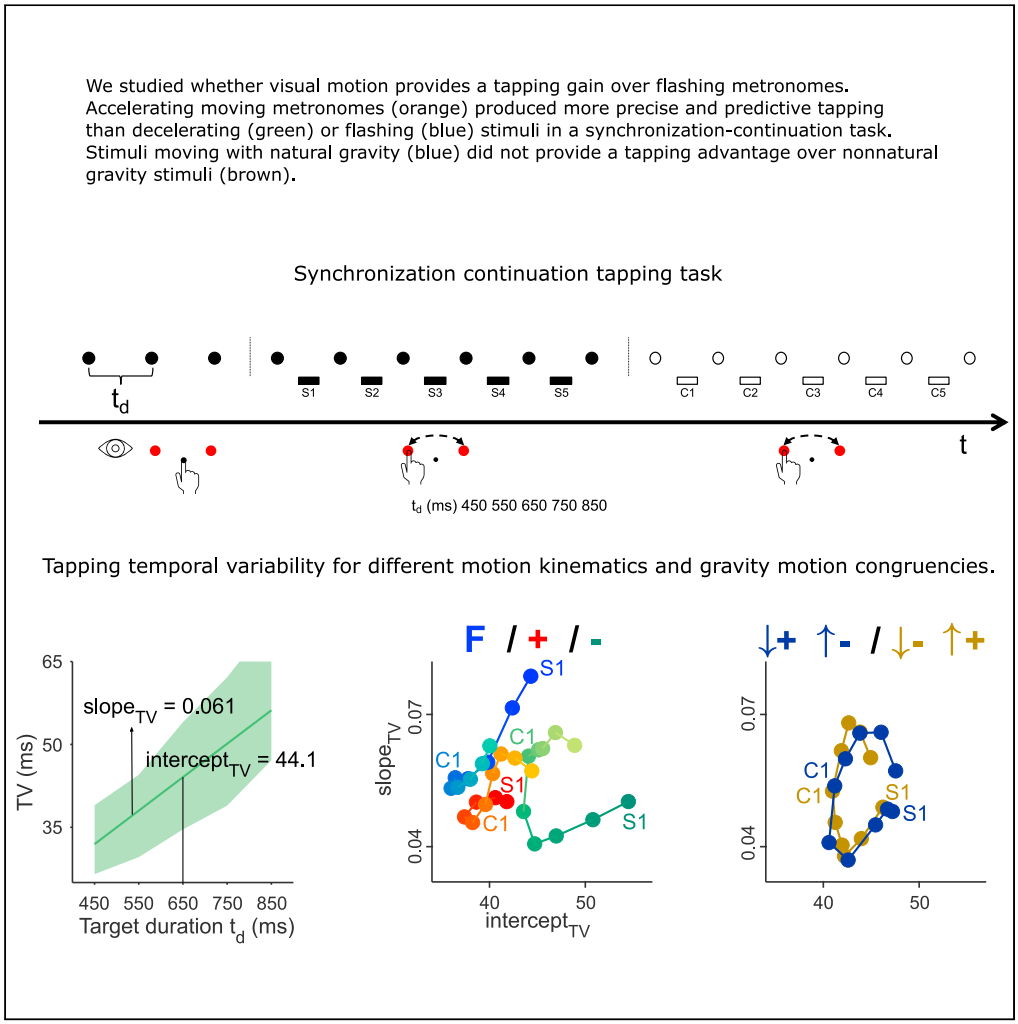
iScience

CellPress
OPEN ACCESS

Article

Rhythmic tapping to a moving beat motion kinematics overrules natural gravity



Oswaldo Pérez, Sergio Delle Monache, Francesco Lacquaniti, Gianfranco Bosco, Hugo Merchant

hugomerchant@unam.mx

Highlights

We studied whether visual motion provides a tapping gain over flashing metronomes

Accelerating moving metronomes produced more precise and predictive tapping

Stimuli moving with natural gravity did not provide a tapping advantage

Error correction or a prior effect was used depending on the tapping contingencies

Pérez et al., iScience 26, 107543
September 15, 2023
https://doi.org/10.1016/j.isci.2023.107543

iScience

CellPress
OPEN ACCESS

Article

Rhythmic tapping to a moving beat motion kinematics overrules natural gravity

Oswaldo Pérez,[1,8] Sergio Delle Monache,[2,3,8] Francesco Lacquaniti,[4,5,6] Gianfranco Bosco,[4,5,6] and Hugo Merchant[7,9,*]

SUMMARY

Beat induction is the cognitive ability that allows humans to listen to a regular pulse in music and move in synchrony with it. Although auditory rhythmic cues induce more consistent synchronization than flashing visual metronomes, this auditory-visual asymmetry can be canceled by visual moving stimuli. Here, we investigated whether the naturalness of visual motion or its kinematics could provide a synchronization advantage over flashing metronomes. Subjects were asked to tap in sync with visual metronomes defined by vertically accelerating/decelerating motion, either congruent or not with natural gravity; horizontally accelerating/decelerating motion; or flashing stimuli. We found that motion kinematics was the predominant factor determining rhythm synchronization, as accelerating moving metronomes in any cardinal direction produced more precise and predictive tapping than decelerating or flashing conditions. Our results support the notion that accelerating visual metronomes convey a strong sense of beat, as seen in the cueing movements of an orchestra director.

INTRODUCTION

Beat induction is the cognitive ability that allows humans to listen to a regular pulse in music and move in synchrony with it. Critically, without beat induction there is no music perception; hence, it is considered a universal human trait.[1–3] A rather useful paradigm to investigate beat induction is the synchronization-continuation task (SCT), in which subjects tap in sync with periodic sensory cues that generate an internal reference interval (synchronization epoch) and keep tapping after the metronome is extinguished using this internal beat representation (continuation epoch).[4,5] Performance in this task shows that the variability of produced intervals increases linearly with the mean, known as scalar property (a form of Weber's law)[6–8] and that shorter and longer intervals are over- and underestimated (termed bias effect or regression toward the mean).[9–12] In addition, subjects use an error correction mechanism that maintains tap synchronization with the metronome since a longer produced interval tends to be followed by a shorter interval and vice versa, to avoid error accumulation and losing the metronome.[13,14] In contrast, during continuation, there is a drift in the produced duration.[15,16] Functional imaging and neurophysiological studies have shown that beat induction and entrainment depend on the cortico-thalamic-striatal, involving the medial premotor cortex and the putamen.[17–21] A neural clock flexibly represents the internal beat in the cyclical resetting of cell population states within the medial premotor cortex.[22–26]

It is widely known that rhythmic tapping is more accurate and consistent when cued by auditory metronomes than flashing visual metronomes.[7,27–30] The superiority of audition across studies supports the notion that the auditory system is specialized for time processing, while vision is specialized for spatial processing.[31–33] Indeed, a recent hypothesis accentuates the role of the audiomotor system in beat perception and entrainment.[34–36] Nevertheless, visual moving metronomes can cancel out the auditory-visual asymmetry,[37] especially with naturalistic stimuli such as videos of a bouncing ball with changes in speed[14,38] or acceleration profiles that are congruent with the effects of gravity.[39] The hypothesis behind these observations is that visual motion may engage the time-to-collision mechanism in the parietal cortex. This mechanism is used for interception of moving targets or collision avoidance.[40–43] Thus, the parieto-premotor system recruited for encoding time to contact for single events could also be involved in the rhythm internalization of periodic collision points, efficiently driving the motor system for beat-based timing.[14,44,45] In other words, the visual parieto-premotor system could predict a visual beat as efficiently as the audiomotor system for auditory stimuli. Remarkably, the degree of congruency between visual motion and motor

[1]Escuela Nacional de Estudios Superiores Unidad Juriquilla, Universidad Nacional Autónoma de México, Boulevard Juriquilla No. 3001, Querétaro, Qro 76230, México

[2]Laboratory of Visuomotor Control and Gravitational Physiology, IRCCS Fondazione Santa Lucia, 00179 Rome, Italy

[3]Department of Civil Engineering and Computer Science Engineering, University of Rome Tor Vergata, 00133 Rome, Italy

[4]Department of Systems Medicine, University of Rome Tor Vergata, Rome, Italy

[5]Centre of Space Biomedicine, University of Rome "Tor Vergata", Rome, Italy

[6]Laboratory of Neuromotor Physiology, IRCCS Santa Lucia Foundation, Rome, Italy

[7]Instituto de Neurobiología, Universidad Nacional Autónoma de México, Campus Juriquilla, Boulevard Juriquilla No. 3001, Querétaro, Qro 76230, México

[8]These authors contributed equally

[9]Lead contact

*Correspondence: hugomerchant@unam.mx

https://doi.org/10.1016/j.isci.2023.107543

response can influence rhythm synchronization to visual motion stimuli, as tapping movements aligned with the direction of the visual motion can accomplish better synchronization than those incongruent with the visual motion direction.[37] These findings support the hypothesis that gravitational cues can facilitate some perceptual functions, like interpreting the causality and naturalness of object motion or discriminating pendular motion,[46–49] and can lead to an advantage in manual interceptions by engaging an internal representation of gravity in the vestibular cortex.[50–52] Interestingly, object kinematics displayed on a computer screen can engage this internal representation of gravity, especially if the pictorial background provides scaling cues to real-world metrics.[53–57]

No studies have used naturalistic videos to determine the effects of Earth's gravity on beat induction. Therefore, we designed an SCT cued by objects that were either vertically or horizontally accelerating over background images that provided elements of naturalness. By pairing upward and downward motion with accelerated and decelerated motion, vertical motion was either compatible or not with Earth's gravity effects. Thus, downward accelerations and upward decelerations are congruent with gravity effects, whereas downward decelerations and upward accelerations are not. In addition, we also used horizontal motion, where all combinations of motion direction (leftward or rightward) and acceleration (positive or negative) could be considered plausible. In principle, this experimental design can dissect the effects of the target kinematics per se from those of the target motion naturalness on the subjects' ability to reproduce the temporal intervals cued by the moving stimuli. Importantly, we developed a Bayesian observer model that fully explained the changes in time precision, accuracy, and correlation in the produced intervals in the tapping sequence. With this framework of statistical inference, we found that subjects used different strategies to optimally switch between the sensory-guided tap synchronization and the internally driven continuation, such as error correction and regression toward the mean, respectively. Overall, the present findings describe profound effects of the acceleration and deceleration profiles of moving metronomes on rhythmic tapping but minimal effects of natural gravity compared to those of arbitrary motion. These results are in line with the sharp acceleration profiles and abrupt changes in direction that conductors use to define the beat of an orchestra.

RESULTS

Sixteen subjects performed a modified version of the SCT (Figure 1, see STAR methods). Briefly, participants produced five intervals by tapping on a touchscreen in sync with visual cued stimuli alternating between two positions (synchronization epoch) and then continued tapping at the same tempo for another five intervals without the visual cues (continuation epoch) (Figure 1A). Tapping was also performed alternating between two positions in the bottom-right portion of the touchscreen. We used two stimuli and two response locations to generate an SCT that simulated natural situations, such as a drummer playing the bongos, congas, or timbales. In one variation of the SCT, eight different motion conditions were used to define the intervals: two vertical (up and down) and two horizontal (right and left) directions, each with either accelerating (1G) or decelerating (-1G) targets, rendered on two distinct quasi-realistic visual scenarios (Figures 1B and 1C). Temporal intervals were cued by two flanking objects moving with the same kinematics (motion duration = 750 ms) but shifted temporally so that the bouncing against flat surfaces defined the interval duration of the visual metronome. Subjects were instructed to tap in sync with the bouncing targets. As control conditions, we used flashing visual metronomes consisting of two flanking targets flashing alternately (150 ms). These targets were identical in appearance and location as those in the moving metronome conditions (Figures 1F and 1G). Five target interval durations (t_d from 450 to 850 ms in steps of 100 ms) across all visual motion and flashing conditions were used.

Precision, accuracy, and prediction of rhythmic timing

First, we assessed the effects of the various experimental manipulations of the moving and flashing metronomes on constant error (CE), temporal variability (TV), and asynchronies. CE is a measure of timing accuracy and corresponds to the difference between produced and target intervals. TV is a measure of timing precision and corresponds to the standard deviation of the produced intervals. Asynchronies are the time differences between stimuli and tap responses and represent measures of rhythmic prediction, which can be evaluated only during the SCT synchronization epoch.[7,24,58]

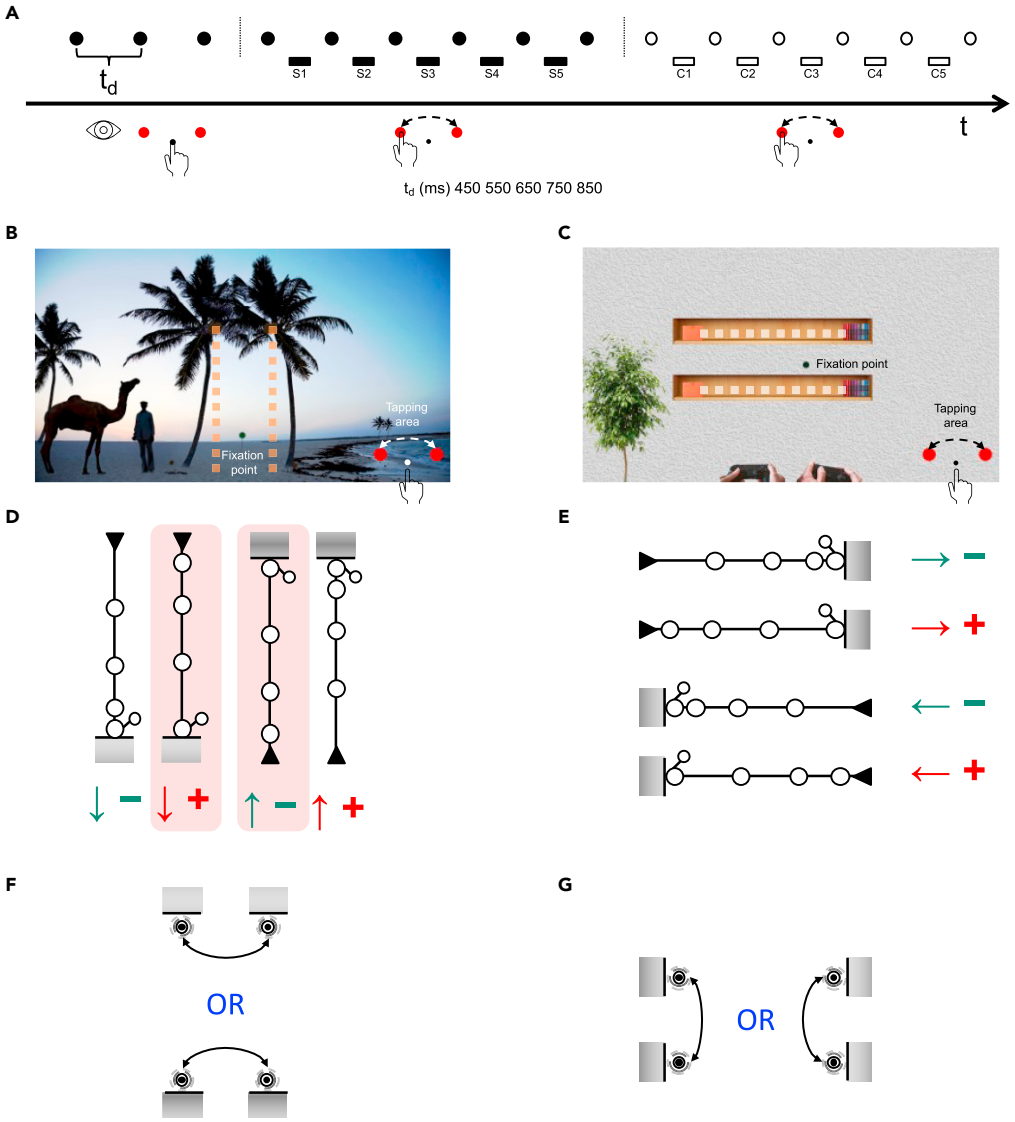


Figure 1. Synchronization-continuation task (SCT)

(A) Sequence of taps and definition of reproduced temporal intervals. A trial started when subjects placed their right index finger on the starting position at the center of the tapping area located at the bottom-right corner of the touchscreen (see also panels B and C). Then, isochronous stimuli were presented, and their motion alternated between two positions. After the first three instruction visual stimuli (observation epoch), the subjects produced six taps in synchrony with the metronome (synchronization epochs $S_{1\text{-}5}$ denote the reproduced intervals) by tapping first on the left red circle and then by alternating between the two red circles in the tapping area. After the visual metronome was extinguished, the subjects continued tapping at the same tempo for another six taps without any visual cue (continuation epochs $C_{1\text{-}5}$ denote the reproduced intervals). Five target interval durations were used (t_d from 450 to 850 ms in steps of 100 ms). During the task, subjects were required to maintain ocular fixation on designated points of the visual scene (see panels B and C).

(B) Vertical Scenario. Visual metronomes were presented in separate trials, either at the bottom or at the top of the palm trees, by alternating visual stimuli at fixed intervals between the right (first) and left trees. The dark green circle between the two palm trees indicates the ocular fixation point when the visual metronomes were presented at the bottom of the trees. The fixation point (not shown for clarity) presented at the top of the trees was located between the two palm trees, at the same visual angle distance as the one used for the bottom visual metronomes.

(C) Horizontal Scenario. Visual metronomes were presented either at the right or left ends of the two bookshelves. The dark green circle between the two shelves denotes the ocular fixation point when visual metronomes were presented at the right end of the shelves. The fixation point for the visual metronomes presented at the left end of the shelves (not

Figure 1. *Continued*

shown for clarity) was located between the bookshelves at the same visual angle distance from the visual metronomes at the right end of the shelves.

(D) Moving metronomes in the Vertical Scenario. Two coconuts were either dropped from both palm trees and bounced on the ground below the trees or were launched from the ground and bounced against the top branches of the palm trees. In each trial, the two coconuts moved in the same direction and with the same kinematics (motion duration = 750 ms), but their motion onsets were shifted temporally at fixed amounts of time so that alternate bounces between the right (occurring first) and the left coconut occurred at fixed intervals corresponding to one of the five possible interval durations of the visual metronome. For each direction of motion, coconuts either accelerated (open circles going further apart as they approach the bounce, coded as a red plus sign) or decelerated (open circles getting progressively closer as they approach the bounce, coded as a minus green sign). The velocity at the bounce was 59 and 5 visual degrees*sec^{-1} for accelerated and decelerated motion, respectively. After the bounce, coconuts moved at an oblique angle from the bouncing surface and disappeared quickly behind the trees at a velocity between 18 and 21 visual degrees*sec^{-1}. Vertical motion conditions congruent with natural gravity effects (i.e., downward accelerated and upward decelerated) are shaded in pink.

(E) Moving metronomes in the Horizontal Scenario. Two toy racing cars ran along the top and bottom bookshelves, starting at one end (symbolized by the black triangle) and bouncing against the pile of books (symbolized by the shaded rectangle) at the opposite end. As in the Vertical Scenario, in each trial, the two moving targets had the same kinematics (either accelerated or decelerated; motion duration = 750 ms), but their motion onsets were shifted temporally at fixed amounts of time so that the alternate bounces between the car on the bottom (occurring first) and top shelves defined one of the five possible interval durations of the visual metronome. The velocity at the bounce was 59 and 5 visual degrees*sec^{-1} for accelerated and decelerated motion, respectively. After the bounce, cars moved at an oblique angle from the bouncing surface and disappeared quickly behind the book piles at a velocity between 18 and 21 visual degrees*sec^{-1}.

(F) Flashing metronomes in the Vertical Scenario. In each trial, static coconut images were alternately flashed for 150 ms between the right (first) and left palm trees, at one of the five possible interval durations of the visual metronome. Alternated flashing of the coconuts occurred at the bouncing locations of the moving metronomes, either at the treetops or on the ground.

(G) Flashing metronomes in the Horizontal Scenario. In each trial, static toy car images were alternately flashed for 150 ms between the bottom (first) and top bookshelves, at one of the five possible interval durations of the visual metronome. Alternate flashing of the toy cars occurred at the bouncing locations of the moving metronomes, either at the left or right end of the bookshelves.

Tapping precision

The slope method is a classical timing model that uses linear regression between TV as a function of target duration (t_d) to reach a generalized form of Weber's law (Figure 2A, data from the first produced interval in the continuation epoch). The resulting slope ($slope_{TV}$) is associated with the time-dependent process since it captures the scalar property of interval timing. The intercept ($intercept_{TV}$) is related to the time-independent component, which is the fraction of variance that remains similar across interval durations and is associated with sensory detection and processing, decision making, memory load, and/or motor execution.[59–61] As a convention, we computed the $intercept_{TV}$ at the intermediate target interval of 650 ms instead of at 0 ms, as usually computed in linear regression (Figure 2A). $Intercept_{TV}$ and $slope_{TV}$ were computed for each subject and condition through linear regression. Figure 2B shows, for the deceleration conditions (mean of the four directions), that $slope_{TV}$ values plotted as a function of $intercept_{TV}$ follow a U-shaped curve across the produced intervals during the synchronization (S1-S5) and continuation (C1-C5) epochs. The time-independent component was large at S1, decreased within the sequence around C1, and rebounded at the end of the continuation epoch. The time-dependent component showed a systematic decrease during synchronization and a slight increase during continuation (Figure 2C). Similar U-shaped curves were observed for the eight motion conditions, with larger $intercept_{TV}$ and $slope_{TV}$ for the deceleration condition compared to the acceleration condition, regardless of the target motion directions (see Figure S1A). To parametrize these effects, we first performed repeated measures ANOVAs using either $intercept_{TV}$ or $slope_{TV}$ as dependent variables and serial order (1–10), stimulus kinematics (acceleration, deceleration), and direction (the four cardinal directions) as factors. The results for $intercept_{TV}$ showed significant main effects of serial order and stimulus kinematics, but neither significant main effects for direction nor significant interactions, except for the significant serial order x stimulus kinematics interaction (Table S1A). In contrast, the $slope_{TV}$ ANOVA did not show significant main or interaction effects (Table S1A). The corresponding repeated measures ANOVAs for the flashing metronome conditions showed only significant main effects of serial order on both $intercept_{TV}$ and $slope_{TV}$ (Table S1B). In Figure 2C, we compared either $intercept_{TV}$ or $slope_{TV}$ among flashing, acceleration, and deceleration conditions. It is evident that the responses to decelerated motion display the largest $intercept_{TV}$, with responses

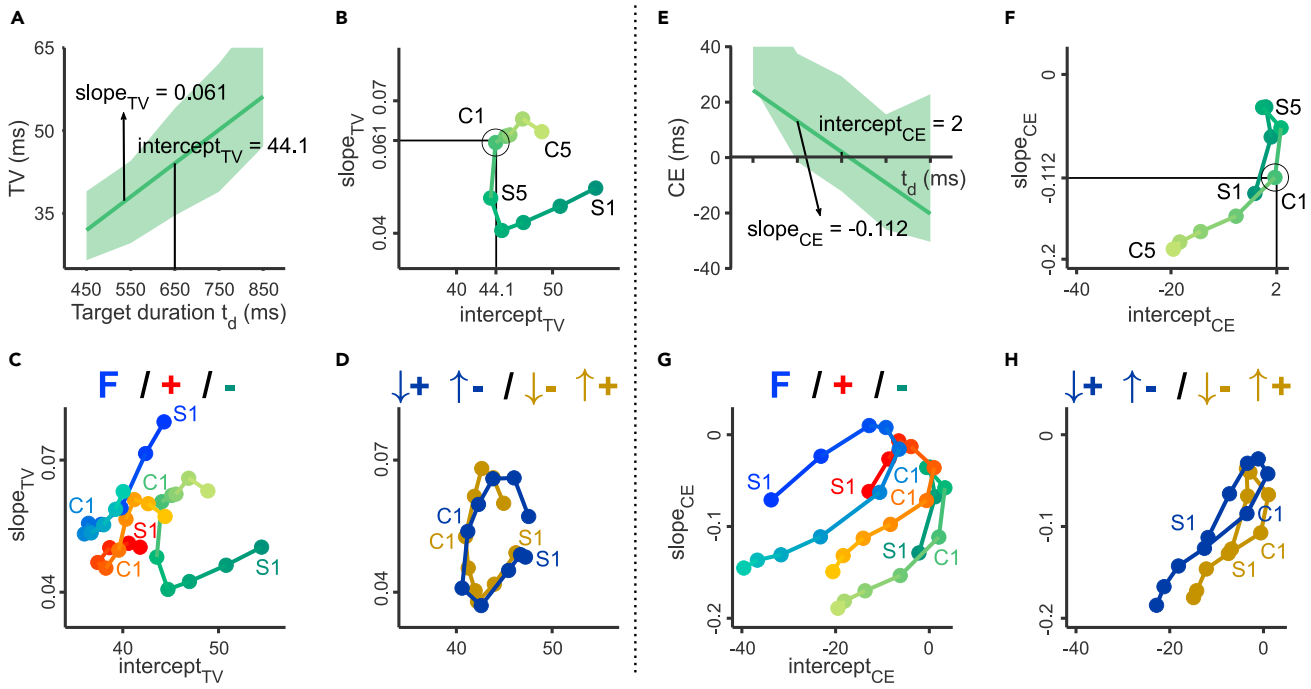


Figure 2. Timing accuracy and precision

(A) The temporal variability (TV; median and interquartiles) of the first interval of the continuation epoch (C1) in the deceleration conditions plotted as a function of target duration (t_d). The corresponding linear regression was used to determine the intercept at 650 ms (intercept$_{TV}$ = 44.1 ms), and the slope (slope$_{TV}$ = 0.061 ms).

(B) Slope$_{TV}$ plotted as a function of intercept$_{TV}$ across the serial order elements (S1-S5, C1-C5) of the SCT for the deceleration conditions. The circle in C1 highlights the values obtained from A.

(C) Slope$_{TV}$ as a function of intercept$_{TV}$ for all serial order elements for the flashing metronomes (blue), all directions with accelerated motion (red +) and all directions with decelerated motion (green-). Note the U-shaped progression of slope$_{TV}$ and intercept$_{TV}$ across the serial order elements of the SCT. The time-independent variability (intercept$_{TV}$) is larger in deceleration conditions, and the time-dependent variability (slope$_{TV}$) is larger in the Flashing condition for the first two serial order elements (S1, S2).

(D) Slope$_{TV}$ plotted as a function of intercept$_{TV}$ across serial orders by pooling motion conditions in natural and non-natural gravity conditions. Note the overlapping temporal variability profiles in the two gravity conditions.

(E) Constant error (median and interquartiles) of the first interval of the continuation epoch (C1) for the decelerated motion condition is plotted as a function of t_d. We fitted a linear regression and extracted the intercept at 650 ms (intercept$_{CE}$ = 2 ms) and the slope (slope$_{CE}$ = −0.112 ms).

(F) Slope$_{CE}$ plotted as a function of intercept$_{CE}$ across the produced intervals of the synchronization and continuation epochs for the deceleration conditions. The open circle depicts the data from E. Note the inverted U-shaped progression of slope$_{CE}$ and intercept$_{CE}$ across the serial order elements of the SCT.

(G) Slope$_{CE}$ as a function of intercept$_{CE}$ for all serial order elements for the flashing metronomes (blue), all directions with accelerated motion (red +) and all directions with decelerated motion (green). Note that intercept$_{CE}$ reached low values for flashing conditions, intermediate values for acceleration conditions, and large values for deceleration conditions. (H) Slope$_{CE}$ as a function of intercept$_{CE}$ across serial orders for the natural and non-natural gravity conditions. Note that the constant error profiles between both gravity conditions are similar but shifted along the abscissa.

to accelerated and flashing metronomes exhibiting comparable intercept$_{TV}$ values across all produced intervals in the SCT sequence. Indeed, significant differences emerged only between flashing and decelerated motion (Table S1C). In addition, the slope$_{TV}$ values in the first synchronization interval (S1) with flashing metronomes are substantial and significantly different from those obtained with moving metronomes for the same S1 interval ($F(2,31) = 6.5$, $p < 0.0001$; flashing against acceleration and deceleration conditions with paired t tests, $p < 0.05$). We also tested the effects of natural and non-natural gravity conditions in a separate repeated measures ANOVA, with no significant differences for either intercept$_{TV}$ or slope$_{TV}$ (Figure 2D; Table S1D).

These results suggest that visual motion produced changes mostly on the time-independent component of the SCT, with lower variability in response to accelerated motion and flashing, and the highest variability for decelerated motion. With respect to the serial order in rhythmic production, larger and smaller variability in both time-independent and dependent components was observed for S1 and S5, respectively. No effect of visual motion naturalness was found.

Tapping accuracy

Figure 2E shows the CE as a function of t_d for the first produced interval of the continuation epoch during the deceleration conditions, with the typical bias effect of overestimation for short intervals and underestimation for long intervals. We defined two measures to characterize this behavior using linear regression on these data: the intercept at 650 ms ($intercept_{CE}$) and the slope of the regression ($slope_{CE}$; see Figure 2E). The former is proportional to the indifference interval, namely, the interval where there is no error in timing.[10,58,62] The latter corresponds to the magnitude of the bias effect (i.e., the larger the negative slope, the larger the regression toward the mean).[11,63,64] Then, we plotted the $slope_{CE}$ as a function of the $intercept_{CE}$ for each produced interval in the SCT sequence. In Figure 2F, it is evident that there is an inverted U shape in the change of these two parameters across the SCT sequence with flashing metronomes. At the beginning of the task (S1), there was a large negative slope with an $intercept_{CE}$ around −10 ms (smaller than the intermediate target interval). Then the slope approached zero in the next synchronization intervals (S2-S5), with the $intercept_{CE}$ migrating toward 0 ms, suggesting that the indifference interval was around 650 ms. During continuation (C1-C5), the slope again became negative and large, and the $intercept_{CE}$ acquired stationary values around −20 ms. These results indicate the crucial effects of the sequence on timing accuracy: (1) the initial interval in the rhythmic sequence shows a large bias effect, (2) during the next synchronization intervals the accuracy is close to perfect, and (3) during the continuation epoch the bias effect resumes with a stable $intercept_{CE}$ that may reflect the subject's preferred interval under internal rhythmic tapping. Similar trends were observed for all motion conditions, as well as for the flashing conditions (Figure S1). Repeated measures ANOVAs showed, for the moving metronome conditions, significant main effects of serial order and motion kinematics, but no significant main effects for direction or significant interactions for both $slope_{CE}$ and $intercept_{CE}$ (Table S1A). Moreover, for the flashing conditions, the repeated measures ANOVA showed only significant main effects for serial order on $slope_{CE}$ and $intercept_{CE}$ (Table S1B). We compared the $slope_{CE}$ and $intercept_{CE}$ patterns across serial orders between flashing, acceleration, and deceleration conditions. The inverted U shape in the serial order of $slope_{CE}$ vs. $intercept_{CE}$ was the prevalent pattern across all conditions (Figure 2G), but with a significant increasing shift in the $intercept_{CE}$ profile between flashing, acceleration, and deceleration conditions (all paired t tests showing $p < 0.0001$, see Table S1C). In addition, $slope_{CE}$, which is a measure of the bias effect, was also larger in the deceleration conditions than in the flashing and acceleration conditions (Table S1C). Crucially, statistical differences were found between natural and non-natural gravity conditions for $intercept_{CE}$ but not for $slope_{CE}$ (Figure 2H; Table S1D).

These findings indicate that $intercept_{CE}$ was shorter for flashing and larger for deceleration. Although there was a strong tendency to reach an indifference interval around 550 ms at the end of the continuation epoch for all motion conditions (see Figure S2), the overall $intercept_{CE}$ was significantly larger in the non-natural gravity condition than in the natural gravity condition. Finally, the bias effect was larger for deceleration than acceleration and flashing conditions.

Tapping prediction

The asynchronies showed large negative values with decelerated motion and similar values with flashing and accelerated motion, which were closer to zero but with a systematic decreasing trend as a function of t_d and of serial order (Figure 3A, top). Indeed, the repeated measures ANOVA applied to the moving metronome conditions showed significant main effects of motion kinematics and t_d, as well as significant interactions for motion kinematics x t_d, motion kinematics x serial order, and serial order x t_d (Table S2A). With flashing metronomes, asynchronies were influenced significantly only by t_d and serial order (blue traces in Figure 3A, top; Table S2B). By comparing mean asynchronies with the static-kinetic model (Table S2C), we did find significant differences between deceleration conditions and either acceleration or flashing conditions, while asynchronies in response to accelerated and flashing stimuli were statistically indistinguishable (compare red and blue traces in Figure 3A; Table S2C). Visual motion naturalness did not have statistically significant effects (Table S2D).

The variability of the asynchronies, which corresponds to their intra-trial standard deviation, was significantly larger for the deceleration condition than the acceleration condition (especially for longer t_d) and showed a systematic increase with serial order, accounted for by a significant serial order x t_d interaction (Figure 3B; Table S2A). With flashing visual metronomes, tapping variability increased significantly with the serial order and t_d (Table S2B). Notably, the variability of the asynchronies was significantly lower not just when comparing the acceleration and deceleration conditions but also when comparing the

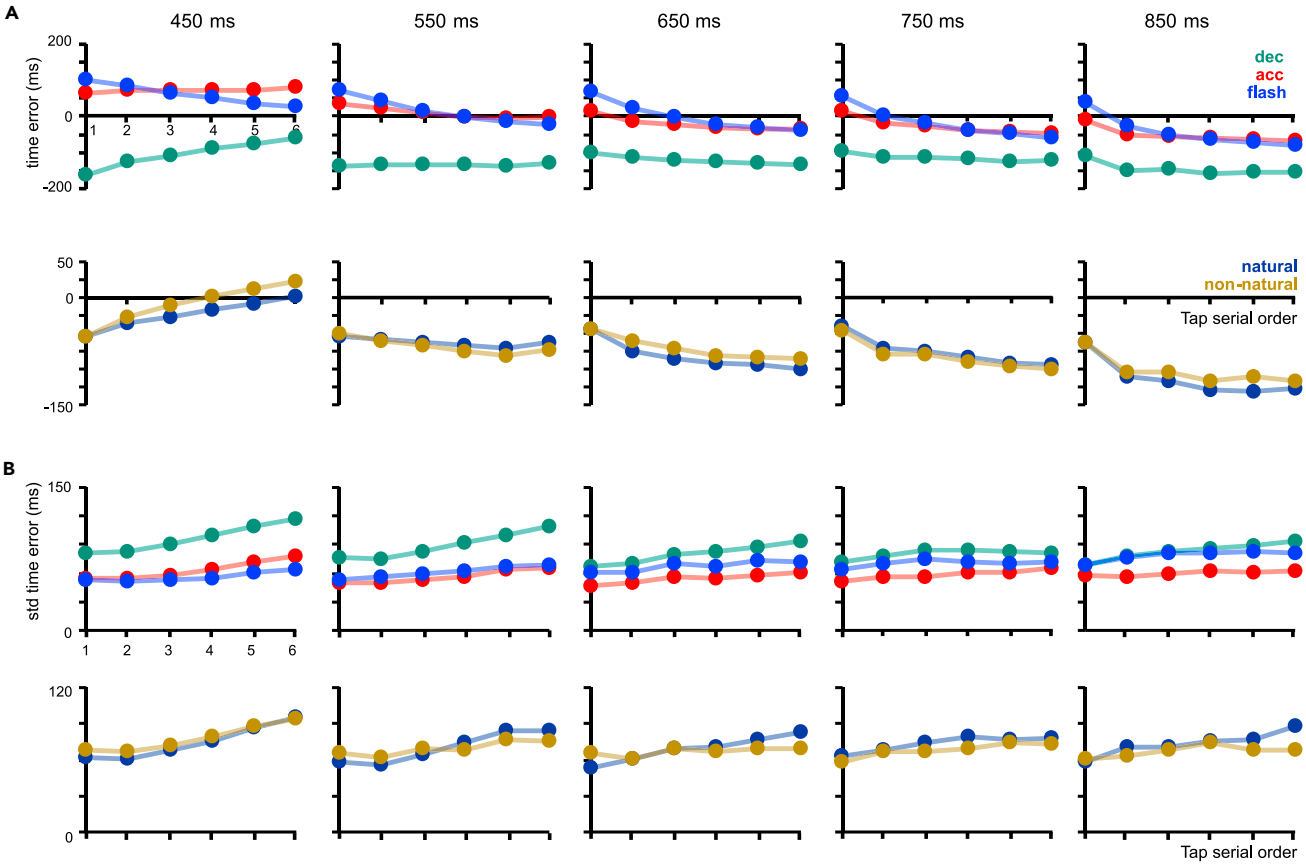


Figure 3. Tapping asynchronies

(A) Top row: Tapping asynchronies in response to accelerated (red), decelerated (green), and flashing (blue) visual stimuli. Each graph shows the mean asynchrony values computed across subjects for the series of six taps in the synchronization epoch for each of the five t_d (label at the top). Bottom row: Tapping asynchronies for the vertical motion conditions congruent with natural (blue) and non-natural (brown) gravity. Each graph shows the mean asynchrony values computed across subjects for the six taps of the synchronization epoch for each t_d.

(B) Top row: Variability in the asynchronies computed as standard deviations across trials as a function of the serial order of taps for the same experimental conditions in A, top row. Bottom row: Variability in the asynchronies for the conditions in A, bottom row.

acceleration and flashing conditions, indicating a more precise prediction for the acceleration condition (Table S2C). Finally, we found a slight interaction effect between the naturalness of the visual motion and the serial order of taps (Table S2D).

In sum, the results of these analyses indicated the following aspects. 1) With moving and flashing metronomes, the initial interval in the rhythmic sequence showed a large bias effect and time-independent component, but both accuracy and precision improved throughout the synchronization epoch. The time-dependent and independent components, as well as the bias effect, increased again during the continuation epoch, with stable intercept_{CE} values at the end of continuation epoch, reflecting the subject's internalized preferred interval. 2) With moving metronomes, the acceleration of the moving objects was the strongest factor determining precision and accuracy in SCT performance, whereas the effects of motion naturalness were limited to the increase of the indifference interval. 3) Compared to flashing metronomes, moving metronomes provided a significant advantage in SCT performance only when time intervals were cued by accelerating objects.

A Bayesian model for rhythmic tapping

In order to gain further insight into the dynamic evolution of the subjects' SCT performance across the different experimental conditions, we adapted the three-stage Bayesian model for single interval production, designed by Jazayeri and Shadlen,[63] to our SCT (Figure 4; see STAR methods). The original model

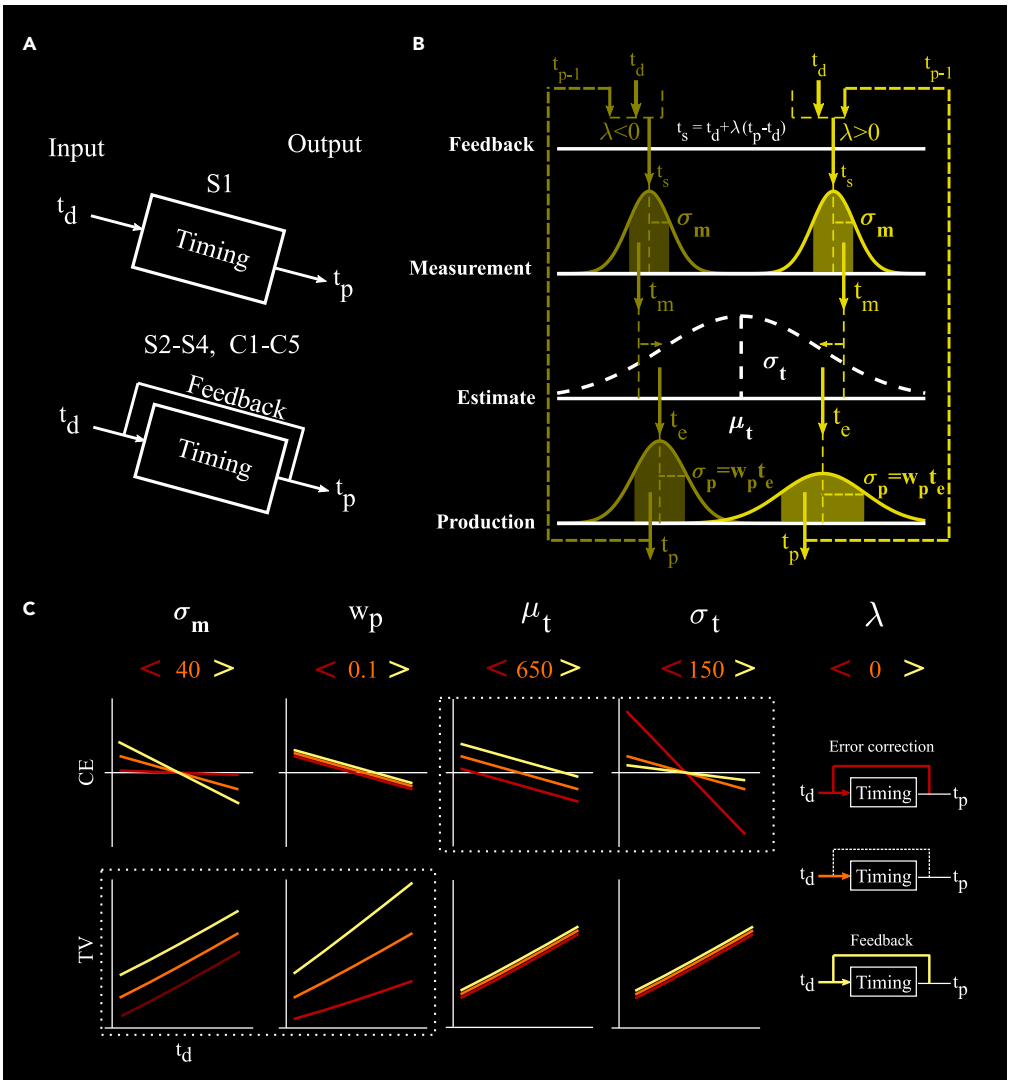


Figure 4. Bayesian model

(A) The model incorporates the change from an open loop (top, initial interval S1) to a closed loop (bottom, S2-S5 and C1-C5) in timing production. During the closed loop, the previous interval influences the current duration, allowing for either error correction (negative feedback) or interval drift (positive feedback).

(B) The model has a four-stage architecture. First, the target duration t_d, defined by the metronome, is measured with t_m, which is modeled with a Gaussian distribution centered in t_s and adds noise σ_m to the relation $p(t_m|t_s)$. Next, the observer computes a time estimate (t_e) that is the maximum of the posterior probability, which is proportional to the product of the likelihood function $p(t_m|t_s)$ and the prior probability distribution $p(t_s)$ (a Gaussian with a mean μ_t, and a standard deviation σ_t). Third, t_e is used to compute the produced interval t_p with a conditional dependence of t_p on t_e, $p(t_p|t_e)$, which adds noise to the production phase by using Gaussian distribution, whose standard deviation is $w_p t_e$. The $w_p t_e$ is larger for longer estimated intervals, simulating the scalar property of timing. The last stage considers the influence of the previous t_{p-1} on the current interval, computing the difference between the previously produced and the target duration as $t_s = t_d + \lambda(t_{p-1} - t_d)$. This difference influences the current measured time with a λ weight. Thus, in the first produced interval $\lambda = 0$, but in the following intervals the impact of the previous produced duration depends on how large the negative feedback error correction or the positive feedback drift adjusts the actual t_p. Simulations of the effect of the small (red), intermediate (orange), and large (yellow) values of the parameters of the model on the constant error (CE) and temporal variability (TV). The increase in σ_m produces large changes in the time-independent variability ($intercept_{TV}$) and a moderate increase in the bias effect. The increase in w_p induces a large increment in $intercept_{TV}$ and $slope_{TV}$ with no changes in CE. An increase in μ_t produces a shift toward the right in the indifference interval and no changes in TV. A rise in σ_t induces an increase in the bias effect. Finally, $\lambda = 0$ implies no feedback, $\lambda < 0$ involves negative feedback error correction, and $\lambda > 0$ entails positive feedback with interval drift. The orange numbers on top of the CE plots correspond

Figure 4. *Continued*
to the actual intermediate values of the model used to obtain the CE and TV. The small and large values for these simulations are close to the minimum and maximum values of the scale (color bar) for each model parameter depicted in Figure 5.

includes an initial measurement stage, where the input interval (t_s) is transformed by a noisy observer to measured time (t_m) using a Gaussian distribution centered on t_s and with a standard deviation σ_m (Figure 4B). Then, the observer computes the maximum likelihood estimator (t_e) as the maximum of the posterior probability, which is proportional to the product of the likelihood function $p(t_m|t_s)$ and the prior probability distribution $p(t_s)$. The prior $p(t_s)$ is a Gaussian distribution with a mean μ_t and a standard deviation σ_t that crucially maps the history of sample intervals that the observer has encountered during a block of trials, producing the bias effect. The third stage uses t_e to compute the produced interval t_p, utilizing the conditional dependence of t_p on t_e as $p(t_p|t_e)$, which adds noise to the production phase by using Gaussian distribution, whose standard deviation is $w_p t_e$. Hence, $w_p t_e$ is larger for longer estimated intervals, simulating the scalar property of timing. The SCT involves the sequential production of intervals in a closed loop (Figure 4A). Typically, during the synchronization epoch, an error correction mechanism generates the tendency for short-produced intervals to be followed by a long interval and vice versa, thereby avoiding large error accumulation that could drive tapping responses out of sync with the metronome. In contrast, during the continuation epoch, subjects use an internal clock that tends to generate a small drift in the produced intervals toward progressively shorter or longer intervals. Hence, in order to capture both phenomena, we added a fourth stage to the model. In this step, the difference between the previously produced interval (t_{p-1}) and the physical interval (t_d) influences the input time as $t_s = t_d + \lambda(t_{p-1} - t_d)$ with a λ weight. In the first produced interval $\lambda = 0$ (Figure 4C, λ middle). In the following intervals of the synchronization epoch, λ depends on how much error correction adjusts the actual t_p, with values below zero since there is a negative feedback loop (Figure 4C, λ top) (see the lag 1 autocorrelation between consecutive intervals in the SCT in Figure S3A). Finally, during the continuation epoch, λ tends to be positive because the positive loop of produced intervals causes a drift (Figure 4C, λ bottom; see also the drift of produced intervals during the continuation epoch in Figure S3B). It is important to consider that the noise parameters σ_m for time measurement and w_p for time production capture the changes in TV during the SCT. In contrast, μ_t and σ_t capture the effects on the CE, with the former linked to the intercept$_{CE}$ and indifference interval and the latter related to the magnitude of the bias effect (slope) (see Figure 4C and S4C).

For each subject, we fitted the model parameters, namely, σ_m, μ_t, σ_t, w_p, and λ to the SCT for the flashing condition and the eight visual motion conditions. Goodness-of-fit values were large, with significant correlations between the predicted and actual data across subjects (Figures S4A and S4B). Figure 5 shows the coefficients of the model across the serial orders of the eight motion conditions and the flashing condition. A repeated measures ANOVA applied to the parameters fitted to the motion conditions indicated that the time measurement noise (σ_m) showed significant main effects of serial order, motion kinematics, and their interaction (Figure 5A top; Table S3A), with larger values for S1 across all kinematic conditions and smaller values for the acceleration condition than for the deceleration condition. σ_m also exhibited significant main effects of serial order and metronome position in the ANOVA applied to the flashing condition (Table S3B). The static-kinetic ANOVA showed significant effects, with pairwise differences between deceleration and either flashing or acceleration conditions (Figure 5A, bottom; Table S3C).

The time production noise (w_p), which is associated with the slope of the scalar property, showed a significant main effect of serial order and kinematics Figure 5B, top; Table S3A) for the moving metronome conditions and only a main effect of serial order with flashing metronomes (Table S3B). A large w_p was observed for S1 Figure 5B, bottom), and a significantly large w_p was observed in the flashing condition compared to the accelerated motion condition Figure 5B bottom; Table S3C).

The mean of the prior distribution (μ_t), a parameter linked to the indifference interval and intercept$_{CE}$, showed significant main effects of serial order and motion kinematics for the moving metronomes ANOVA and only of serial order for the flashing metronomes ANOVA Figure 5C, top; Tables S3A and S3B). This parameter was larger for deceleration than for acceleration and flashing conditions (Table S3C) and larger for the synchronization epoch than for the continuation epoch (Figure 5C, bottom).

The sigma of the prior (σ_t), a parameter associated with the magnitude of the bias effect, showed a significant main effect of serial order for both the motion and flashing metronome ANOVAs (Figure 5D,

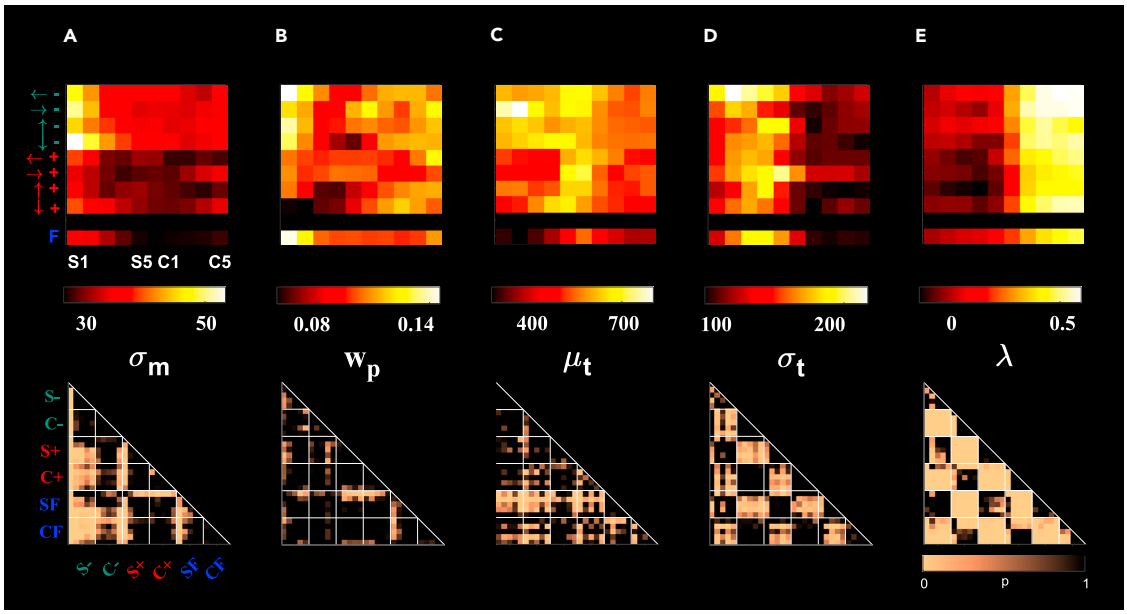


Figure 5. Parameters of the Bayesian model across the flashing and motion conditions

(A) Top. σ_m is plotted for the eight motion conditions (arrows specify the direction: green for deceleration, red for acceleration) and the flashing condition (blue F) as a function of serial order in the SCT, ranging from S1 to C5. The scale on the heatmap is at the bottom. Bottom. Diagonal matrix of the pairwise Bonferroni post-hoc test between the five serial order elements of the synchronization epochs and the corresponding five of the continuation epochs for the four deceleration conditions (green), four acceleration conditions (red), and the flashing condition (blue F).

(B, C, D, E) Values of μ_t, σ_t, w_p, and λ, respectively, plotted as in A.

top; Table S3A and S3B), since it was smaller for S1 and C1-C5, where the bias effect was also more evident (Figure 5D, bottom). σ_t was larger for deceleration than for acceleration and flashing conditions (Table S3C).

Finally, λ, a parameter that captures the correlation between consecutively produced intervals in the SCT, showed significant main effects of serial order and motion kinematics for the moving metronomes ANOVA and only of serial order for the flashing metronomes ANOVA Figure 5E, top; Tables S3A and S3B). λ was negative only for the synchronization epoch of the acceleration condition and significantly different from the deceleration condition, indicating robust error correction (Table S3C). Moreover, λ was positive and larger for the continuation epoch than for the synchronization epoch (Figure 5E, bottom), indicating the drift in produced intervals during the internally driven phase of the SCT.

It is striking that none of the Bayesian model parameters, except for μ_t, showed significant differences between the natural and non-natural gravity conditions (Table S3D). Overall, the Bayesian model captured all the precision and accuracy properties of rhythmic performance in the SCT across the target conditions described earlier. Moreover, regardless of the target condition, subjects optimally switched between the sensory-guided tap synchronization and the internally driven continuation using mechanisms of error correction and regression toward the mean, respectively.

DISCUSSION

The results of the present study support three main conclusions. First, when syncing to moving metronomes, timing accuracy, precision, and the correlation of consecutively produced intervals were more efficient with accelerated motion than with decelerated motion and hardly influenced by object motion naturalness, implying that motion kinematics per se was the predominant factor determining rhythm synchronization. Second, metronomes cued by accelerated motion provided an overall performance advantage, also relative to flashing visual metronomes. Third, a Bayesian observer model could account for the changes in the bias effect and scalar property, as well as the error correction or drift in produced intervals within the synchronization and continuation epochs of the SCT across motion conditions,

providing a parametric account of isochronic rhythmic behavior that indicates an optimal shift of behavior that maximizes temporal precision and accuracy depending on the epoch of the SCT.

Effects of motion kinematics and naturalness on rhythmic timing

It has been previously shown that the auditory-visual asymmetry for beat induction can be canceled out by visual moving metronomes,[37] especially with naturalistic stimuli such as videos of a continuous bouncing ball with changes in speed.[14,38,39] One hypothesis to explain these observations is that visual motion may engage the time-to-collision mechanism underlying interceptive action control or collision avoidance.[41,42,65] According to this idea, the parieto-premotor system recruited for encoding time to contact for discrete events could also be involved in the internal representation of periodic collision points, efficiently driving the motor system for beat-based timing.[14,44,45] The present study was designed to test whether physically realistic motion trajectories with downward acceleration or upward deceleration profiles induced better rhythm entraining in the SCT than non-natural vertical acceleration/downward deceleration or horizontal acceleration/deceleration motion metronomes. In effect, a natural gravity gain on rhythmic timing might be expected based on the evidence that diving gannets use gravitational signals to compute the time to contact and fold their wings before entering the water[66] and humans can use an internal representation of gravity effects residing in the vestibular cortex to accurately time the interception of vertically falling objects.[52,67–69] However, our findings were not congruent with this expectation since the timing accuracy, precision, and correlation of consecutively produced intervals were similar between natural and non-natural gravity and vertical and horizontal motion conditions. Rather, they suggest that the parieto-premotor internal representation of periodic collision points either does not integrate or downweighs signals from the parieto-insular vestibular cortex, which are known to encode *a priori* information about gravity,[50] to generate beat representation during the SCT. The only consistent effect of the naturalness of visual motion was a decrease in the indifference interval, implying that gravity-related information may tune the preferred tempo representation.

Our results also indicated that the parieto-premotor system entrains rhythmic tapping more efficiently with metronomes cued by accelerated motion than with decelerated motion or flashing metronomes.[70] Neurons in the middle temporal area (MT) of the extrastriate visual cortex, which is reciprocally connected with the posterior parietal cortex,[71] are tuned for direction and speed of moving visual stimuli but not for acceleration-deceleration.[72,73] However, recent studies have shown that acceleration information can be extracted by assessing changes over time in the stimulus speed population code in the MT.[74,75] We could thus hypothesize that the reader of speed population signals from the MT in the parietal-premotor system could build a stronger rhythmic timing prediction with accelerating motion than with decelerating motion. For example, this could be achieved by a better estimation of the difference between the initial object velocity and the velocity at the bounce in the acceleration condition due to a bias in the reading system toward increasing speeds. In other words, the stronger internal representation of periodic collision points may depend on the more robust estimation of motion acceleration. In addition, the high terminal velocity and sudden speed changes observed in the accelerating objects could have entrained sharper and less variable ramping activities in parietal and premotor neurons that encode event timing and control rhythmic tapping behavior by producing higher and sharper activity levels in MT neurons tuned to high speeds.[76,77]

Preferred tempo and the internally produced intervals

Classical tapping studies coined the notion of preferred rhythmic tempo, which refers to the interval produced naturally when individuals tap in the absence of external cues.[78–80] The preferred tempo in human adults is around 600 ms. In addition, since Fraisse (1963), indifference interval phenomena[78,81] have been linked to the preferred internal periodicity. Thus, the convergence of the CE toward zero is considered the result of the dominance of an internal clock with prevailing timing. For example, Jones and McAuley (2005)[82] computed the preferred period from the CE of an interval discrimination task and found that the preferred period is not fixed and varies depending on the range of intervals used in an experimental session. The present findings indicate that, during the last three intervals of the continuation epoch, the subjects' indifference interval converged to a particular value that was shorter for the flashing condition than for accelerating and decelerating conditions and was larger for non-natural gravity than natural gravity conditions (Figure 2). These findings suggest that the internal timing mechanism uses the bias effect to cope with the lack of entraining sensory input across conditions, but the preferred tempo is directly affected by motion properties of the metronome, changing the indifference interval even during

continuation where the stimuli were turned off several intervals before. These changes in indifference interval could be due to changes in the mean of the internal prior function. Neurophysiological studies have shown that the neural population dynamics in the primate medial premotor areas can represent the prior distribution of timing.[5,83] Hence, we suggest that the motion condition and naturalness had a differential effect on the encoding of timing priors in the motor system through the visual motion encoders in the parietal cortex.[26]

Bayesian modeling of rhythmic timing

The results of the model revealed interesting features concerning the synchronization strategies adopted with moving and flashing metronomes: a) subjects used three different optimal strategies to produce intervals within the tapping sequence: a large mix between prior knowledge of the statistics of used intervals and the time measurement at the beginning of the synchronization epoch, which is sensitive to the properties of the metronome; b) an error correction mechanism during the successive intervals of the synchronization that compensated for the effect of the prior and increased the timing precision; c) a balanced use of prior and time measurement during the internally driven epoch of the task, with a large feedforward influence between consecutive intervals producing the characteristic drift of the continuation; and d) the indifference interval was modulated by serial order but reached homogeneous values at the end of the continuation epoch, suggesting the expression of a preferred interval that was modulated by the metronome's properties.

The use of Bayesian inference to describe timing performance has been widely employed to characterize the optimal combination of prior knowledge about the statistics of the presented intervals of an experiment with the likelihood function of the measured time.[64,84] This modeling approach offers some advantages over more classical approaches, such as the iconic Wing and Kristofferson model,[85,86] which focused only on the continuation epoch and was designed to split the total variability into a motor and a scalar timing component without including timing accuracy or the change in correlation between consecutively produced intervals in the two epochs of the SCT. Bayesian models, like the one we adopted, can not only capture the scalar property of timing by adding noise that is proportional to the interval during time production but also explain the bias, as well as the negative or positive correlation of lag 1 intervals. Particularly, in the Bayesian framework, the prior represents the bias or regression toward the mean effect, and it has been suggested to be the result of an error minimization strategy.[63,87,88] In contrast, our model explains the bias and scalar properties of timing, as well as the negative or positive correlation of lag 1 intervals. Our results indicate that this error optimization strategy was used at the beginning of the SCT sequence and during the internally driven epoch of the trial. However, during synchronization, it was replaced by an error correction mechanism that could adjust the duration of consecutive intervals in a negative feedback fashion (short followed by long, long followed by short). This mechanism allows for the compensation of movement timing to avoid large error accumulation.[28,89] Previous psychophysical studies in humans have reported strong error correction for auditory metronomes and weak error correction for flashing visual metronomes.[14,45] Here we found that accelerated but not decelerated motion produced robust error correction, comparable if not superior to that observed with flashing metronomes. Crucially, the parameter lambda that we included in the new version of the Bayesian model by Jazayeri and Shadlen[26,63] accounted for the error correction strategy during synchronization and for the drift in the duration of the produced intervals during the continuation epoch. These findings support the notion that, during rhythmic timing, humans dynamically use prior knowledge or error correction to optimize time interval production depending on the task contingencies, adding a new dimension to the optimal control of rhythmic behavior.

Visual timing in practice: Kinematics of the conductor's baton

A cogent, real-world example of visual rhythmic timing is in an orchestra, where a conductor uses visual gestures, among a range of other expressive features, to communicate clear timing cues that enable individuals in an ensemble to produce highly synchronized sound.[90] In addition to gesturing beats, the conducting pattern typically follows standard shapes depending on the musical meter, as the two-, three-, and four-beat hand-arm trajectories. Notably, the shapes produced by the conductor are not solely in the vertical plane. Rather, they involve movements in different directions to produce the typical shapes which cue the ensemble. Each shape represents an entire bar of music consisting of two, three, or four beats. Therefore, beats are not conveyed only through vertical movements, which may exploit the effects of gravity. Indeed, empirical studies investigating conducting movement kinematics on synchronization

through tapping indicated that the acceleration of the baton is the key kinematic variable explaining synchronization behavior,[91] more so than the direction of motion, the changes in direction, or the instantaneous speed.[92] Notwithstanding the differences between the biological motion patterns generated by the orchestra conductor and the inanimate object motion employed in the present study, our finding that accelerating metronomes, regardless of movement direction, produced the most precise and accurate tapping is consistent with these studies of orchestral conducting.

Conclusion

In sum, our experimental findings and computational modeling link together and add detail to a consistent account of the kinematic properties of movement that can induce a visual beat in both laboratory and real-world settings, and they provide insight into the neural dynamics in parieto-premotor regions that underlie these abilities. Furthermore, our model provides a parametric account not just for the occurrence of taps but also for the dynamic production of taps on a beat-by-beat basis, given previous sensory and motor history. In particular, the four stages of our model (measurement, internal estimation, motor production, and feedback) may provide a useful framework for deeper quantitative investigation of synchronization to visual cues provided by a conductor. These insights may also be relevant for computational gesture analysis and interactive tools based on conducting,[93,94] and for our understanding of whole-body movement dynamics during dance.[95]

Limitations of the study

This study has three main caveats. The first one is the lack of an auditory metronome condition. This would allow comparing the gain in the precision and predictability of tapping between the classical auditory SCT and the present accelerating and decelerating visual motion conditions. Second, a task condition with no changes in the position of the moving visual stimuli and the tapping locations could definitively rule out the possible effects of stimulus-response position on the observed changes in tapping behavior. Finally, the subjects produced five intervals in the synchronization and continuation epochs. With a larger number of intervals produced in the SCT sequence, subjects could have reached more stable error correction and bias effects in the two epochs of our task.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Synchronization-continuation task
  - Slope analysis
  - Bayesian model
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107543.

ACKNOWLEDGMENTS

Hugo Merchant is supported by CONACYT: A1-S-8430, UNAM-DGAPA-PAPIIT IN201721.

Gianfranco Bosco is supported by #NEXTGENERATIONEU (NGEU) and funded by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006) – A Multiscale integrated approach to the study of the nervous system in health and disease (DN. 1553 11.10.2022).

We are very grateful for the valuable comments that Vani Rajendran provided to our manuscript. We also thank Raul Paulín and Luis Prado for their technical assistance.

AUTHOR CONTRIBUTIONS

HM and GFB conceived the study; GFB and SDM collected the data. OP and SDM performed data analyses. OP contributed analytical tools and guidance. HM, FL, and GFB supervised the project. HM and GFB wrote the manuscript. FL, GFB, OP, and SDM edited the manuscript.

DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

Received: March 13, 2023
Revised: May 30, 2023
Accepted: August 1, 2023
Published: August 22, 2023

REFERENCES

1. Honing, H. (2012). Without it no music: beat induction as a fundamental musical trait. Ann. N. Y. Acad. Sci. *1252*, 85–91. https://doi.org/10.1111/J.1749-6632.2011.06402.X.

2. Honing, H., Merchant, H., Háden, G.P., Prado, L., and Bartolo, R. (2012). Rhesus Monkeys (Macaca mulatta) Detect Rhythmic Groups in Music, but Not the Beat. PLoS One *7*, e51369. https://doi.org/10.1371/JOURNAL.PONE.0051369.

3. Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., and Fitch, W.T. (2015). Finding the beat: a neural perspective across humans and non-human primates. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20140093. https://doi.org/10.1098/RSTB.2014.0093.

4. Wing, A.M. (2002). Voluntary Timing and Brain Function: An Information Processing Approach. Brain Cogn. *48*, 7–30. https://doi.org/10.1006/BRCG.2001.1301.

5. Balasubramaniam, R., Haegens, S., Jazayeri, M., Merchant, H., Sternad, D., and Song, J.H. (2021). Neural Encoding and Representation of Time for Sensorimotor Control and Learning. J. Neurosci. *41*, 866–872. https://doi.org/10.1523/JNEUROSCI.1652-20.2020.

6. Gibbon, J., Malapani, C., Dale, C.L., and Gallistel, C. (1997). Toward a neurobiology of temporal cognition: advances and challenges. Curr. Opin. Neurobiol. *7*, 170–184. https://doi.org/10.1016/S0959-4388(97)80005-0.

7. Merchant, H., Zarco, W., and Prado, L. (2008). Do We Have a Common Mechanism for Measuring Time in the Hundreds of Millisecond Range? Evidence From Multiple-Interval Timing Tasks. J. Neurophysiol. *99*, 939–949.

8. García-Garibay, O., Cadena-Valencia, J., Merchant, H., and de Lafuente, V. (2016). Monkeys share the human ability to internally maintain a temporal rhythm. Front. Psychol. *7*, 1971. https://doi.org/10.3389/FPSYG.2016.01971/BIBTEX.

9. Woodrow, H. (1934). The temporal indifference interval determined by the method of mean error. J. Exp. Psychol. *17*, 167–188. https://doi.org/10.1037/h0070235.

10. McAuley, J.D., and Jones, M.R. (2003). Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing. J. Exp. Psychol. Hum. Percept. Perform. *29*, 1102–1125.

11. Pérez, O., and Merchant, H. (2018). The Synaptic Properties of Cells Define the Hallmarks of Interval Timing in a Recurrent Neural Network. J. Neurosci. *38*, 4186–4199. https://doi.org/10.1523/JNEUROSCI.2651-17.2018.

12. Merchant, H., Bartolo, R., Méndez, J.C., Perez, O., Zarco, W., and Mendoza, G. (2011). What can be inferred from multiple-task psychophysical studies about the mechanisms for temporal processing? In Multidisciplinary aspects of time and time perception (Springer), pp. 207–229.

13. Repp, B.H. (2005). Sensorimotor synchronization: A review of the tapping literature. Psychon. Bull. Rev. *12*, 969–992. https://doi.org/10.3758/BF03206433.

14. Iversen, J.R., Patel, A.D., Nicodemus, B., and Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals. Cognition *134*, 232–244. https://doi.org/10.1016/J.COGNITION.2014.10.018.

15. Madison, G. (2001). Variability in isochronous tapping: Higher order dependencies as a function of intertap interval. J. Exp. Psychol. Hum. Percept. Perform. *27*, 411–422. https://doi.org/10.1037/0096-1523.27.2.411.

16. Collier, G.L., and Ogden, R.T. (2004). Adding drift to the decomposition of simple isochronous tapping: An extension of the Wing-Kristofferson model. J. Exp. Psychol. Hum. Percept. Perform. *30*, 853–872. https://doi.org/10.1037/0096-1523.30.5.853.

17. Grahn, J.A., and Rowe, J.B. (2009). Feeling the Beat: Premotor and Striatal Interactions in Musicians and Nonmusicians during Beat Perception. J. Neurosci. *29*, 7540–7548. https://doi.org/10.1523/JNEUROSCI.2018-08.2009.

18. Merchant, H., Zarco, W., Perez, O., Prado, L., and Bartolo, R. (2011). Measuring time with different neural chronometers during a synchronization-continuation task. Proc. Natl. Acad. Sci. USA *108*, 19784.

19. Merchant, H., Pérez, O., Zarco, W., and Gámez, J. (2013). Interval Tuning in the Primate Medial Premotor Cortex as a General Timing Mechanism. J. Neurosci. *33*, 9082–9096.

20. Bartolo, R., Prado, L., and Merchant, H. (2014). Information Processing in the Primate Basal Ganglia during Sensory-Guided and Internally Driven Rhythmic Tapping. J. Neurosci. *34*, 3910–3923.

21. Sánchez-Moncada, I., Concha, L., and Merchant, H. (2022). Pre-supplementary motor cortex mediates learning transfer from perceptual to motor timing. Preprint at bioRxiv. https://doi.org/10.1101/2020.12.17.423301.

22. Crowe, D.A., Zarco, W., Bartolo, R., and Merchant, H. (2014). Dynamic Representation of the Temporal and Sequential Structure of Rhythmic Movements in the Primate Medial Premotor Cortex. J. Neurosci. *34*, 11972–11983.

23. Merchant, H., Pérez, O., Bartolo, R., Méndez, J.C., Mendoza, G., Gámez, J., Yc, K., and Prado, L. (2015). Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. Eur. J. Neurosci. *41*, 586–602. https://doi.org/10.1111/ejn.12811.

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

25. Merchant, H., and Bartolo, R. (2018). Primate beta oscillations and rhythmic behaviors. J. Neural. Transm. *125*, 461–470. https://doi.org/10.1007/S00702-017-1716-9/FIGURES/3.

26. Betancourt, A., Pérez, O., Gámez, J., Mendoza, G., and Merchant, H. (2022). Premotor population dynamics as neural substrate for auditory and visual rhythmic entrainment. bioRxiv. https://doi.org/10.1101/2022.08.14.503904.

27. Repp, B.H., and Penel, A. (2002). Auditory Dominance in Temporal Processing: New Evidence from Synchronization with Simultaneous Visual and Auditory Sequences. J. Exp. Psychol. Hum. Percept. Perform. *28*, 1085–1099. https://doi.org/10.1037/0096-1523.28.5.1085.

28. Repp, B.H., and Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. Psychol. Res. *68*, 252–270. https://doi.org/10.1007/S00426-003-0143-8/FIGURES/7.

29. Chen, Y., Repp, B.H., and Patel, A.D. (2002). Spectral decomposition of variability in synchronization and continuation tapping: Comparisons between auditory and visual pacing and feedback conditions. Hum. Mov. Sci. *21*, 515–532. https://doi.org/10.1016/S0167-9457(02)00138-0.

30. Merchant, H., and de Lafuente, V. (2014). Introduction to the neurobiology of interval timing. Adv. Exp. Med. Biol. *829*, 1–13.

31. Welch, R.B., and Warren, D.H. (1980). Immediate perceptual response to intersensory discrepancy. Psychol. Bull. *88*, 638–667. https://doi.org/10.1037/0033-2909.88.3.638.

32. Guttman, S.E., Gilroy, L.A., and Blake, R. (2005). Hearing what the eyes see: Auditory encoding of visual temporal sequences. Psychol. Sci. *16*, 228–235. https://doi.org/10.1111/J.0956-7976.2005.00808.X/ASSET/IMAGES/LARGE/10.1111_J.0956-7976.2005.00808.X-FIG2.JPEG.

33. Kubovy, M. (1988). Should We Resist the Seductiveness of the Space: Time:: Vision:Audition Analogy? J. Exp. Psychol. Hum. Percept. Perform. *14*, 318–320. https://doi.org/10.1037/0096-1523.14.2.318.

34. Patel, A.D., and Iversen, J.R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. Front. Syst. Neurosci. *8*, 57. https://doi.org/10.3389/FNSYS.2014.00057/ABSTRACT.

35. Merchant, H., and Honing, H. (2013). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. Front. Neurosci. *7*, 274. https://doi.org/10.3389/FNINS.2013.00274/BIBTEX.

36. Lenc, T., Merchant, H., Keller, P.E., Honing, H., Varlet, M., and Nozaradan, S. (2021). Mapping between sound, brain and behaviour: Four-level framework for understanding rhythm processing in humans and non-human primates. Philos. Trans. R. Soc. Lond. B Biol. Sci. *376*, 20200325. https://doi.org/10.1098/rstb.2020.0325.

37. Hove, M.J., Spivey, M.J., and Krumhansl, C.L. (2010). Compatibility of Motion Facilitates Visuomotor Synchronization. J. Exp. Psychol. Hum. Percept. Perform. *36*, 1525–1534. https://doi.org/10.1037/A0019059.

38. Hove, M.J., Iversen, J.R., Zhang, A., and Repp, B.H. (2013). Synchronization with competing visual and auditory rhythms: Bouncing ball meets metronome. Psychol. Res. *77*, 388–398. https://doi.org/10.1007/s00426-012-0441-0.

39. Gan, L., Huang, Y., Zhou, L., Qian, C., and Wu, X. (2015). Synchronization to a bouncing ball with a realistic motion trajectory. Sci. Rep. *5*, 11974–11979. https://doi.org/10.1038/srep11974.

40. Merchant, H., Battaglia-Mayer, A., and Georgopoulos, A.P. (2003). Functional organization of parietal neuronal responses to optic-flow stimuli. J. Neurophysiol. *90*, 675–682. https://doi.org/10.1152/JN.00331.2003/ASSET/IMAGES/LARGE/9K0833271005.JPEG.

41. Merchant, H., Battaglia-Mayer, A., and Georgopoulos, A.P. (2004). Neural Responses during Interception of Real and Apparent Circularly Moving Stimuli in Motor Cortex and Area 7a. Cereb. Cortex *14*, 314–331. https://doi.org/10.1093/CERCOR/BHG130.

42. Merchant, H., and Georgopoulos, A.P. (2006). Neurophysiology of perceptual and motor aspects of interception. J. Neurophysiol. *95*, 1–13. https://doi.org/10.1152/JN.00422.2005/ASSET/IMAGES/LARGE/Z9K0120551470014.JPEG.

43. Li, Y., Wang, Y., and Cui, H. (2022). Posterior parietal cortex predicts upcoming movement in dynamic sensorimotor control. Proc Natl Acad Sci USA *119*, e2118903119. https://doi.org/10.1073/PNAS.2118903119/SUPPL_FILE/PNAS.2118903119.SAPP.PDF.

44. Mendoza, G., and Merchant, H. (2014). Motor system evolution and the emergence of high cognitive functions. Prog. Neurobiol. *122*, 73–93. https://doi.org/10.1016/J.PNEUROBIO.2014.09.001.

45. Comstock, D.C., Hove, M.J., and Balasubramaniam, R. (2018). Sensorimotor synchronization with auditory and visual modalities: Behavioral and neural differences. Front. Comput. Neurosci. *12*, 53. https://doi.org/10.3389/FNCOM.2018.00053/BIBTEX.

46. Kim, I.K., and Spelke, E.S. (1992). Infants' Sensitivity to Effects of Gravity on Visible Object Motion. J. Exp. Psychol. Hum. Percept. Perform. *18*, 385–393. https://doi.org/10.1037/0096-1523.18.2.385.

47. Twardy, C.R., and Bingham, G.P. (2002). Causation, causal perception, and conservation laws. Percept. Psychophys. *64*, 956–968. https://doi.org/10.3758/BF03196799.

48. Bingham, G., Bishop, R., Brody, M., Bromley, D., Clark, E.T., Cooper, W., Costanza, R., Hale, T., Hayden, G., Kellert, S., et al. (1995). Issues in ecosystem valuation: improving information for decision making. Ecol. Econ. *14*, 73–90. https://doi.org/10.1016/0921-8009(95)00021-Z.

49. Pittenger, J.B. (1990). Detection of Violations of the Law of Pendulum Motion: Observers' Sensitivity to the Relation Between Period and Length. Ecol. Psychol. *2*, 55–81. https://doi.org/10.1207/s15326969eco0201_3.

50. Indovina, I., Maffei, V., Bosco, G., Zago, M., Macaluso, E., and Lacquaniti, F. (2005). Representation of Visual Gravitational Motion in the Human Vestibular Cortex. Science *308*, 416–419. https://doi.org/10.1126/SCIENCE.1107961.

51. Lacquaniti, F., Bosco, G., Indovina, I., La Scaleia, B., Maffei, V., Moscatelli, A., and Zago, M. (2013). Visual gravitational motion and the vestibular system in humans. Front. Integr. Neurosci. *7*, 101. https://doi.org/10.3389/FNINT.2013.00101/BIBTEX.

52. Delle Monache, S., Indovina, I., Zago, M., Daprati, E., Lacquaniti, F., and Bosco, G. (2021). Watching the Effects of Gravity. Vestibular Cortex and the Neural Representation of "Visual" Gravity. Front. Integr. Neurosci. *15*, 793634. https://doi.org/10.3389/FNINT.2021.793634.

53. La Scaleia, B., Zago, M., and Lacquaniti, F. (2015). Hand interception of occluded motion in humans: A test of model-based vs.on-line control. J. Neurophysiol. *114*, 1577–1592. https://doi.org/10.1152/JN.00475.2015/ASSET/IMAGES/LARGE/Z9K0131532310010.JPEG.

54. Delle Monache, S., Lacquaniti, F., and Bosco, G. (2017). Differential contributions to the interception of occluded ballistic trajectories by the temporoparietal junction, area hMT/V5+, and the intraparietal cortex. J. Neurophysiol. *118*, 1809–1823. https://doi.org/10.1152/JN.00068.2017/ASSET/IMAGES/LARGE/Z9K0091743060008.JPEG.

55. Moscatelli, A., and Lacquaniti, F. (2011). The weight of time: Gravitational force enhances discrimination of visual motion duration. J. Vis. *11*, 5. https://doi.org/10.1167/11.4.5.

56. Ceccarelli, F., La Scaleia, B., Russo, M., Cesqui, B., Gravano, S., Mezzetti, M., Moscatelli, A., d'Avella, A., Lacquaniti, F., and Zago, M. (2018). Rolling motion along an incline: Visual sensitivity to the relation between acceleration and slope. Front. Neurosci. *12*, 406. https://doi.org/10.3389/FNINS.2018.00406/BIBTEX.

57. Delle Monache, S., Lacquaniti, F., and Bosco, G. (2019). Ocular tracking of occluded ballistic trajectories: Effects of visual context and of target law of motion. J. Vis. *19*, 13. https://doi.org/10.1167/19.4.13.

58. García-Saldivar, P., de León, C., Concha, L., and Merchant, H. (2022). White matter structural bases for predictive tapping synchronization. Preprint at bioRxiv. https://doi.org/10.1101/2022.09.05.506691.

59. Ivry, R.B., and Hazeltine, R.E. (1995). Perception and Production of Temporal Intervals Across a Range of Durations: Evidence for a Common Timing Mechanism. J. Exp. Psychol. Hum. Percept. Perform. *21*, 3–18. https://doi.org/10.1037/0096-1523.21.1.3.

60. Merchant, H., Luciana, M., Hooper, C., Majestic, S., and Tuite, P. (2008). Interval timing and Parkinson's disease: heterogeneity in temporal performance. Exp. Brain Res. *184*, 233–248. https://doi.org/10.1007/s00221-007-1097-7.

61. Zarco, W., Merchant, H., Prado, L., and Mendez, J.C. (2009). Subsecond Timing in Primates: Comparison of Interval Production Between Human Subjects and Rhesus Monkeys. J. Neurophysiol. *102*, 3191–3202.

62. Rajendran, V.G., Teki, S., and Schnupp, J.W.H. (2018). Temporal Processing in Audition: Insights from Music. Neuroscience *389*, 4–18. https://doi.org/10.1016/J.NEUROSCIENCE.2017.10.041.

63. Jazayeri, M., and Shadlen, M.N. (2010). Temporal context calibrates interval timing. Nat. Neurosci. *13*, 1020–1026. https://doi.org/10.1038/nn.2590.

64. Petzschner, F.H., Glasauer, S., and Stephan, K.E. (2015). A Bayesian perspective on magnitude estimation. Trends Cogn. Sci. *19*, 285–293.

65. Merchant, H., Battaglia-Mayer, A., and Georgopoulos, A.P. (2004). Neural responses in motor cortex and area 7a to real and apparent motion. Exp. Brain Res. *154*, 291–307. https://doi.org/10.1007/S00221-003-1664-5/FIGURES/13.

66. Lee, D.N., and Reddish, P.E. (1981). Plummeting gannets: a paradigm of ecological optics. Nature *293*, 293–294. https://doi.org/10.1038/293293a0.

67. Lacquaniti, F., and Maioli, C. (1989). The role of preparation in tuning anticipatory and reflex responses during catching. J. Neurosci. *9*, 134–148. https://doi.org/10.1523/JNEUROSCI.09-01-00134.

68. Lacquaniti, F., Bosco, G., Gravano, S., Indovina, I., La Scaleia, B., Maffei, V., and Zago, M. (2015). Gravity in the Brain as a Reference for Space and Time Perception. Multisens. Res. *28*, 397–426. https://doi.org/10.1163/22134808-00002471.

69. Indovina, I., Maffei, V., Bosco, G., Zago, M., Macaluso, E., and Lacquaniti, F. (2005). Representation of visual gravitational motion in the human vestibular cortex. Science *308*, 416–419. https://doi.org/10.1126/SCIENCE.1107961/SUPPL_FILE/INDOVINA.SOM.PDF.

70. Merchant, H., Crowe, D.A., Robertson, M.S., Fortes, A.F., and Georgopoulos, A.P. (2011). Top-down spatial categorization signal from prefrontal to posterior parietal cortex in the primate. Front. Syst. Neurosci. *5*, 69. https://doi.org/10.3389/FNSYS.2011.00069/BIBTEX.

71. Cavada, C., and Goldman-Rakic, P.S. (1993). Chapter 12 Multiple visual areas in the posterior parietal cortex of primates. Prog. Brain Res. *95*, 123–137. https://doi.org/10.1016/S0079-6123(08)60363-5.

72. Movshon, J.A., Lisberger, S.G., and Krauzlis, R.J. (1990). Visual Cortical Signals Supporting Smooth Pursuit Eye Movements. Cold Spring Harb. Symp. Quant. Biol. *55*, 707–716. https://doi.org/10.1101/SQB.1990.055.01.066.

73. Lisberger, S.G., and Movshon, J.A. (1999). Visual Motion Analysis for Pursuit Eye Movements in Area MT of Macaque Monkeys. J. Neurosci. *19*, 2224–2246. https://doi.org/10.1523/JNEUROSCI.19-06-02224.1999.

74. Schlack, A., and Albright, T.D. (2007). Remembering Visual Motion: Neural Correlates of Associative Plasticity and Motion Recall in Cortical Area MT. Neuron *53*, 881–890. https://doi.org/10.1016/J.NEURON.2007.02.028.

75. Price, N.S.C., Ono, S., Mustari, M.J., and Ibootson, M.R. (2005). Comparing acceleration and speed tuning in macaque MT: Physiology and modeling. J. Neurophysiol. *94*, 3451–3464. https://doi.org/10.1152/JN.00564.2005/ASSET/IMAGES/LARGE/Z9K0110550490014.JPEG.

76. Merchant, H., Zarco, W., Pérez, O., Prado, L., and Bartolo, R. (2011). Measuring time with different neural chronometers during a synchronization-continuation task. Proc Natl Acad Sci USA *108*, 19784–19789. https://doi.org/10.1073/PNAS.1112933108/SUPPL_FILE/PNAS.201112933SI.PDF.

77. Merchant, H., and Averbeck, B.B. (2017). The Computational and Neural Basis of Rhythmic Timing in Medial Premotor Cortex. J. Neurosci. *37*, 4552–4564. https://doi.org/10.1523/JNEUROSCI.0367-17.2017.

78. Fraisse, P. (1963). The Psychology of Time (Harper & Row).

79. Fraisse, P. (1978). Time and rhythm perception. Perceptual Coding, 203–254. https://doi.org/10.1016/B978-0-12-161908-4.50012-7.

80. McAuley, J.D., Jones, M.R., Holub, S., Johnston, H.M., and Miller, N.S. (2006). The time of our lives: Life span development of timing and event tracking. J. Exp. Psychol. Gen. *135*, 348–367. https://doi.org/10.1037/0096-3445.135.3.348.

81. Hirsh, I.J., and Watson, C.S. (1996). Auditory psychophysics and perception. Annu. Rev. Psychol. *47*, 461–484. https://doi.org/10.1146/ANNUREV.PSYCH.47.1.461.

82. Jones, M.R., and Mcauley, J.D. (2005). Time judgments in global temporal contexts. Percept. Psychophys. *67*, 398–417. https://doi.org/10.3758/BF03193320.

83. Sohn, H., Narain, D., Meirhaeghe, N., and Jazayeri, M. (2019). Bayesian Computation through Cortical Latent Dynamics. Neuron *103*, 934–947.e5. https://doi.org/10.1016/J.NEURON.2019.06.012.

84. Donnet, S., Bartolo, R., Fernandes, J.M., Cunha, J.P.S., Prado, L., and 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.

85. Wing, A.M., and Kristofferson, A.B. (1973). The timing of interresponse intervals. Percept. Psychophys. *13*, 455–460. https://doi.org/10.3758/BF03205802.

86. Wing, A.M., and Kristofferson, A.B. (1973). Response delays and the timing of discrete motor responses. Percept. Psychophys. *14*, 5–12. https://doi.org/10.3758/BF03198607.

87. Cicchini, G.M., Arrighi, R., Cecchetti, L., Giusti, M., and Burr, D.C. (2012). Optimal encoding of interval timing in expert percussionists. J. Neurosci. *32*, 1056–1060.

88. Petzschner, F.H., and Glasauer, S. (2011). Iterative Bayesian Estimation as an Explanation for Range and Regression Effects: A Study on Human Path Integration. J. Neurosci. *31*, 17220–17229. https://doi.org/10.1523/JNEUROSCI.2028-11.2011.

89. Jantzen, K.J., Ratcliff, B.R., and Jantzen, M.G. (2017). Cortical Networks for Correcting Errors in Sensorimotor Synchronization Depend on the Direction of Asynchrony. J. Mot. Behav. *50*, 235–248. https://doi.org/10.1080/00222895.2017.1327414.

90. Globerson, E., Flash, T., Eitan, Z., Globerson, E., Flash, T., and Eitan, Z. (2021). Space, Time and Expression in Orchestral Conducting, pp. 199–212. https://doi.org/10.1007/978-3-030-57227-3_10.

91. Deliège, I., Davidson, J., and Sloboda, J.A. (2011). Music and the Mind: Essays in Honour of John Sloboda (Oxford University Press).

92. Luck, G., and Sloboda, J.A. (2009). Spatio-temporal cues for visually mediated synchronization. Music Percept. *26*, 465–473.

93. Toh, L.W., Chao, W., and Chen, Y.S. (2013). An interactive conducting system using Kinect. SAVE Proc. https://doi.org/10.1109/ICME.2013.6607481.

94. Schramm, R., Jung, C.R., and Miranda, E.R. (2015). Dynamic time warping for music conducting gestures evaluation. IEEE Trans. Multimedia *17*, 243–255. https://doi.org/10.1109/TMM.2014.2377553.

95. Burger, B., Thompson, M.R., Luck, G., Saarikallio, S.H., and Toiviainen, P. (2014). Hunting for the beat in the body: On period and phase locking in music-induced movement. Front. Hum. Neurosci. *8*, 903. https://doi.org/10.3389/FNHUM.2014.00903/BIBTEX.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
MATLAB	MathWorks, USA	https://la.mathworks.com/products/matlab.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Hugo Merchant (hugomerchant@unam.mx).

Materials availability

This study did not generate any new unique reagents.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Sixteen Italian volunteers (eight females, age 22.94 years ±2.82 SD) were enrolled in the study. Experimental procedures were previously approved by the ethics committee of the University of Rome "Tor Vergata" (protocol #: 23/18), and participants signed a written consent form before participating in the experiment. Subjects were either right-handed or ambidextrous (one subject), according to the Edinburgh questionnaire, and had normal or corrected-to-normal vision.

METHOD DETAILS

Participants sat 35 cm in front of a 22-inch touchscreen (Elo Touch ET2201L, Elo TouchSystems, USA) where visual scenes were displayed with 1920 × 1080 pixel resolution and a refresh rate of 60 Hz. Scenarios were created using the MATLAB (2017b; Mathworks, USA) Psychtoolbox-3 toolbox.[59] Tapping responses were acquired via the USB interface of the touchscreen at a sampling frequency of 120 Hz. Eye movements were recorded with an Eyelink 1000 tracker system (SR Research Ltd., Canada) at a sampling frequency of 500 Hz.

Synchronization-continuation task

Subjects performed two variations of the Synchronization-Continuation Task (SCT). In one variation, they synchronized finger-tapping responses to events produced by moving visual objects (Moving metronome condition). In the second variation, which was used as a control condition, participants tapped in sync with the flashing of the same visual objects for 150 ms at fixed screen locations (Flashing metronome condition).

Each trial started when subjects placed their right index finger in the starting position at the center of the tapping area located at the bottom right corner of the touchscreen (see Figure 1). Then, stimuli were presented alternately in two positions as an isochronous metronome. After the first three instruction visual stimuli (observation epoch), the subjects produced six taps in synchrony with the metronome (synchronization epoch), first on the left red circle and then alternating between the two red circles in the tapping area. We used two stimuli and two response locations to generate a tapping task that simulated more naturalistic situations, such as a drummer playing the bongos, congas, or timbales. In addition, the alternating response locations were used to generate a congruent response to two falling objects in the vertical

condition. It is important to consider that the present results on tapping precision, accuracy, and prediction are similar to the ones reported in the literature for tapping to one location during an SCT,[7,24] suggesting that the use of two response locations slightly affects rhythmic tapping performance. After the visual metronome was extinguished, the subjects tapped six more times in the absence of any visual cue (continuation epoch). Five target interval durations (t_d) were used, from 450 to 850 ms in steps of 100 ms. During the task, subjects were required to maintain ocular fixation on designated points of the visual scenes (see below).

Vertical and horizontal scenarios

Visual metronomes were displayed in two different visual scenarios, which were designed specifically to represent either vertical or horizontal object motion (Vertical and Horizontal Scenarios, respectively). The two scenarios were presented in separate experimental sessions, performed at least 24 h apart, and their order of presentation was counter-balanced across subjects. In the Vertical Scenario, visual metronomes were rendered over a structured background scene reproducing a tropical beach, in which two palm trees surrounded by other naturalistic graphic elements provided cues for scaling the scene to real-world size (Figure 1B).

For the Moving metronome condition, two coconuts (9.4 visual degrees apart) were either dropped from both palm tree and bounced on the ground below the trees or were launched from the ground and bounced against the top branches of the palm trees. For each direction of motion, the coconuts' motion could be either accelerated or decelerated at a magnitude of 9.81 $m*s^{-2}$ scaled to the virtual scene size. The velocity at the bounce was 59 and 5 visual $degrees*sec^{-1}$ for accelerated and decelerated motion, respectively. After bouncing, coconuts moved at an oblique angle from the bouncing surface and disappeared quickly behind the trees by moving at a velocity between 18 and 21 visual $degrees*sec^{-1}$ (Figure 1D). In each trial, the two coconuts moved in the same direction and with the same kinematics (motion duration was fixed at 750 ms), but their motion onsets were shifted temporally at fixed amounts of time so that alternated bounces between the right (always occurring first) and the left coconut defined one of the five possible t_d of the visual metronome. Note that, in this scenario, downward accelerated and upward decelerated motion of the coconuts represented conditions compatible with the effects of natural gravity, whereas upward accelerated and downward decelerated motion were conditions incongruent with a natural setting. Thus, these conditions could be used to test for the effects of motion naturalness on the synchronization performance.

In the Flashing metronome condition, static images of the coconuts were flashed for 150 ms at the bouncing locations of the Moving metronome conditions, either on the ground or on the tree branches. In each trial, alternated flashing of the two visual objects started always from the right side and occurred at fixed intervals corresponding to one of the five t_d of the visual metronome (Figure 1F).

In the Horizontal Scenario, the two visual objects were represented by remotely controlled toy racing cars positioned upon two bookshelves, separated vertically by 9.4 visual degrees (see Figure 1C). In the Moving metronome condition, the toy cars ran along the upper and lower bookshelves, starting from one end of the bookshelves, and bounced against the pile of books at the opposite side of the bookshelves. The objects' velocity at the bounce was identical to that described for the vertically moving objects. After the bounce, cars moved obliquely from the bouncing surface at a velocity between 18 and 21 visual $degrees*sec^{-1}$ and disappeared quickly behind the book piles. In separate trials, the toy cars moved either rightward or leftward and their motion could be either accelerated or decelerated at 9.81 $m*s^{-2}$ scaled to the virtual scene size (Figure 1E). Like in the Vertical Scenario, the two cars had the same kinematics (motion duration = 750 ms), but their motion onsets were shifted temporally at fixed amounts of time so that alternate bounces between the cars on the bottom (occurring first) and on the top shelf defined one of the five t_d of the visual metronome.

Note that because the acceleration/deceleration profiles for the leftward/rightward motion of the two motorized vehicles could be equally plausible in a real-life setting, we did not use these conditions to test for the effect of motion naturalness. In the Flashing condition, static images of the toy cars were flashed for 150 ms at the bouncing locations of the Moving metronome condition, either at the left or at the right end of the bookshelves. In each trial, alternated flashing of the two visual targets started always from the bottom shelf and corresponded to one of the five possible t_d (Figure 1F).

During each SCT trial, subjects maintained ocular fixation on a dark green dot (diameter 1.1 visual degrees) positioned ~8 visual degrees from the visual metronome (see the examples shown in Figures 1B and 1C for visual metronomes presented at the bottom of the palm trees in the Vertical Scenario and at the right end of the bookshelves in the Horizontal Scenario). The fixation point color turned red between trials to instruct subjects they could temporarily break fixation.

Each experimental session comprised twenty Moving metronome conditions resulting from factoring two motion directions (UM | DM in the Vertical Scenario; RM | LM in the Horizontal Scenario), two motion accelerations (deceleration | acceleration) and five t_d (450 | 550 | 650 | 750 | 850 ms). Similarly, ten Flashing metronome conditions were obtained by factoring two flashing locations (top | bottom for the Vertical Scenario; right | left for the Horizontal Scenario) and five t_d (450 | 550 | 650 | 750 | 850 ms). A pseudorandom sequence of eight repetitions of these thirty conditions (240 trials total) was presented in six mini-blocks of forty trials each. The first two mini-blocks involved Moving metronome trials, grouped with respect to the visual objects' movement direction; the third mini-block included Flashing metronome trials presented at the location of the first Moving metronome mini-block. After a short break (~5 min), subjects performed the remaining two Moving metronome mini-blocks (presented in the same order as the first two mini-blocks) and the final mini-block with Flashing metronomes presented at the location of the second Moving metronome mini-block. The order of mini-blocks with different motion directions was counterbalanced across participants.

Data pre-processing. We evaluated the subjects' performance during the SCT by analyzing their tapping response times, acquired at 120 Hz from the touchscreen via a standard USB interface. The series of twelve taps produced during each trial delimited 11 time intervals. The first five produced intervals (t_p) represented the subjects' reproduction of the target duration (t_d) during the synchronization epoch, whereas the last five corresponded to the reproduction of the memorized t_d during the continuation epoch. The sixth interval, delimited by the last tap of the synchronization epoch and the first of the continuation epoch, was discarded. For each subject we computed the mean constant error (CE) and temporal variability (TV) of each t_p of the synchronization and continuation epochs across trials of the same experimental condition. CE can be defined as $t_p - t_d$ and represents a measure of timing accuracy. TV is a measure of timing precision and corresponds to the standard deviation of the t_p.

From the response times, we also determined the tap asynchronies; that is, the time differences between the taps and visual events cueing the subjects' responses during the synchronization epoch (object bouncing or flashing, depending on the type of visual metronome). Negative and positive asynchrony values indicated whether subjects anticipated or followed the cueing visual events. Then, for each subject we computed the mean and the standard deviation values of the asynchronies for each tap of the SCT series across trials of a given experimental condition.

These datasets were screened for trials during which subjects either broke ocular fixation or did not perform the tapping sequence correctly. Following this screening procedure, we discarded 127 out of 3840 trials (3.3%) in which ocular fixation deviated more than 2° of visual angle for longer than 200 ms; an additional 176 trials were discarded (4.6% of the total) because participants did not complete the series of 12 taps; 263 other trials (6.9% of the total) were discarded because the t_p was three standard deviations above the mean values observed in that experimental condition.

Custom-made scripts implemented in MATLAB (Mathworks, USA) were used for data pre-processing.

Slope analysis

The slope method is a classical timing model that uses a linear regression between TV as a function of target duration (t_d) to arrive at a generalized form of Weber's law (see Figure 2A). The resulting slope (slopeTV) is associated with the time-dependent process, capturing the scalar property of interval timing. The intercept (interceptTV) is related to the time-independent component, which is the fraction of variance that is relatively invariant across interval durations and is associated with sensory detection and processing, decision making, memory load, and/or motor execution.[59–61] As a convention, we computed the interceptTV at the intermediate target interval of 650 ms instead of at 0 ms, as usually computed in linear regression (Figure 2A). InterceptTV and slopeTV were computed for each subject and condition though linear regression.

Similarly, we regressed the CE on the t_d and defined two other measures to further characterize the SCT behavior: 1) the intercept at 650 ms (interceptCE), which is proportional to the indifference interval, namely, the interval where there is no error in timing[10,62]; 2) the slope of the regression line, called slopeCE, which corresponds to the magnitude of the bias effect (i.e., the larger the negative slope, the larger the regression toward the mean).[11,63,64]

Bayesian model

We used the target t_d as initial input data and produced intervals t_p. Our four-step model included a measurement process characterized by $p(t_m|t_s, b_m) = N(t_m|t_s, \sigma_m)$, an estimation process given by $t_e = f(t_m, \mu_t, \sigma_t) = max_t\{p(t|t_m)\}$, a production process described by $p(t_p|t_e, w_p) = N(t_p|t_e, w_p t_e)$, and finally a feedback process given by $t_s = t_d + \lambda(t_{p-1} - t_d)$. All these processes were summarized in the conditional probability:

$$p(t_p|t_d) = \iint [p(t_p|f(t_m))p(t_m|t_d + \lambda(t_{p-1} - t_d))p(t_{p-1}|t_d)]\, dt_m dt_{p-1}$$

where σ_m, μ_t, σ_t, w_p, and λ were the model parameters. Assuming independence of the pairs (t_d, t_p) across different trials, the total conditional probability could be written as $p(t_p^1, t_p^2, \ldots, t_p^n | t_d^1, t_d^2, \ldots, t_d^n) = \prod_i p(t_p^i \vee t_d^i)$.

We used this equation to maximize the likelihood probability for the different model parameters across all pairs of behavioral measures (t_d, t_p). Maximization was carried out with the MATLAB gradient-based method.

QUANTIFICATION AND STATISTICAL ANALYSIS

The parameters derived from the slope analysis ($slope_{TV}$, $intercept_{TV}$, $slope_{CE}$, $intercept_{CE}$) and Bayesian model (σ_m, μ_t, σ_t, w_p, λ) were analyzed with four different repeated measures ANOVA models. Two of these models were applied separately to data from either Moving or Flashing metronome conditions, the third (Static-Kinetic model) tested differences among acceleration, deceleration, and flashing conditions (thus, across the two types of visual metronomes), and the fourth (Gravity model) tested the effects of motion naturalness.

Specifically, for the Moving metronomes model, we included the following "within subject" factors: motion direction (upward | downward | leftward | rightward), motion kinematics (accelerated | decelerated), serial order of t_p (10 levels), and all their two-way interactions.

For the Flashing metronome condition, we considered the following "within subjects" factors: target flashing position (up | down | left | right), serial order of t_p (10 levels) and all their two-way interactions. The Static-Kinetic model was a one-way repeated measures ANOVA with a three-level factor (accelerated | decelerated | flashing). The Gravity model included data from the vertical session only, collapsed with respect to the naturalness of the vertical motion, and considered the following "within subjects" factors: gravity (natural | non-natural), serial order of t_p (10 levels), and their two-way interaction.

The datasets of mean and standard deviation asynchrony values pooled from the 16 participants were submitted to four repeated measures ANOVA models similar to those described above, with the difference that we considered the serial order of taps during the Synchronization epoch (6 levels) rather than the serial order of t_p, and that we included the "within subjects" factor target duration (t_d: 450 | 550 | 650 | 750 | 850 ms) to the Moving metronome, Flashing metronome and Gravity models.

Greenhouse-Geisser corrections were applied to the ANOVA factors' p values with a statistical significance cut-off of $p = 0.05$.

Finally, pairwise comparisons between levels of the ANOVA factors were evaluated by means of two-tailed paired t-tests, Bonferroni corrected (cut-off of $p = 0.05$).