ANOVA, correct vs. incorrect: $F(1,16) = .37, p = .55$, Bayes factor = .29; condition $F(7,112) = 5.12, p = .0032$; interaction: $F(7,112) = 1.02, p = .41$). We also compared the performance of participants in trials where the tapping CV was low or high, by using a median-split procedure (Fig. S2c inset). Again, while the CV in these two groups of trials was highly different (repeated-measures ANOVA, low vs. high CV: $F(1,19) = 79.7, p < .001$; condition: $F(7,133) = 6.82, p < .001$; interaction: $F(7,133) = 4.41, p = .003$), we observed similar performance between these two groups of trials (repeated-measures ANOVA, low vs. high CV: $F(1,19) = .42, p = .52$, Bayes factor = .28; condition: $F(7,133) = 16.1, p < .001$; interaction: $F(7,133) = .47, p = .79$; Fig. S2c). Overall, these results indicate that while the optimal rate of rhythmic movements and of auditory temporal attention is similar on average, there is no direct mechanistic relation between the rhythmicity of motor acts and the quality of auditory temporal attention.

Second, we investigated the temporal distance between motor acts and the beat, i.e. the degree of simultaneity of motor acts relative to the beat (sensorimotor simultaneity, see [Methods](#)). Participants tended to anticipate the beat in this modality, except when the beat was too fast (≥ 2.9 Hz; Fig. S4a). We observed an overall better sensorimotor simultaneity in trials where participants' temporal attention was accurate than in incorrect trials (repeated-measures ANOVA, correct vs. incorrect: $F(1,16) = 16.4, p < .001$; condition $F(7,112) = 1.65, p = .2$; interaction: $F(7,112) = .73, p = .6$; Fig. 2e). We also split trials in which the temporal distance between motor acts and the beat was low or high (Fig. 2f inset; repeated-measures ANOVA, low vs. high: $F(1,19) = 209, p < .001$; condition : $F(7,133) = 2.3, p = .098$; interaction: $F(7,133) = 11.84, p < .001$). We observed a significant difference of performance between these two groups of trials (repeated-measures ANOVA, low vs. high: $F(1, 19) = 30.81, p < .001$, condition: $F(7,133) = 16.13, p < .001$; interaction: $F(7,133) = 1.62, p = .16$; Fig. 2f), indicating that the ability of participants to closely track the auditory beat, vis. the quality of motor tracking, directly benefits performance accuracy.

Visual periodic temporal attention optimally operates at ~0.6-0.8 Hz. In a first visual passive experiment (exp. 4), we used visual grating stimuli. Ten conditions were investigated with isochronous beats of 0.3, 0.4, 0.6, 0.7, 1, 1.3, 1.7, 2.2, 2.9 and 3.8 Hz (Fig 3c; see [Methods](#)). Two participants did not complete the experiment and were excluded. The average difficulty level (density of distractors) was around 0.3 ($M = .28$; $SD = .18$), significantly lower than in the auditory tasks (comparison of exp. 1 & 4: unpaired welsh t test, $t(57) = -5.54$, $p < .001$; Fig. 3b). The comparison of conditions revealed significant fluctuations in performance across beats (repeated-measures ANOVA, condition: $F(9,243) = 53.6$, $p < .001$; Fig. 3c). They moreover had an inverse U-shape profile (3rd order fit: $R^2(10) = .93$, $p < .001$). The estimated local maximum of the individual level performance revealed that visual temporal attention has an optimal rhythmic sampling frequency of ~0.8 Hz ($M = .84$ Hz; $SD = .34$ Hz; Fig. 3d). These results also reveal different preferred sampling rates of temporal attention among sensory modalities, with a significantly lower optimal beat in the visual as compared to the auditory modality (comparison of individual estimates of the optimal beat in exp. 1 & 4: unpaired welsh t test: $t(57) = 3.38$, $p = .001$; Fig. 3d).

Disruptive motor contribution to visual temporal attention. In a second visual experiment (exp. 5), we investigated the motor influence on visual temporal attention across 8 conditions (Fig. 4; see [Methods](#)). We first observed significant fluctuations in performance across beats, which replicates the results of experiment 4 (repeated-measures ANOVA, condition: $F(7,133) = 62.9$, $p < .001$; Fig. 4a). The comparison of passive and tracking sessions did not revealed a significant difference in overall categorization performance (session: $F(1,19) = .56$, $p = .46$) but a beat-selective significant difference (interaction: $F(7,133) = 2.8$, $p = .03$). In contrast to auditory perception, post-hoc t-tests indicated that overt motor tracking significantly decreased performance when participants performed the task between 1.66 and 2.2 Hz (paired t-tests: 1.66 Hz: $t = -2.23$, $p = .038$; 2.20 Hz: $t = -2.67$, $p = .015$; all other beats: $t < 1.91$; $p > .05$). Like in the previous experiments, the inverse U-shape profile of performance could be properly approximated with a 3rd order fit for both sessions (passive: $R^2(8) = .95$, $p < .001$; tracking: $R^2(8) = .86$, $p = .002$). The optimal beat estimated with individual fits was around 0.7 Hz in both sessions (passive: $M = .83$ Hz; $SD = .46$ Hz; tracking: $M =$

.65 Hz; $SD = .17$ Hz; passive vs. tracking: paired welsh t-tests: $t(38) = -1.56$, $p = .13$; Bayes factor = .66; Fig. 4c).

Divergent optimal rates for motor tapping and visual temporal attention. We analysed the tapping precision of participants across conditions during the tracking session of experiment 5, which indicated that participants tended to tap too fast in all conditions (Fig. S3a). Like in the previous auditory experiment (exp. 2), the CV of tapping had a U-shape profile (3rd order fit: $R^2(8) = .78$, $p = .007$) and individual estimates of the beat associated to an optimal tapping rhythmicity were not significantly different across modalities (vision: $M = 1.40$ Hz, $SD = .77$ Hz; audition: $M = 1.35$ Hz, $SD = .55$ Hz; unpaired t test: $t(38) = .22$, $p = .83$; Bayes factor = .23). However, in the visual modality the CV of motor tapping was not significantly different across conditions (repeated-measures ANOVA, condition: $F(7,133) = .88$, $p = .48$) and overall the tapping CV was significantly lower in the auditory than the visual modality (comparison of CV averaged across comparable conditions, i.e. between 0.7 and 3.8 Hz; unpaired t test: $t(14) = 3.4$, $p = .005$).

Counter to the auditory experiment, the optimal beat for motor tapping was statistically different to the optimal beat of visual temporal attention, in both passive and tracking sessions (paired welsh t tests, tapping vs. passive $t(38) = 2.84$, $p = .01$; tapping vs. tracking $t(38) = 4.22$, $p < .001$; Fig. 4c). Thus, the optimal frequency of rhythmic movements in the presence of synchronous periodic visual stimuli reflects natural motor dynamics (~1.5 Hz) but differs from the optimal frequency of visual temporal attention which is ~0.7 Hz (in both presence and absence of concomitant rhythmic movements).

The quality of motor tracking negatively impacts visual performance accuracy. As in the auditory experiment, we compared the CV of tapping for correct and incorrect trials of the tracking session (Fig. S3b) and observed an absence of difference between trials where participants responded correctly or incorrectly (repeated-measures ANOVA, correct vs. incorrect: $F(1,18) = 1.74$, $p = .20$; condition $F(7,126) = .85$, $p = .49$; interaction: $F(7,126) = .10$, $p = .92$; Bayes factor = .55). We also compared the performance of participants in trials where the tapping CV was low or high (Fig. S3c inset; repeated-measures ANOVA, low vs. high CV: $F(1,19) = 48.5$, $p < .001$; condition: $F(7,133) = .88$, $p = .48$; interaction: $F(7,133) = .62$, $p = .61$), and observed similar

performance between these two groups of trials (repeated-measures ANOVA, low vs. high CV: $F(1,19) = .06, p = .81$; condition: $F(7,133) = 41.8, p < .001$; interaction: $F(7,133) = .33, p = .87$; Bayes factor = .23; Fig. S3c).

Investigation of the temporal distance between motor acts and the beat revealed that in the visual modality participants were not anticipating the beat but tapped in reaction to it (Fig. S4b). In contrast to the auditory modality, correct trials were moreover associated with a lower degree of sensorimotor simultaneity than incorrect trials (Fig. 4d; repeated-measures ANOVA, correct vs. incorrect: $F(1,18) = 9.38, p = .007$; condition $F(7,126) = 27.9, p < .001$; interaction: $F(7,126) = 3.45, p = .012$). These effects were most pronounced at 1 and 1.7 Hz (post-hoc paired t-tests: 1 Hz: $t = 3.18, p = .005$; 1.7 Hz: $t = 2.3, p = .034$; all other beats: $t < 1.81; p > .087$). Splitting trials in which sensorimotor simultaneity was low or high (Fig. 4e inset; repeated-measures ANOVA, low vs. high: $F(1,19) = 796, p < .001$; condition: $F(7,133) = 33.1, p < .001$; interaction: $F(7,133) = 81, p < .001$) revealed a significant difference of performance between these two groups of trials (Fig. 4e; repeated-measures ANOVA, low vs. high: $F(1,19) = 6.2, p = .022$; condition: $F(7,133) = 41.76, p < .001$; interaction: $F(7,133) = 4.56, p = .002$). The ability of participants to closely track the beat was detrimental to performance accuracy, and this effect was most pronounced at 1, 1.7 and 2.2 Hz (post-hoc paired t-tests: 1 Hz: $t = -2.93, p = .009$; 1.7 Hz: $t = -2.59, p = .018$; 2.2 Hz: $t = -2.76, p = .012$; all other beats: $t < 1.82; p > .084$). These results elucidate the observed disruptive motor contribution to visual temporal attention (Fig. 4a) by showing, in sharp contrast to the auditory modality, that the ability of participants to closely track the visual beat, vis. the quality of motor tracking, directly impairs performance accuracy. Moreover, this effect is selective for beats presented close to natural motor dynamics (~1.7 Hz; Fig. 2b). In line with the auditory results, these analyses highlight that motor impact on temporal attention crucially depends on the temporal simultaneity of motor acts relative to the beat, supporting a synergistic modulation of sensory processing that relies on the temporal alignment between motor and attention fluctuations. However, this does not explain why motor involvement positively impacts auditory temporal attention, but negatively impacts visual temporal attention.

A model of three delay-coupled phase oscillators replicates the behavioural results. Finally, we investigated whether our results could be explained by a simple neural network model. To understand the specific motor contribution to auditory and visual periodic temporal attention, each having its own optimal sampling rate, we implemented a model in which sensory-specific temporal attention behaves like a self-sustained oscillator (a structure with an intrinsic rhythm capable of being entrained coupled to a motor oscillator and entrained by an external beat⁴⁹. In its simplest realization, this results in a model of three coupled phase oscillators (stimulus (S), attention (A) and motor (M) oscillators) with time-delays and noise⁵⁰ (Fig. 5a; see [Methods](#)). We varied the frequency of the external beat (S) to mirror our different experimental conditions (between 0.3 and 3.8 Hz). The natural frequency of the sensory-specific oscillator (A) was fixed to reflect the optimal sampling rate of temporal attention, at 1.5 Hz for the auditory modality (after exp. 2; Fig. 2a) and at 0.7 Hz for the visual modality (after exp. 5; Fig. 4a). Finally, the natural frequency of the motor oscillator (M) was fixed at 1.7 Hz to reflect the spontaneous tapping frequency (after exp. 3). Coupling strengths (K), time-delays (τ) and the strength of the internal noise (D) were then adjusted to fit the different behavioural results (passive and tracking sessions in auditory and visual modalities; Fig. 6b-c). Behavioural performance was approximated by the phase-locking value (PLV⁵¹ between the external beat (S) and the sensory-specific temporal attention oscillator (A), as it reflects the capacity of the sensory-specific oscillator to entrain to the external beat.

First, the model reproduced the results of the passive auditory (exp. 2; Fig. 2a) and visual (exp. 5; Fig. 4a) experiments (Fig. 6b-c). We approximated the results of these passive experiments with very high accuracy (auditory: fit quality: $R^2 = .92$, $p < .001$; visual: fit quality: $R^2 = .95$, $p < .001$). Importantly, apart from the natural frequency of the sensory-specific temporal attention oscillator (A; which differed between auditory (1.5 Hz) and visual (0.7 Hz) experiments) and the time-delay between the stimulus (S) and the motor (M) oscillator (auditory: $\tau_{S-M} = .1$ s; visual: $\tau_{S-M} = .35$ s) all other parameters (coupling strength K, time-delay τ , and noise ξ) were similar across modalities ($K_{S-M} = 8$; $K_{S-A} = 10$, $\tau_{S-A} = .1$; $K_{A-M} = 10$, $\tau_{A-M} = 0$; $K_{M-A} = 2$, $\tau_{M-A} = 0$; $\xi_A = 5$; $\xi_M = 10$). Even if there is no explicit motor act in the passive session we assume that the motor system is already involved ($K_{M-A} = 2$), in line with a previous study (Morillon & Baillet, 2017). Second, the model also successfully reproduced the results of the tracking auditory (exp. 2; Fig.

2a) and visual (exp. 5; Fig. 4a) experiments (auditory: fit quality: $R^2 = .95$, $p < .001$; visual: fit quality: $R^2 = .95$, $p < .001$), with a notably selective modulation of performance around 1.5-2 Hz in the tracking as compared to the passive sessions, which, crucially, was respectively positive and negative in the auditory and visual modalities. The only parameter that varied between passive and active sessions was the strength of the coupling between motor and temporal attention ($K_{M-A} = 10$; vs. 2 for the modelling of the passive sessions). Overall, three parameters played a key role in reproducing the behavioural results. In addition to the natural frequency of the sensory-specific temporal attention oscillator, which varied between modalities, and the coupling strength between motor and attention oscillators K_{M-A} , which varied between passive and tracking sessions, the time-delay between the stimulus (S) and the motor oscillator (M; τ_{S-M}) was crucial for reproducing the difference of results across modalities.

Discussion

Our findings reveal the natural sampling rate of periodic temporal attention. Attention supports the allocation of resources to relevant locations, objects, or moments in a scene¹. Recent studies have revealed the rhythmic nature of sustained attention, showing that spatial^{44,53–59} or featured-based⁴⁴ attention samples visual stimuli rhythmically, tethered by the phase of a theta (4-8 Hz) neural oscillation. Importantly, this temporal constraint is orthogonal to the attended (spatial or object-based) dimension, and hence does not hinder the quality of sensory selection. Here, in contrast, we reveal a surprisingly limited capacity – restricted to a lower range (0.5-2 Hz; Fig. 1c, 2a, 3c, 4a) – of humans to flexibly adapt and adjust their temporal attention to the natural dynamics of a scene.

On the one hand, these results retrospectively explain why studies investigating periodic temporal attention are consistently designed with rhythms in the 1-2 Hz frequency range^{2,6–27}. On the other hand, they fuel recent frameworks postulating that the functional architecture of cognition is inherently rhythmic and underpinned by neural oscillatory activity generated at the population level^{60–63}.

Confirming classic work focusing on motor synchronisation to periodic stimuli^{29–31,47}, our study shows that the temporal precision of motor acts is optimal when (auditory or visual) stimuli unfold at around 1.7 Hz (Fig. 2C and 4B). But crucially, this set of experiments reveal the perceptual consequences of such sensorimotor synchronisation, by highlighting the intricate role of the motor system in temporal attention. The motor system is critically implicated in timing and time perception^{64–67} and periodic – beat-based – timing, in particular, is underpinned by striato-thalamo-cortical circuits^{64–67}. Our results confirm previous findings showing that overt motor activity optimizes auditory periodic temporal attention^{13,14}. They furthermore reveal that such an overt motor impact on temporal attention is rate-restricted and maximal around 1.7 Hz (Fig. 2a and 4a). This belongs to the delta (0.5-4 Hz) range of neural oscillations, which governs the dynamics of motor behaviour and motor neural processes³³. Our findings also support previous results showing that motor delta oscillations represent temporal information and modulate perceptual processing^{14,68–71}. Perception is thus shaped by motor activity, which unfolds at a delta rate and imposes

temporal constraints on the sampling of sensory information. Strikingly, in our experiments, outside the range of natural motor dynamics overt movements had no significant impact on temporal attention, either positive or negative. While participants were able to produce rhythmic movements between ~0.6 and 4.7 Hz (Fig. 2b), motor rhythmicity was less accurate outside ~1.7 Hz (Fig. 2b and 4b), and the inability of participants to closely track the beat – the lack of temporal simultaneity between motor acts and the beat – was associated with an absence of performance gain (Fig. 2f and 4e). Overt motor impact on temporal attention thus appears to be conditional upon the temporal alignment between motor and attention fluctuations.

An important question relates to the ubiquity of rhythmic sampling attentional mechanisms across modalities⁶⁰. In this set of experiments, we directly applied the same paradigm in two – auditory and visual – modalities. We observed in both of them, first the existence of an optimal beat at which temporal attention operates, and second that overt motor activity impacts temporal attention selectively for beats presented close to the natural motor dynamics (~1.7 Hz; Fig. 2b). We furthermore highlighted that this effect crucially depends on the temporal simultaneity of motor acts relative to the beat. Nevertheless, several crucial differences across modalities exist. First, while auditory periodic temporal attention operates around 1.5 Hz, close to natural motor dynamics, visual periodic temporal attention operates around 0.7 Hz, that is, is twice slower (Fig. 3d). Our paradigm used transient stimuli which are known to be more suited to auditory than visual perception^{72–74}. Indeed, the visual modality is ecologically more precise for capturing movement whereas audition is more adapted to transient stimuli⁷⁵. Accordingly, participants were overall much more accurate in auditory than visual temporal attention (Fig. 3b)^{73,76–80}. However, independently of the overall accuracy effect, our results highlight a sensory-specific constrained sampling of temporal regularities, rather than an amodal optimal beat at which temporal attention operates. These specific rhythmic sampling rates would thus emerge from the specific configuration of large-scale neural networks encompassing sensory (in addition to attentional and motor) regions^{81,82}. Second, we reveal that the quality of motor tracking directly benefits

performance accuracy in auditory attention, but negatively impacts it in visual attention (Fig. 2f and 4e).

These two crucial differences between auditory and visual modalities were accurately captured in a model of coupled oscillators representing the periodic stimulus, a sensory-specific temporal attention oscillator and a motor oscillator (Fig. 5). The difference of optimal sampling rate across modalities was directly related to the natural frequency of the sensory-specific temporal attention oscillator. More importantly, the time-delay between the stimulus and the motor oscillator was key in reproducing the differential impact of overt motor tracking on performance across modalities. While a small time-delay (100 ms) results in a positive motor impact on the quality of periodic temporal attention, a longer time-delay (350 ms) is associated with a disruptive effect. The presence of such a long delay in the visual modality is compatible with previous models of the visuomotor system⁸³. Overall, this model captures the motor contribution to temporal attention in two sensory modalities. It reveals the structural constraints governing the temporal alignment between motor and attention fluctuations.

In conclusion, our results reveal the limited capacities of periodic temporal attention and its optimal sampling rate in two sensory modalities. They furthermore characterize the structural constraints governing the motor contribution to temporal attention. Whether our results are specific to periodic temporal attention or generalise to other forms of temporal attending remains to be investigated^{3,19,84–87}.

Methods

Participants. 30, 20, 50, 30, 20 and 15 participants (age range: 18–35 years; 69% of females) were respectively recruited for experiments 1 to 6. The experiments followed the local ethics guidelines from Aix-Marseille University. Informed consent was obtained from all participants before the experiments. All had normal audition and vision and reported no history of neurological or psychiatric disorders. We did not select participants based on musical training.

Experimental design of the auditory experiments (n°1, 2 and 6). Auditory stimuli were sampled at 44 100 Hz and presented binaurally at a comfortable hearing level via headphones (Sennheiser HD 250 linear) in an anechoic room, using the Psychophysics-3 toolbox⁸⁸ and additional custom scripts written for MATLAB (The Mathworks). Instructions were visually displayed on a mid-grey background on a screen laptop computer (Lenovo Thinkpad T470s) situated at a viewing distance of 50 cm. The screen had a spatial resolution of 1920 by 1080 pixels and vertical refresh rate of 60 Hz. On each trial, participants had to fixate on a cross, located at the centre of the screen, to get a constant visual stimulation.

Each trial consisted of a sequence of pure tones, qualified as reference, targets and distractors (Fig. 1a). Three *reference* tones defining the beat of the sequence preceded a mixture of on- (*target*) and off-beat (*distractor*) tones. Participants performed a beat discrimination task at the end of each trial, by deciding whether the last tone of the sequence, a spectral deviant (785 Hz vs. 660 Hz), was on or off beat. Tones frequencies were selected to avoid potential bones transmission. The beat varied across conditions (8 conditions, with beats of 0.6, 0.7, 1, 1.3, 1.7, 2.2, 2.9 and 3.8Hz) to span the entire range of perceivable beats^{45,46}. In experiments 1 and 2, tones lasted 10 % of the inter-stimulus interval (ISI; e.g. for a beat of 1 Hz, the ISI would be of 1000 ms and the tones would last 100 ms). In experiment 6, we orthogonalized across conditions beat and tone durations, by fixing across conditions tones duration to 22.5 ms. Tones dampening length was of 10 % of their duration and tones attenuation was of 40 dB. Trials had pseudo-random durations (~2 to 10 s) but included at least four targets and lasted at least 2 seconds. These constraints were chosen to enable the deployment of temporal attention in all conditions. The density of distractors per sequence (i.e. the number of distractors per beat) was titrated individually

(see below). Distractors appeared randomly between targets with the constraint that all ISI within the sequence should be of at least 9 % of the beat period (e.g. ISI > 90ms for a 1 Hz beat).

In all auditory experiments, participants performed a *passive* session, in which they executed the task while staying completely still during the duration of the trial, not moving any part of their body. In experiment 2, additionally, participants performed a 'tracking' session, in which they were required to follow the beat by moving their (left or right, at their convenience) index finger on a noiseless pad from the beginning of the sequence (the 2nd reference tone). The pad was home-made and included a microphone connected to a Focusrite Saffire Pro24 sound card to record participants movements. In essence, the tracking session is a variation of the synchronization-continuation paradigm ³¹.

Each participant started the experiment with a short training session. The beat was fixed to 2 Hz and the density of distractors was at first equal to zero and increased progressively up to 0.4 distractor per beat. Participants were instructed not to move during the trials (as in the passive session). Then, participants listened to the 8 conditions at least 1 time each. Following this short training session, participants performed a psychophysical staircase where the density of distractors was the varying parameter. The staircase was set to obtain 75 % of categorization performance. Each experiment was divided into multiple sessions, each lasting around 1 hour.

Participants performed 40 trials per condition per session. In experiments 1, 2 and 6, participants performed 2, 1 and 1 passive session, corresponding to 640, 320 and 320 trials, respectively. In experiment 2, they also performed 1 tracking session (320 trials). Conditions (beats) were pseudo-randomly alternating in blocs of 20 trials each. Feedback was provided after each trial to indicate correct/incorrect responses, and more general performance feedback indicating the total number of correct responses was given after every bloc, for motivational purposes.

Experimental design of the visual experiments (n°4 and 5). These experiments are the transposition of experiments 1 and 2, respectively, to the visual modality. Each trial consisted of a sequence of centred visual gratings (visual extent 5°; Fig. 3a). Visual stimuli were sampled at 60 Hz. To impose a constant auditory stimulation on each trial, participants were presented with

auditory pink noise binaurally at a comfortable hearing level via headphones. Participants performed a beat discrimination task at the end of each trial, by deciding whether the last grating of the sequence, a colour deviant (blue vs. yellow), was on or off beat. In experiment 4, 10 conditions were investigated, with beats of 0.3, 0.4, 0.6, 0.7, 1, 1.3, 1.7, 2.2, 2.9 and 3.8 Hz. Compared to the auditory experiments, two extra conditions (corresponding to beats 0.3 and 0.4 Hz) were included, after pilot experiments, as it appeared that the optimal beat was of lower range in the visual than auditory modality. In experiment 5, only 8 conditions were presented (for time constraints issues), corresponding to beats of 0.4, 0.7, 1, 1.3, 1.7, 2.2, 2.9 and 3.8 Hz. Gratings duration was longer than tones duration, to avoid presenting subliminal stimuli, and lasted 18 % of the inter-stimulus interval (ISI; e.g. for a beat of 1 Hz, the gratings would last 180 ms). Participants performed 40 trials per condition per session. In experiments 4 and 5, participants performed 3 and 1 passive session, corresponding to 880 and 320 trials, respectively. In experiment 5, they also performed 1 tracking session (320 trials).

Free tapping experiment (n° 3). This experiment correspond to a subset of BASTAA (Battery for the Assessment of Auditory Sensorimotor and Timing Abilities ⁴⁷. To assess participants' spontaneous tapping rate and motor variability without a pacing stimulus, participants were asked to tap regularly at a comfortable rate for 60 seconds, the only instruction being to maintain the tapping rate as constant as possible. In two additional conditions, participants were instructed to tap as fast and as slow as possible, for 30 and 60 seconds, respectively. Participants were required to tap with their index finger on the noiseless home-made pad.

Timing of motor acts in tracking sessions. To investigate the ability of participants to actively follow the beat, we extracted the timing of individual motor acts in the tracking sessions. For each trial, we computed the mean and standard deviation of the inter-tap intervals. We then derived the tapping precision, expressed as the ratio between the average tapping frequency and the beat, and the coefficient of variation (CV), expressed as the relative standard deviation, i.e. the ratio between the standard deviation of the inter-tap intervals and the beat. For each trial, we also estimated the temporal distance between each individual motor acts and the beat, which indexes the sensorimotor simultaneity, i.e. the degree of simultaneity of motor acts relative to the beat. This

temporal distance was then normalized relative to the beat period (in a 2π space), with zero corresponding to perfect simultaneity between a motor act and the beat. We then either derived a relative or absolute sensorimotor simultaneity index, which respectively allows to estimate if participants tended to tap in anticipation or reaction to the beat, or to quantify the degree of sensorimotor simultaneity.

Estimation of an optimal beat. To estimate the beat at which performance (/CV) would be maximal (/minimal), variations of performance (/CV) across conditions were approximated with a third order polynomial function $f(x) = ax^3+bx^2+cx+d$, and the coordinates of the local maxima α (/minima β) were extracted according to the functions:

$$\alpha = \frac{-b - \sqrt{\delta}}{3a}; \beta = \frac{-b + \sqrt{\delta}}{3a}$$

where $\delta = b^2-3ac$ with $\delta > 0$. We used a third order polynomial function, as it is the best (ie most flexible) model that allows estimating one maximum without ambiguity (higher order models accept multiple maxima).

Statistical procedures. All analyses were performed at the single-subject level and followed by standard parametric two-sided tests at the group level (repeated-measures ANOVAs, paired and unpaired t-tests, Spearman correlations). For no compliance of Fisher's test, we used two by two t-test Welsh correction. When necessary, to provide an unbiased decision criterion with regards to the null hypothesis, we additionally used Bayesian statistics to derive a Bayes factor. We used a standard approach to compute the Bayes factors between "null" and "effect" hypotheses at the population level using the Akaike Information Criterion ⁸⁹. According to this symmetric hypothesis comparison measure, a Bayes factor of $< 1/3$ provides significant evidence in favour of the null hypothesis. Bayes factors were also computed for correlation coefficients ⁹⁰.

Model of coupled oscillators. We implemented a model of three coupled phase oscillators⁹¹ with time-delays and noise⁵⁰ to approximate the selective coupling between the external beat (S), sensory-specific periodic temporal attention (A) and motor tapping (M; Fig. 5a). The model was implemented with a set of differential equations, as:

$$\dot{\theta}_S = \omega_S,$$

$$\dot{\theta}_A = \omega_A + K_{AM} \sin[\theta_M(t - \tau_{AM}) - \theta_A] + K_{AS} \sin[\theta_S(t - \tau_{AS}) - \theta_A] + \xi_A(t),$$

$$\dot{\theta}_M = \omega_M + K_{MA} \sin[\theta_A(t - \tau_{MA}) - \theta_M] + K_{MS} \sin[\theta_S(t - \tau_{MS}) - \theta_M] + \xi_M(t),$$

where ω_i , θ_i and ξ_i are the natural frequency, phase and noise of an oscillator i , and for each pair of oscillators i and j , K_{ij} and τ_{ij} represent the coupling strength and time-delay from oscillator i to j . The noise ξ_i is additive and Gaussian with an intensity D , such as $\langle \xi_i(t) \rangle = 0$ and $\langle \xi_i(t) \xi_j(t') \rangle = 2D_{ij} \delta(t - t') \delta_{ij}$ (where $\langle \cdot \rangle$ denotes the time-average operator and δ the delta function). In line with studies implementing models of coupled oscillators, we ran the simulation during $1e4$ seconds, in order to obtain an equilibrium in the interaction between the coupled oscillators^{50,92}. The sampling rate of the simulation was 25 ms. We thus set internal time-delays to 0 ms (i.e., < 25 ms).

The level of coherence between the external beat S and the sensory-specific oscillator A was computed with the phase-locking value (PLV)^{50,51}. It estimates the capacity of A to entrain to S and is hence used as an approximation of behavioural performance. PLV is defined as:

$$PLV_{S-A} = \frac{1}{N} \sum_{t=1}^N e^{i\Delta\theta_{SA}(t)},$$

where the phase angle $\Delta\theta$ between oscillators S and A at time t is averaged across time points from $t = 1$ to N .

References

1. Nobre, A. C. & Van Ede, F. Anticipated moments: Temporal structure in attention. *Nat. Rev. Neurosci.* **19**, 34–48 (2018).
2. Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I. & Schroeder, C. E. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science (80-.)*. **320**, 110–113 (2008).
3. Morillon, B., Schroeder, C. E., Wyart, V. & Arnal, L. H. Temporal Prediction in lieu of Periodic Stimulation. *J. Neurosci.* **36**, 2342–2347 (2016).
4. Nobre, A. C. Orienting attention to instants in time. *Neuropsychologia* **39**, 1317–1328 (2001).
5. Cui, X., Stetson, C., Montague, P. R. & Eagleman, D. M. Ready...go: Amplitude of the fMRI signal encodes expectation of cue arrival time. *PLoS Biol.* **7**, (2009).
6. Lange, K. Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain Cogn.* **69**, 127–137 (2009).
7. Tal, I. *et al.* Neural Entrainment to the Beat: The “Missing-Pulse” Phenomenon. *J. Neurosci.* **37**, 6331–6341 (2017).
8. Teki, S., Grube, M., Kumar, S. & Griffiths, T. D. Distinct Neural Substrates of Duration-Based and Beat-Based Auditory Timing. *J. Neurosci.* **31**, 3805–3812 (2011).
9. Cope, T. E., Grube, M. & Griffiths, T. D. Temporal predictions based on a gradual change in tempo. *J. Acoust. Soc. Am.* **131**, 4013–4022 (2012).
10. Rimmele, J., Jolsvai, H. & Sussman, E. Auditory target detection is affected by implicit temporal and spatial expectations. *J. Cogn. Neurosci.* **23**, 1136–1147 (2011).
11. Arnal, L. H., Doelling, K. B. & Poeppel, D. Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cereb. Cortex* **25**, 3077–3085 (2015).
12. Di Luca, M. & Rhodes, D. Optimal Perceived Timing: Integrating Sensory Information with Dynamically Updated Expectations. *Sci. Rep.* **6**, 1–15 (2016).
13. Morillon, B., Schroeder, C. E. & Wyart, V. Motor contributions to the temporal precision of auditory attention. *Nat. Commun.* **5**, 1–9 (2014).
14. Morillon, B. & Baillet, S. Motor origin of temporal predictions in auditory attention. *Proc. Natl. Acad. Sci.* **114**, E8913–E8921 (2017).
15. Jones, M. R. Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychol. Rev.* **83**, 323–355 (1976).
16. Jones, M. R., Moynihan, H., MacKenzie, N. & Puente, J. Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychol. Sci.* **13**, 313–319 (2002).
17. Rhodes, D. & Di Luca, M. Temporal regularity of the environment drives time perception. *PLoS One* **11**, 1–18 (2016).
18. Moon, J., Choe, S., Lee, S. & Kwon, O. S. Temporal Dynamics of Visual Attention Allocation. *Sci. Rep.* **9**, 1–11 (2019).
19. Breska, A. & Deouell, L. Y. Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLoS Biol.* **15**, 1–30 (2017).
20. Cravo, A. M., Rohenkohl, G., Wyart, V. & Nobre, A. C. Temporal Expectation Enhances Contrast Sensitivity by Phase Entrainment of Low-Frequency Oscillations in Visual Cortex. *J. Neurosci.* **33**, 4002–4010 (2013).
21. ten Oever, S., Schroeder, C. E., Poeppel, D., van Atteveldt, N. & Zion-Golombic, E. Rhythmicity and cross-modal temporal cues facilitate detection. *Neuropsychologia* **63**, 43–50 (2014).
22. Besle, J. *et al.* Tuning of the Human Neocortex to the Temporal Dynamics of Attended Events. *J. Neurosci.* **31**, 3176–3185 (2011).
23. Bulthé, J., De Smedt, B. & Op de Beeck, H. P. Visual Number Beats Abstract Numerical Magnitude. *J. Cogn. Neurosci.* 1–10 (2015). doi:10.1162/jocn
24. McAuley, J. D. & Fromboloti, E. K. Attentional entrainment and perceived event duration. *Philos. Trans. R. Soc. B Biol. Sci.* **369**, (2014).
25. Auksztulewicz, R. *et al.* Not All Predictions Are Equal: “What” and “When” Predictions Modulate Activity in Auditory Cortex through Different Mechanisms. *J. Neurosci.* **38**, 8680–8693 (2018).
26. Doherty, J. R. Synergistic Effect of Combined Temporal and Spatial Expectations on Visual Attention. *J. Neurosci.* **25**, 8259–8266 (2005).

27. Rajendran, V. G., Teki, S. & Schnupp, J. W. H. Temporal Processing in Audition: Insights from Music. *Neuroscience* **389**, 4–18 (2018).
28. Ding, N. *et al.* Temporal modulations in speech and music. *Neurosci. Biobehav. Rev.* **81**, 181–187 (2017).
29. MacDougall, H. G. & Moore, T. S. Marching to the beat of the same drummer: the spontaneous tempo of human locomotion. *J. Appl. Physiol.* **99**, 1164–1173 (2005).
30. McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M. & Miller, N. S. The time of our lives: Life span development of timing and event tracking. *J. Exp. Psychol. Gen.* **135**, 348–367 (2006).
31. Repp, B. H. & Su, Y. Sensorimotor synchronization : A review of recent research (2006 – 2012). (2013). doi:10.3758/s13423-012-0371-2
32. Moelants, D. Preferred tempo reconsidered. *Proc. 7th Int. Conf. Music Percept. Cogn.* 580–583 (2002). doi:10.1080/00140138208924946
33. Morillon, B., Arnal, L. H., Schroeder, C. E. & Keitel, A. Prominence of delta oscillatory rhythms in the motor cortex and their relevance for auditory and speech perception. *Neurosci. Biobehav. Rev.* (2019).
34. Chartier, J., Anumanchipalli, G. K., Johnson, K. & Chang, E. F. Encoding of Articulatory Kinematic Trajectories in Human Speech Sensorimotor Cortex. *Neuron* **98**, 1042-1054.e4 (2018).
35. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51–6 (2012).
36. Hall, T. M., deCarvalho, F. & Jackson, A. A Common Structure Underlies Low-Frequency Cortical Dynamics in Movement, Sleep, and Sedation. *Neuron* **83**, 1185–1199 (2014).
37. Sussillo, D., Churchland, M. M., Kaufman, M. T. & Shenoy, K. V. A neural network that finds a naturalistic solution for the production of muscle activity. *Nat. Neurosci.* **18**, 1025–1033 (2015).
38. Keitel, A. & Gross, J. Individual Human Brain Areas Can Be Identified from Their Characteristic Spectral Activation Fingerprints. *PLoS Biol.* **14**, 1–22 (2016).
39. Hall, T. M., deCarvalho, F. & Jackson, A. A Common Structure Underlies Low-Frequency Cortical Dynamics in Movement, Sleep, and Sedation. *Neuron* **83**, 1185–1199 (2014).
40. Hatsopoulos, N. G. & Suminski, A. J. Sensing with the motor cortex. *Neuron* **72**, 477–487 (2011).
41. Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H. & Lakatos, P. Dynamics of Active Sensing and perceptual selection. *Curr. Opin. Neurobiol.* **20**, 172–176 (2010).
42. Wachowiak, M. All in a Sniff: Olfaction as a Model for Active Sensing. *Neuron* **71**, 962–973 (2011).
43. Fiebelkorn, I. C. & Kastner, S. A Rhythmic Theory of Attention. *Trends Cogn. Sci.* **23**, 87–101 (2019).
44. Re, D., Inbar, M., Richter, C. G. & Landau, A. N. Feature-Based Attention Samples Stimuli Report Feature-Based Attention Samples Stimuli Rhythmically. *Curr. Biol.* 1–7 (2019). doi:10.1016/j.cub.2019.01.010
45. Woodrow, H. Time Perception. *A Handb. Exp. Psychol.* (1951).
46. Fraisse, P. Les Structures Rhythmiques. *Erasmus Paris* (1956).
47. Bella, S. D. *et al.* BAASTA : Battery for the Assessment of Auditory Sensorimotor and Timing Abilities. 1128–1145 (2017). doi:10.3758/s13428-016-0773-6
48. Mishima, J. On the factors of the mental tempo. *Jpn. Psychol. Res.* **1956**, 27–38 (1956).
49. Arkady, P., Rosenblum, M. & Kurths, J. Synchronization A universal concept in nonlinear sciences. *Cambridge Univ. Press* 2–6 (2001).
50. Petkoski, S., Palva, J. M. & Jirsa, V. K. Phase-lags in large scale brain synchronization: Methodological considerations and in-silico analysis. *PLoS Comput. Biol.* **14**, 1–30 (2018).
51. Lachaux, J.-P., Rodriguez, E., Martinerie, J. & Varela, F. J. Measuring Phase Synchrony in Brain Signals. *Hum. Brain Mapp.* **52**, 562–569 (1999).
52. Coull, J. T. & Nobre, A. C. Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *J. Neurosci.* **18**, 7426–7435 (1998).
53. Fiebelkorn, I. C., Foxe, J. J., Butler, J. S. & Molholm, S. Auditory facilitation of visual-target detection persists regardless of retinal eccentricity and despite wide audiovisual misalignments. *Exp. Brain Res.* **213**, 167–174 (2011).
54. Fiebelkorn, I. C. & Kastner, S. A Rhythmic Theory of Attention. *Trends Cogn. Sci.* **23**, 87–101 (2019).
55. VanRullen, R. & Dubois, J. The psychophysics of brain rhythms. *Front. Psychol.* **2**, 1–10 (2011).
56. VanRullen, R., Zoefel, B. & Ilhan, B. On the cyclic nature of perception in vision versus audition. *Philos. Trans. R.*

- Soc. B Biol. Sci.* **369**, (2014).
57. Landau, A. N. & Fries, P. Attention samples stimuli rhythmically. *Curr. Biol.* **22**, 1000–1004 (2012).
 58. Song, K., Meng, M., Chen, L., Zhou, K. & Luo, H. Behavioral Oscillations in Attention: Rhythmic Pulses Mediated through Band. *J. Neurosci.* **34**, 4837–4844 (2014).
 59. Buschman, T. J. & Kastner, S. From Behavior to Neural Dynamics: An Integrated Theory of Attention. *Neuron* **88**, 127–144 (2015).
 60. VanRullen, R. Perceptual Cycles. *Trends Cogn. Sci.* **20**, 723–735 (2016).
 61. Saskia Haegens, E. Z. G. Rhythmic facilitation of sensory processing: A critical review. *Neurosci. Biobehav. Rev.* **1791**, 524–539 (2018).
 62. Helfrich, R. F. *et al.* Neural Mechanisms. *Neuron* **99**, 854–865.e5 (2018).
 63. Womelsdorf, T. & Everling, S. Long-Range Attention Networks: Circuit Motifs Underlying Endogenously Controlled Stimulus Selection. *Trends Neurosci.* **38**, 682–700 (2015).
 64. Merchant, H. & Yarrow, K. How the motor system both encodes and influences our sense of time. *Curr. Opin. Behav. Sci.* **8**, 22–27 (2016).
 65. Coull, J. T., Cheng, R. K. & Meck, W. H. Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology* **36**, 3–25 (2011).
 66. Teki, S., Grube, M. & Griffiths, T. D. A unified model of time perception accounts for duration-based and beat-based timing. *Front. Integr. Neurosci.* **5**, 1–7 (2012).
 67. Kotz, S. A., Brown, R. M. & Schwartze, M. Cortico-striatal circuits and the timing of action and perception. *Curr. Opin. Behav. Sci.* **8**, 42–45 (2016).
 68. Saleh, M., Reimer, J., Penn, R., Ojakangas, C. L. & Hatsopoulos, N. G. Fast and Slow Oscillations in Human Primary Motor Cortex Predict Oncoming Behaviorally Relevant Cues. *Neuron* **65**, 461–471 (2010).
 69. Arnal, L. H., Doelling, K. B. & Poeppel, D. Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cereb. Cortex* **25**, 3077–3085 (2015).
 70. Arnal, L. H. & Kleinschmidt, A. K. Entrained delta oscillations reflect the subjective tracking of time. *Commun. Integr. Biol.* **0889**, (2017).
 71. Stefanics, G. *et al.* Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *J. Neurosci.* **30**, 13578–13585 (2010).
 72. Iversen, J. R., Patel, A. D., Nicodemus, B. & Emmorey, K. Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition* **134**, 232–244 (2015).
 73. Hove, M. J., Iversen, J. R., Zhang, A. & Repp, B. H. Synchronization with competing visual and auditory rhythms: Bouncing ball meets metronome. *Psychol. Res.* **77**, 388–398 (2013).
 74. Hove, M. J., Spivey, M. J. & Krumhansl, C. L. Compatibility of Motion Facilitates Visuomotor Synchronization. *J. Exp. Psychol. Hum. Percept. Perform.* **36**, 1525–1534 (2010).
 75. Casasanto, D. & Boroditsky, L. Time in the mind: Using space to think about time. *Cognition* **106**, 579–593 (2008).
 76. Fraisse, P. II. - Rythmes auditifs et rythmes visuels. *Annee. Psychol.* **49**, 21–42 (1948).
 77. Dunlap, K. Reaction to rhythmic stimuli with attempt to synchronize. *Psychol. Rev.* **17**, 399–416 (1910).
 78. Burr, D., Banks, M. S. & Morrone, M. C. Auditory dominance over vision in the perception of interval duration. *Exp. Brain Res.* **198**, 49–57 (2009).
 79. Patel, A. D., Iversen, J. R., Chen, Y. & Repp, B. H. The influence of metricality and modality on synchronization with a beat. *Exp. Brain Res.* **163**, 226–238 (2005).
 80. Comstock, D. C., Hove, M. J. & Balasubramaniam, R. Sensorimotor Synchronization With Auditory and Visual Modalities: Behavioral and Neural Differences. *Front. Comput. Neurosci.* **12**, 1–8 (2018).
 81. Siegel, M., Donner, T. H. & Engel, A. K. Spectral fingerprints of large-scale neuronal interactions. *Nat. Rev. Neurosci.* **13**, 121–134 (2012).
 82. Donner, T. H. & Siegel, M. A framework for local cortical oscillation patterns. *Trends Cogn. Sci.* **15**, 191–199 (2011).
 83. Miall, R. C., Weir, D. J., Stein, J. F. & Road, P. MANUAL TRACKING OF VISUAL TARGETS BY TRAINED MONKEYS R.C. *Behav. Brain Res.* **20**, 185–201 (1986).

84. Rimmele, J. M., Morillon, B., Poeppel, D. & Arnal, L. H. Proactive Sensing of Periodic and Aperiodic Auditory Patterns. *Trends Cogn. Sci.* **22**, 870–882 (2018).
85. Correa, Á., Cona, G., Arbula, S., Vallesi, A. & Bisiacchi, P. Neural dissociation of automatic and controlled temporal preparation by transcranial magnetic stimulation. *Neuropsychologia* **65**, 131–136 (2014).
86. Ren, Y. *et al.* Does Temporal Expectation Driven by Rhythmic Cues Differ From That Driven by Symbolic Cues Across the Millisecond and Second Range? *Perception* **48**, 515–529 (2019).
87. Breska, A. & Deouell, L. Y. Automatic Bias of Temporal Expectations following Temporally Regular Input Independently of High-level Temporal Expectation. *J. Cogn. Neurosci.* 1–10 (2014). doi:10.1162/jocn
88. Kleiner, M. *et al.* What 's new in Psychtoolbox-3 ? (2007).
89. Adrian E. Raftery & Kass Robert E. Bayes Factor. (1995).
90. Wetzels, R. & Wagenmakers, E. A default Bayesian hypothesis test for correlations and partial correlations. 1057–1064 (2012). doi:10.3758/s13423-012-0295-x
91. Kuramoto, Y. SELF-ENTRAINMENT OF A POPULATION OF COUPLED NON-LINEAR OSCILLATORS. 2548 (1984).
92. Asllani, M., Expert, P. & Carletti, T. A minimally invasive neurostimulation method for controlling abnormal synchronisation in the neuronal activity. *PLoS Comput. Biol.* **14**, 1–18 (2018).

Acknowledgments: We thank Viktor Jirsa and Jennifer Coull for their valuable insights into this manuscript. Research supported by grants ANR-16-CONV-0002 (ILCB), ANR-11-LABX-0036 (BLRI) and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

Author contributions: A.Z. and B.M. designed the experiments; A.Z. acquired data; A.Z. and B.M. analysed data; S.P. implemented the model; and A.Z. and B.M. wrote the manuscript.

Competing interests: The authors declare no competing interests.

Figures

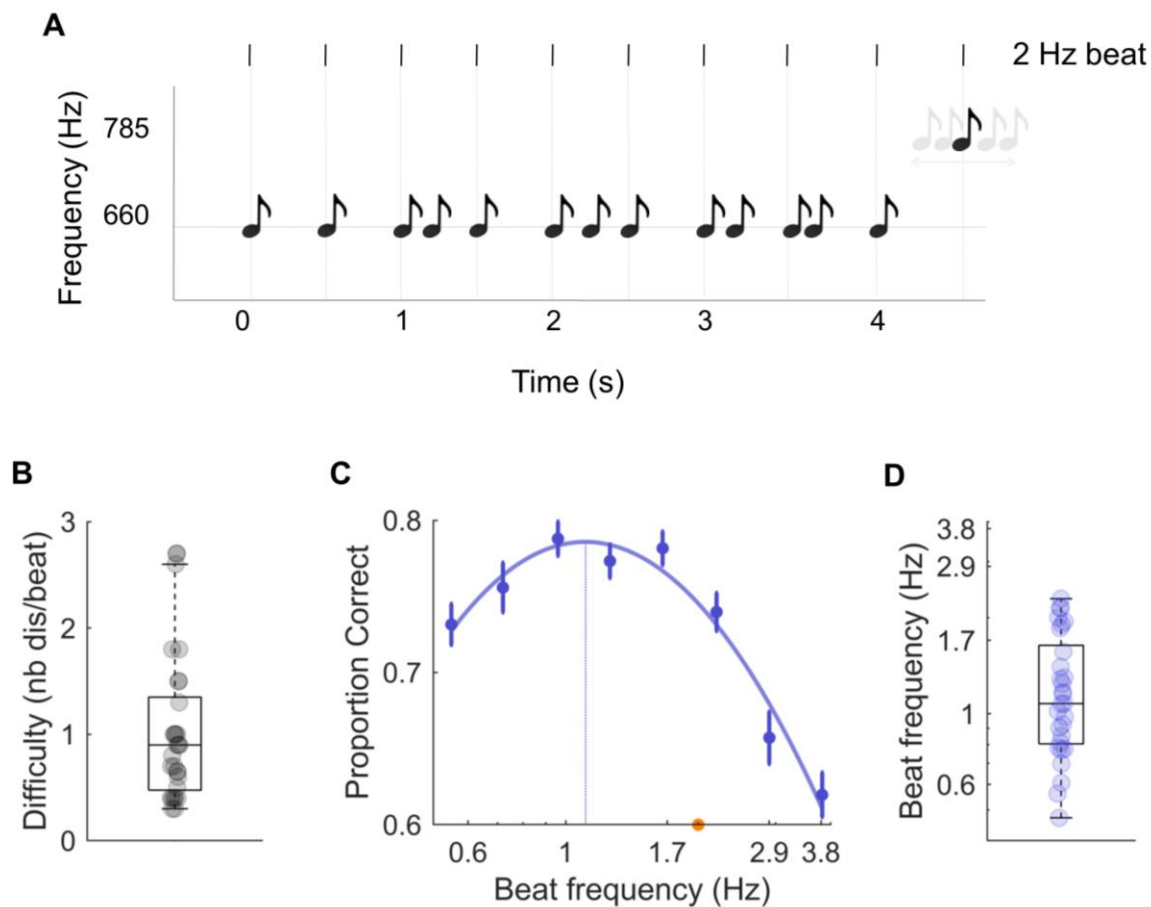


Figure 1 | Experimental design and results of the auditory periodic temporal attention

experiment. Experiment 1. **A.** Sequences of pure tones were presented binaurally on each trial. Three reference tones defining the beat of the sequence preceded a mixture of on- and off-beat tones. Participants performed a beat discrimination task at the end of each trial, by deciding whether the last tone of the sequence, a spectral deviant (785 Hz vs. 660 Hz), was on or off beat. The beat varied across conditions (from 0.6 Hz to 3.8 Hz). **B.** Individual difficulty level to reach threshold performance for a 2 Hz beat. Difficulty was modulated by adjusting the density of distractors in the sequence (number of distractors per beat). **C.** Average performance per condition (beat). Data were approximated with a polynomial function (plain line), and an optimal beat (leading to a maximal performance) could be estimated (dashed line). The orange dot indicates the beat (2 Hz) that served to estimate the individual difficulty level. **D.** Individual estimates of the optimal beat. Error bars indicate s.e.m. ($n = 30$).

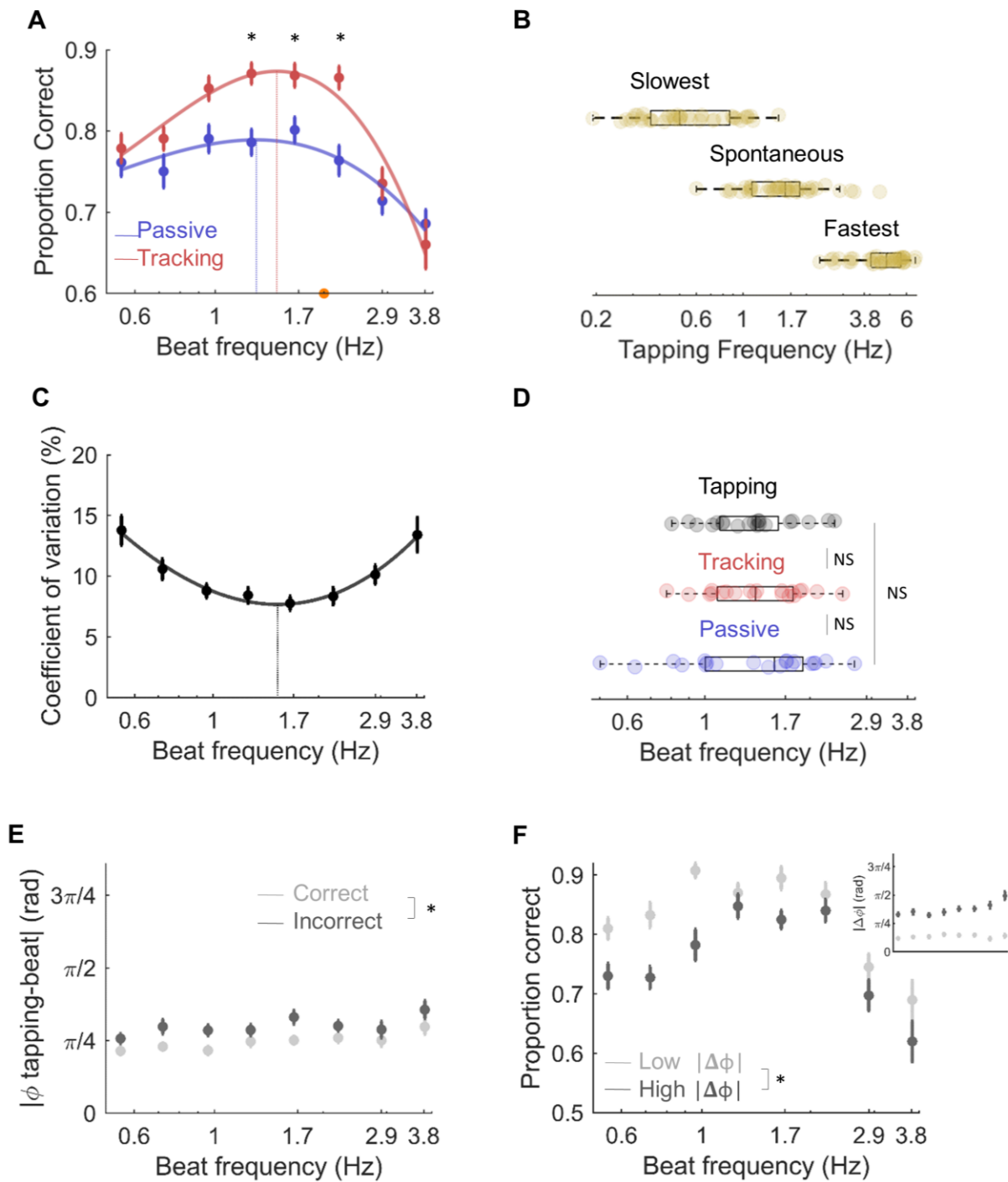


Figure 2 | Motor-tracking influence on auditory attention and tapping characteristics. A.

Experiment 2. Average performance per condition (beat) in the passive (blue) and tracking (red) sessions. In the passive session, participants performed the task without moving before the end of the sequence. In the tracking session, participants performed the task while expressing the beat by moving their index finger. Same conventions as in Figure 1C. **B.** Free tapping experiment (exp. 3). Individual average frequency of tapping in three conditions where participants had to rhythmically

tap at their slowest, spontaneous and fastest rate. **C-F. Experiment 2:** **C.** Coefficient of variation (CV; i.e. relative standard deviation) of tapping across conditions in the tracking session. Same conventions as in Figure 1C. **D.** Individual estimates of the optimal beat of auditory temporal attention in the passive (blue) and tracking (red) sessions (from A) and of motor tapping in the tracking session (black; from C). **E.** Sensorimotor simultaneity (Φ ; in radian), i.e. temporal distance between motor acts and the beat (in absolute value, normalized to the beat period) across conditions in the tracking session, for correct (light grey) and incorrect (dark grey) trials. **F.** Average performance per condition in the tracking session for trials with low (light grey) and high (dark grey) sensorimotor simultaneity indexes. Trials were sorted according to a median-split procedure. The inset plot indicates the associated sensorimotor simultaneity indexes. Error bars indicate s.e.m. ($n = 20$).

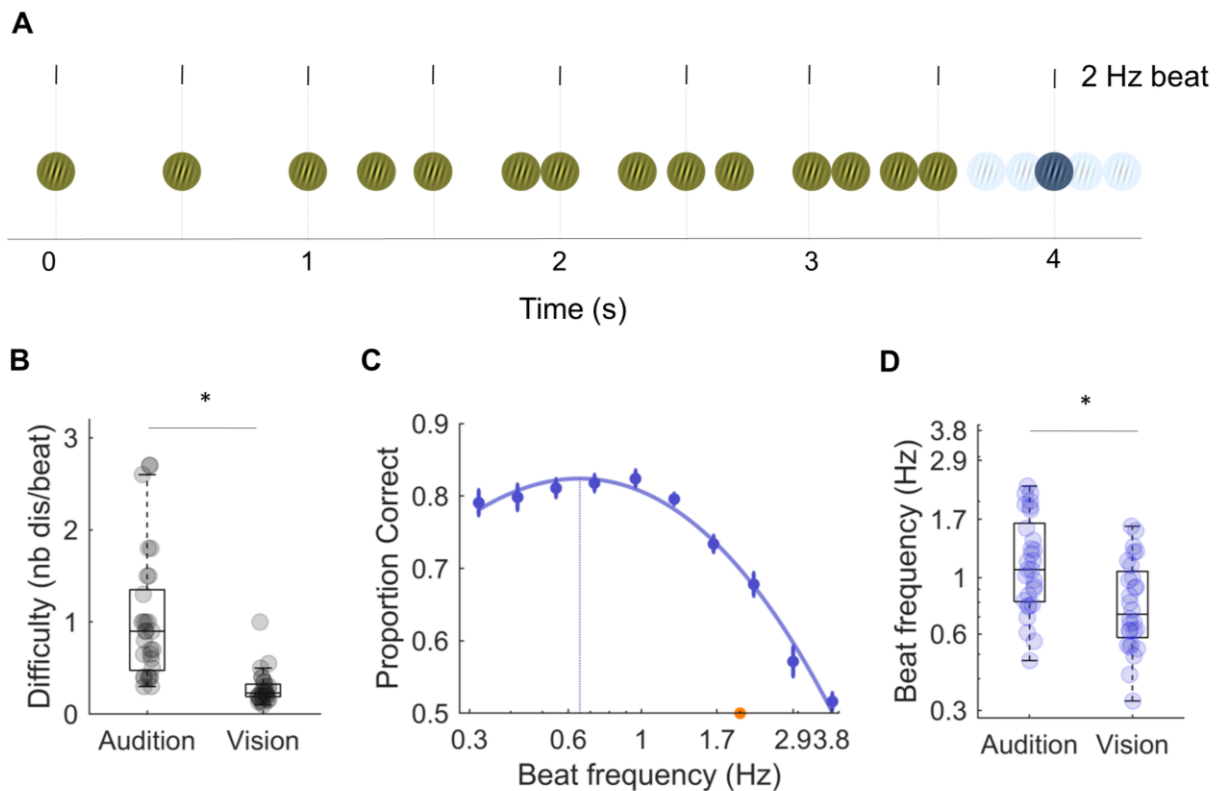


Figure 3 | Experimental design and results of the visual periodic temporal attention

experiment. Experiment 4. **A.** In contrast to auditory experiment 1, sequences of visual gratings were presented on each trial. Participants performed a beat discrimination task by deciding whether the last grating of the sequence, a colour deviant (blue vs. yellow), was on or off beat. The beat varied across conditions (from 0.3 Hz to 3.8 Hz) **B.** Comparison of the individual difficulty level to reach threshold performance for a 2 Hz beat, between auditory (exp. 1) and visual (exp. 4) experiments. **C.** Average performance per condition. Same conventions as in Figure 1C. **D.** Comparison of individual estimates of the optimal beat, between auditory (exp. 1) and visual (exp. 4) experiments. Error bars indicate s.e.m. Stars indicate significant differences ($n = 30$; unpaired t -tests; $p < .05$).

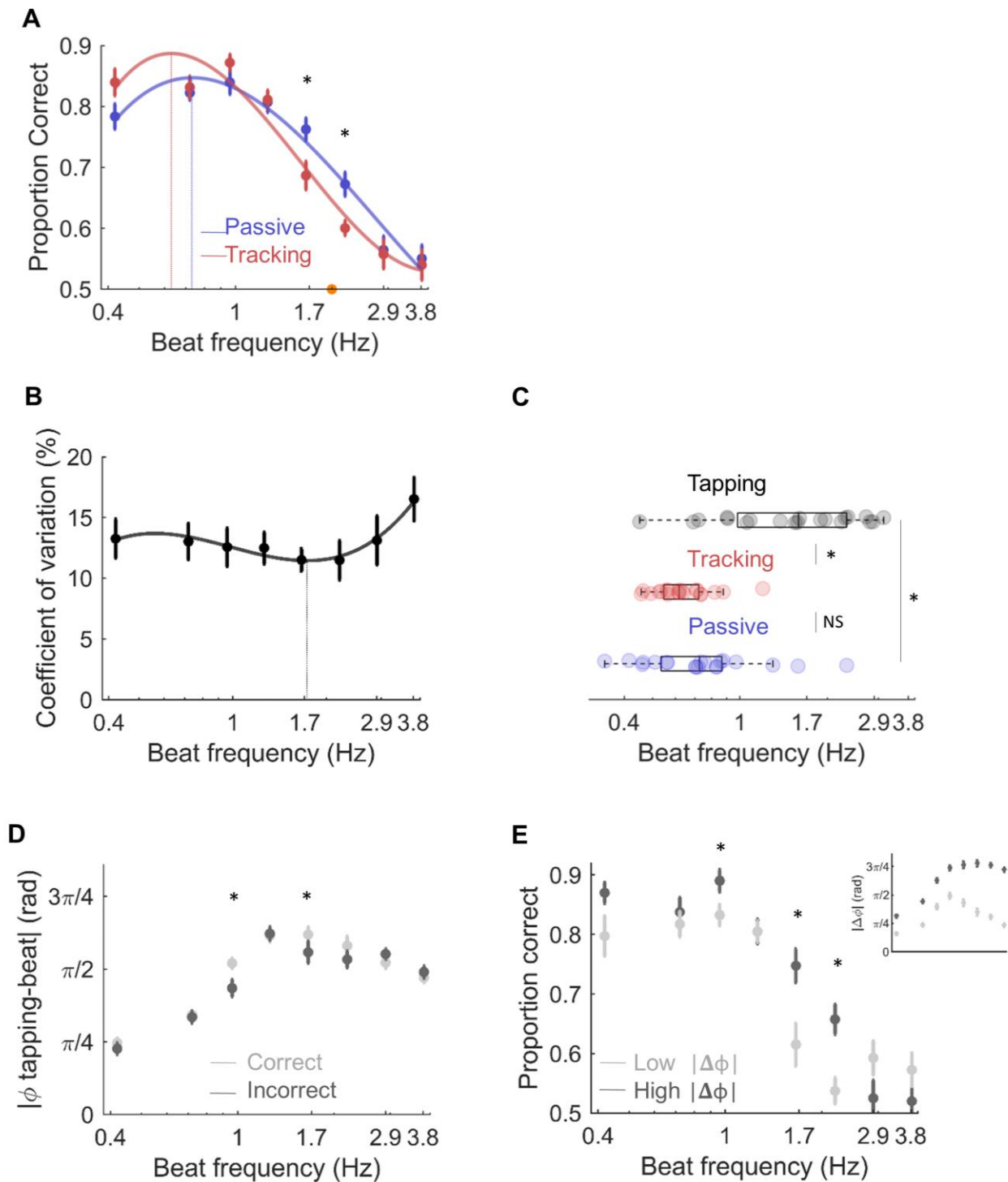


Figure 4 | Motor-tracking influence on visual attention and tapping characteristics.

Experiment 5. **A.** Average performance per condition in the passive (blue) and tracking (red) sessions. **B.** Coefficient of variation (CV) of tapping across conditions in the tracking session. **C.** Individual estimates of the optimal beat of visual temporal attention in the passive (blue) and tracking (red) sessions (from A) and of motor tapping in the tracking session (black; from B). **D.**

Sensorimotor simultaneity (Φ ; in radian) across conditions in the tracking session, for correct (light grey) and incorrect (dark grey) trials. **E.** Average performance per condition in the tracking session for trials with low (light grey) and high (dark grey) sensorimotor simultaneity indexes. The inset plot indicates the associated sensorimotor simultaneity indexes. Same conventions as in Figure 2 (n = 20).

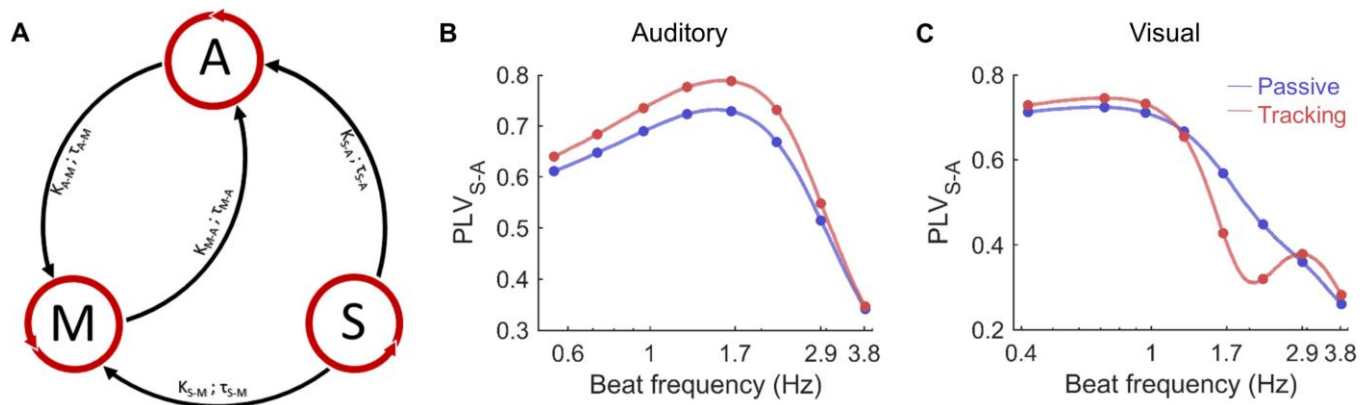
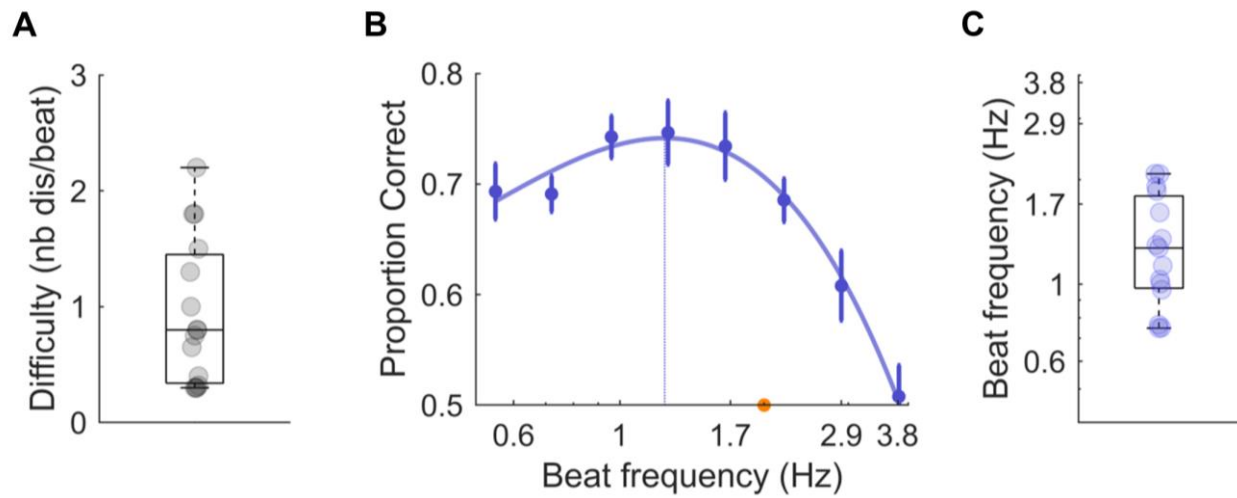
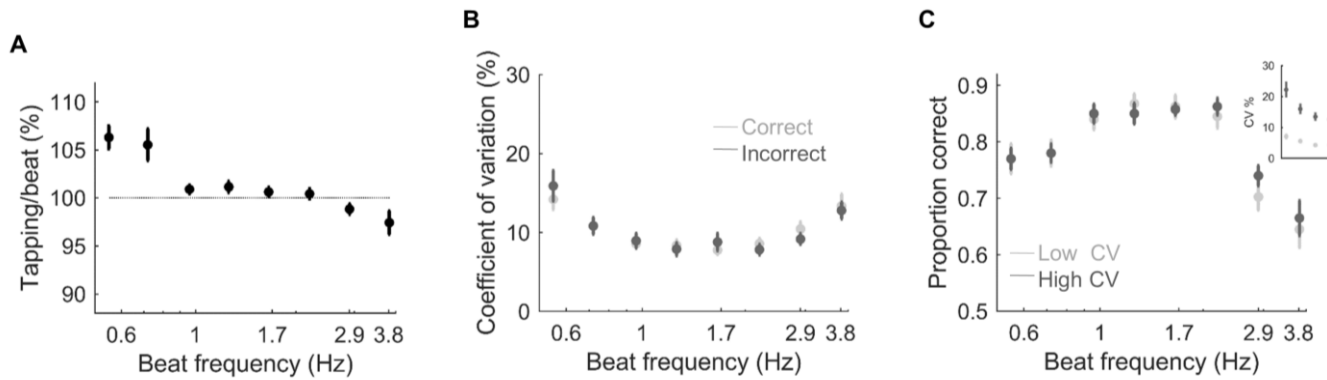


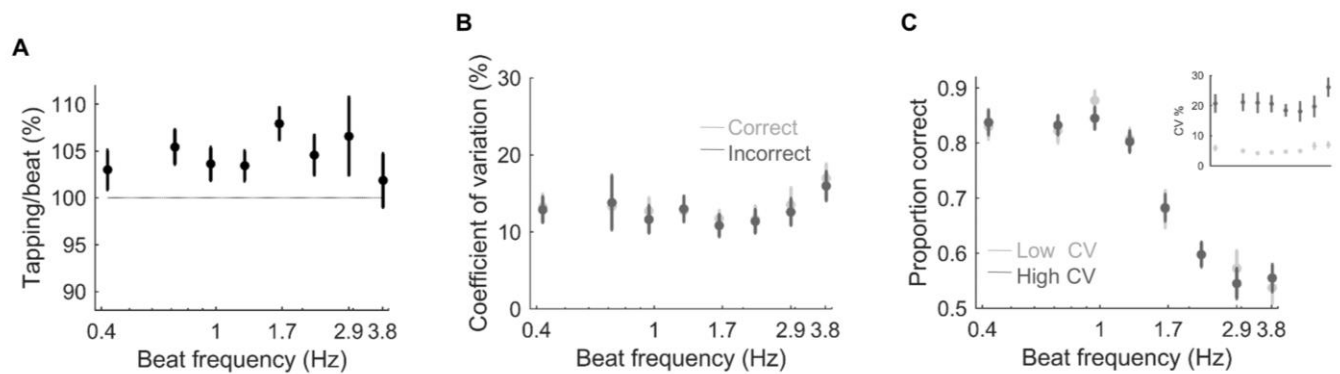
Figure 5 | Replication of the results with a model of coupled oscillators. A. Model of three delay-coupled phase oscillators approximating the selective coupling between the external beat (S), sensory-specific temporal attention (A), and natural motor dynamics (M). The external beat oscillator influences attention and motor oscillators with a specific strength K and delay τ , and attention and motor oscillators reciprocally influence each other. The phase-locking value (PLV) between the external beat (S) and the sensory-specific temporal attention oscillator (A) reflects the capacity of A to entrain to S and thus is used as an approximation of behavioural performance. **B-C.** Replication of the **(A)** auditory (exp. 2) and **(C)** visual (exp. 5) passive (blue) and tracking (red) experiments. Difference between conditions was obtained by adjusting three key parameters: the natural frequency of the sensory-specific temporal attention oscillator (A) and the time-delay between the stimulus (S) and the motor oscillator (M) which varied across modalities (auditory: $\omega_A = 1.5$ Hz, $\tau_{S-M} = 0.1$ s; visual: $\omega_A = 0.7$ Hz, $\tau_{S-M} = 0.35$ s), and the coupling strength between motor and attention oscillators, which varied between passive ($K_{M-A} = 2$) and tracking ($K_{M-A} = 10$) sessions.



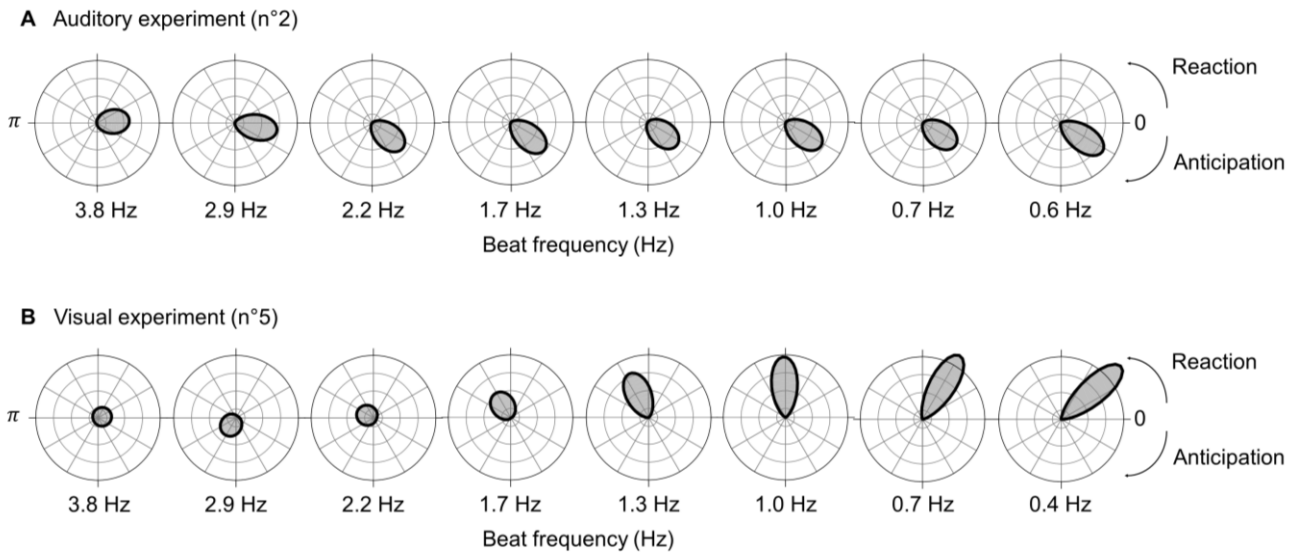
Supplementary figure 1 | Experiment 6. In this alternate version of experiment 1, the duration of the pure tones was kept constant across conditions and equal to 22.5ms. **A.** Individual difficulty level to reach threshold performance for a 2 Hz beat. **B.** Average performance per condition. Same conventions as in Figure 1C. **C.** Individual estimates of the optimal beat. Error bars indicate s.e.m. (n = 15).



Supplementary figure 2 | Experiment 2 (auditory). **A.** Tapping precision across conditions of the tracking session. The precision is expressed as the ratio between the average tapping frequency and the beat. The dashed line indicates the ideal ratio. **B.** Coefficient of variation (CV) of tapping across conditions in the tracking session, for correct (light grey) and incorrect (dark grey) trials. **C.** Average performance per condition in the tracking session, for trials with low (light grey) and high (dark grey) CV. Trials were sorted according to a median-split procedure. The inset plot indicates the associated CV. Error bars indicate s.e.m. ($n = 20$).



Supplementary figure 3 | Experiment 5 (visual). **A.** Tapping precision across conditions of the tracking session. **B.** Coefficient of variation (CV) of tapping across conditions in the tracking session, for correct (light grey) and incorrect (dark grey) trials. **C.** Average performance per condition in the tracking session, for trials with low (light grey) and high (dark grey) CV. The inset plot indicates the associated CV. Same conventions as in Figure S2.



Supplementary figure 4 | Inter-individual distribution of the temporal distance between motor acts and the beat (in *relative* value, normalized to the beat period) across conditions in the tracking session of the **(A)** auditory (exp. 2) and **(B)** visual (exp. 5) experiments. Negative (/positive) values indicate that motor acts anticipated (lagged) the beat (0: in-phase, π : antiphase).