UC Davis
UC Davis Previously Published Works

Title

The role of consciously timed movements in shaping and improving auditory timing

Permalink

Journal

ISSN

Authors

De Kock, Rose
Zhou, Weiwei
Datta, Poorvi
et al.

Publication Date

DOI

Author for correspondence:
Martin Wiener
e-mail: mwiener@gmu.edu

The role of consciously timed movements in shaping and improving auditory timing

Rose De Kock[1], Weiwei Zhou[1], Poorvi Datta[1], Wilsaan Mychal Joiner[1] and Martin Wiener[2]

[1]Department of Neurobiology, Physiology and Behaviour, University of California, Davis, CA, USA
[2]Department of Psychology, George Mason University, Fairfax, VA, USA

MW, 0000-0001-5963-5439

Our subjective sense of time is intertwined with a plethora of perceptual, cognitive and motor functions, and likewise, the brain is equipped to expertly filter, weight and combine these signals for seamless interactions with a dynamic world. Until relatively recently, the literature on time perception has excluded the influence of simultaneous motor activity, yet it has been found that motor circuits in the brain are at the core of most timing functions. Several studies have now identified that concurrent movements exert robust effects on perceptual timing estimates, but critically have not assessed how humans consciously judge the duration of their own movements. This creates a gap in our understanding of the mechanisms driving movement-related effects on sensory timing. We sought to address this gap by administering a sensorimotor timing task in which we explicitly compared the timing of isolated auditory tones and arm movements, or both simultaneously. We contextualized our findings within a Bayesian cue combination framework, in which separate sources of temporal information are weighted by their reliability and integrated into a unitary time estimate that is more precise than either unisensory estimate. Our results revealed differences in accuracy between auditory, movement and combined trials, and (crucially) that combined trials were the most accurately timed. Under the Bayesian framework, we found that participants' combined estimates were more precise than isolated estimates, yet were sub-optimal when compared with the model's prediction, on average. These findings elucidate previously unknown qualities of conscious motor timing and propose computational mechanisms that can describe how movements combine with perceptual signals to create unified, multimodal experiences of time.

1. Introduction

Motor control functions are critical to our survival in the world and diverse in nature, spanning multiple time scales and integrating a flood of neural signals to guide us through various tasks [1–3]. Important to movement is the monitoring of sensory information to update movement plans according to errors or environmental demands [4]. In addition, individuals often calibrate ongoing movements to amplify or suppress channels of sensory information according to their goals via 'active sensing' [5], reflecting the continuous and bidirectional nature of the relationship.

It follows that time perception, a high-level, cumulative evaluation of one or more sensory channels, is highly malleable in response to movement characteristics. Although time perception studies have largely excluded *concurrent* motor components, recent studies focused on arm movements reveal that the accuracy and precision of perceptual timing are affected by a number of movement characteristics such as direction, speed, distance and movement environment [6–10]. Notably, movement can also improve time perception. In a set of complementary experiments, auditory intervals were presented either during arm movements or in the absence of movement. For both temporal categorization and reproduction tasks, intervals encoded during movement were timed more precisely [11]. In a

separate temporal discrimination study using auditory intervals, timing precision was enhanced for intervals for which the stimulus onset was determined by the participant rather than passively presented [12]. The benefit of movement is also highlighted in animal behaviour. Rats trained to estimate a fixed interval to receive a reward learned to use stereotyped movements to enhance the accuracy and precision of their estimates [13].

Given the clear temporal benefit of movement, it is reasonable to ask whether and to what extent it is timed with different levels of accuracy and precision than the sensory channels that much of this work focuses on, and whether these differences can explain how the channels of information are combined. To address these questions, we contextualize our study under a framework of Bayesian cue combination, which posits that multiple channels of timing information are evaluated by their reliability and optimally integrated into a more precise estimate than either alone [14]. Accordingly, the mean of the combined estimates is predicted to gravitate towards the mean of the more precise (i.e. more influential) modality. This framework aligns with principles of multisensory cue combination [15–17], which places importance on signal reliability when combining multiple inputs. Neural data reflect this differential weighting via population responses in multisensory areas [18,19].

This evidence has offered insights upon which to build our understanding of the intersection of motor and timing processes [20]. However, little is known about how self-movements are timed without an added perceptual event. We note that motor control studies certainly include timing components, but the key caveat is that they examine implicit rather than explicit timing. For example, a task might require participants to synchronize their movements to a beat [21] or interact dynamically with a stimulus without probing their conscious evaluation of the passage of time. To the latter point, we refer to conscious awareness of time as explicitly measuring its passage, as distinct from implicit timing mechanisms [22], which may reflect internalizations of cause-and-effect relationships or correlated spatiotemporal measurements [23]. One preliminary study of interest [24] employed a unique paradigm to test explicit timing of durations that were implicitly encoded. Participants were trained on a 'skittles' task requiring them to hold and release a virtual ball to hit a target. Repeated practice led them to internalize an optimal duration range for which holding and releasing led to a successful trial. When tested in an explicit timing task, they exhibited a selective improvement for timing the target interval, while participants who did not play the game showed no benefit. Combined with the previous studies discussed, this supports the hypothesis that movements offer highly reliable temporal measurements that in turn improve timing of concurrent events and even future timing performance. This is further evidenced by increased timing acuity in motor 'experts' such as athletes and drummers [25,26].

Our goal in the current study was twofold: first, we sought to understand how duration of self-movement is evaluated, given that most movement-timing tasks have either focused solely on implicit timing or have not isolated movements from a concurrent perceptual event. Second, we sought to describe differences in motor and auditory timing, and importantly, how these sources of information are combined to form a unitary estimate. We found evidence that there are differences in timing accuracy between motor and auditory estimates, and that durations are both more accurately and precisely timed using both sources of information. Finally, we synthesized these results under a Bayesian cue combination framework to account for the pronounced benefit that resulted from combining motor and perceptual sources.

2. Methods

(a) Participants
We tested 20 right-handed participants (13 females, 7 males, mean age = 25.45 (9.17)). Handedness was confirmed by the Edinburgh Handedness Inventory [27].

(b) Procedure
Participants performed the experiment using a robotic arm manipulandum (KINARM End-Point Lab, BKIN Technologies [28,29]) that allowed movement along a flat workspace using the right arm. Direct viewing of the robotic arm was occluded by a flat display that allowed viewing of targets and cues via a downward-facing monitor mounted above the workspace. Motor output was sampled at 1000 Hz. Participants were free to adjust the chair so that they could comfortably view the full display.

Trials were divided into encoding and reproduction phases and were structured as follows (figure 1): first, the robotic arm guided participants to 1 of 16 locations in a grid-like array. Then, they experienced one of the three trial conditions. In 'movement' trials, subjects began moving until interrupted by an imposed brake (a 100 ms linear increase in resistive force from 0 to 50 N). In 'auditory' trials, the robotic arm was locked in the random location and the participant heard an auditory tone. In 'combined' trials, subjects were cued to move while timing a concurrent auditory tone; in this condition, the tone began as soon as the apparatus detected movement at the velocity threshold of $5 \, \text{cm s}^{-1}$, and the brake was applied synchronously with the auditory tone offset. After the encoding phase, they were guided to a central target for the reproduction phase. When this target turned green, subjects reproduced the encoded duration by holding and releasing a button attached to the handle. The tested durations were 1000, 1500, 2000, 2500, 3000, 3500 and 4000 ms. Trial conditions were experienced in blocks of 14 trials (for a total of 210 trials) in a pseudorandomized order such that no condition was experienced twice in a row.

(c) Analysis
Robotic arm manipulandum data were sampled at 1000 Hz to produce vectors for position, velocity, force and other movement parameters over the course of time for each trial. Trials were excluded if reproduced times fell outside three deviations from the mean (<1% of trials excluded). In addition, trials with movement (movement-only and combined) were excluded if the stop latency after the brake was applied fell outside three scaled absolute deviations from the median (2.9% of trials excluded).

Our first goal was to investigate the relationship between duration and reproduced time to determine timing accuracy. Importantly, movement-only and combined trials were analyzed with respect to time spent moving rather than the pre-specified duration. This is because, in general, participants exhibited a short delay to respond to the movement brake, usually adding up to a few hundred milliseconds on most trials. We directly assessed constant error as a measure of accuracy, defined as the difference between the reproduced duration using the button press and the actual target duration. To measure performance, we applied a linear mixed model (LMM) design, in which response
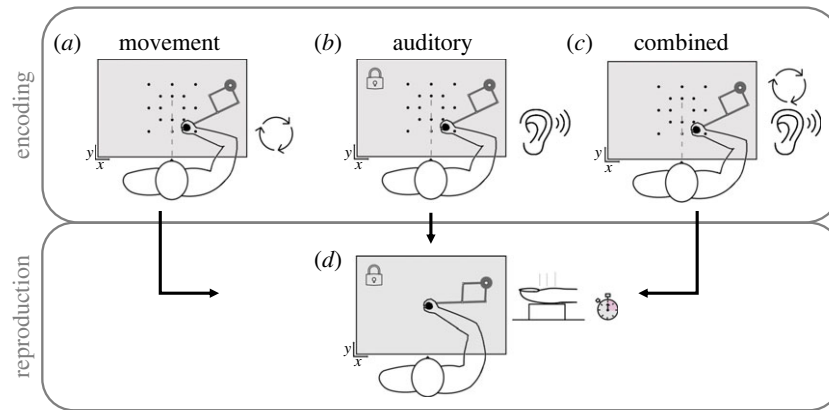
Figure 1. Task schematic. (*a*) Movement-only trial. The robotic arm guided subjects to a random location (dots not visible during task), and the participant moved freely until interrupted by a 'brake'. (*b*) Auditory-only trial. The robot handle was locked in random location, and the participant listened to an auditory tone. (*c*) Combined movement-auditory trial. The robotic arm was guided to a random location, and the participant moved freely and listened to the auditory tone until the 'brake' was applied along with the tone offset. (*d*) Reproduction phase. The interval was estimated by holding and releasing a button attached to the handle.

error was the predicted variable, encoded duration and condition were fixed effects, and subject was treated as a random effect.

In addition, we examined the relationships between movement parameters and timing performance. Previous works have shown that arm movements covering a greater distance lead to longer perceived durations [11,30]. Here, we extracted the movement distance for each trial (defined as the Euclidean distance travelled between duration onset and offset) and performed a Spearman partial correlation test between movement distance and reproduced time, controlling for target duration. Partial correlation values were calculated for each subject separately for movement and combined trial types; each was assessed with a one-sample t-test against a baseline of zero. Multiple comparisons were corrected using the Bonferroni correction.

Statistical analyses were performed using R, JASP (http://www.jasp-stats.org), and Matlab. For accuracy and coefficient of variation (CV) analyses, we report results from LMMs with subject as a random effect. Results are reported at a significance level of 0.05.

(d) Computational modelling

To determine the sensitivity of each of the unisensory modality conditions, as well as the multisensory combined one, we employed a Bayesian observer–actor model (figure 4) previously described by Remington *et al.* [31] (see also [32]) and used previously by our group [30]. In this model, sample durations (t_s) are inferred as draws from noisy measurement distributions (t_m) that scale in width according to the length of the presented interval. These measurements, when perceived, may be offset from veridical estimates as a result of perceptual bias or other outside forces (*b*). Due to the noise in the measurement process, the brain combines the perceived measurement with the prior distribution of presented intervals in a statistically optimal manner to produce a posterior estimate of time (t_e). The mean of the posterior distribution is then, in turn, used to guide the reproduced interval (t_p), corrupted by production noise (*p*). The resulting fits to this model thus produce an estimate of the measurement noise (*m*), the production noise (*p*) and the offset shift in perceived duration (*b*). Note that the offset term is also similar to that employed for other reproduction tasks as a shift parameter [33]. In addition, the prior used for the model can be either uniform or Gaussian in shape [26], with implications for how these are combined; uniform priors are characterized by the range of intervals presented, whereas Gaussian priors are centred on the average duration presented, with a width dependent on their precision. We chose here to model the prior as a Gaussian, as each individual subject will have experienced a slightly different set of intervals during the

estimation phase for each of the three conditions. This is because in the two movement conditions, the offsets were variable from trial to trial. As such, we modelled the width of individual subject priors to match the width of encoded intervals for each subject (σt_s). Model fits were conducted by minimizing the negative log-likelihood of subject responses given the sample values using Matlab's *fminsearch* function.

To determine if subjects combined auditory and movement modalities in an optimal manner, consistent with cue combination, we used outputs of the Bayesian model to compare between unisensory and multisensory conditions. Specifically, cue combination predictions that the multisensory combination σ_C of two unisensory estimates (auditory σ_A and movement σ_M) when modelled as Gaussians should equal:

$$\sigma_c = \frac{\sigma_A \sigma_M}{\sigma_A + \sigma_M}.$$

That is, the combined width is the product of the unisensory widths divided by their sum. Since the measurement widths were modelled as Gaussians in our Bayesian model, the unisensory widths from model fits can be used in the aforementioned equation to produce an estimate of the predicted width. Further, as the model also provides a width for the multisensory, combined condition, we can compare the width observed with that predicted by cue combination [17,34]. If the predicted and observed widths *are not statistically different*, then subjects combined unisensory estimates optimally, whereas larger observed widths than predicted indicate sub-optimal combination [35].

3. Results

The temporal reproduction group data are illustrated in figure 2. As described earlier, movement-only and combined trials were analyzed with respect to time spent moving, which we defined as the predetermined interval (1000, 1500, 2000, 2500, 3000, 3500 and 4000 ms) plus the 'stop latency' (figure 2*d*), defined as the time it took participants to stop movement after the brake was applied. We used a linear mixed effects model with trial type (movement and combined levels only) and duration as fixed effects and subject as a random effect to test for stop latency differences related to trial type or duration. There were no significant main effects of trial type [$F(1, 2650.03) = 0.056$, $p = 0.812$] or duration [$F(1, 2650.36) = 0.014$, $p = 0.906$], and no significant interaction [$F(1, 2650.01) = 1.622$, $p = 0.203$]. In the rest of the figures, the
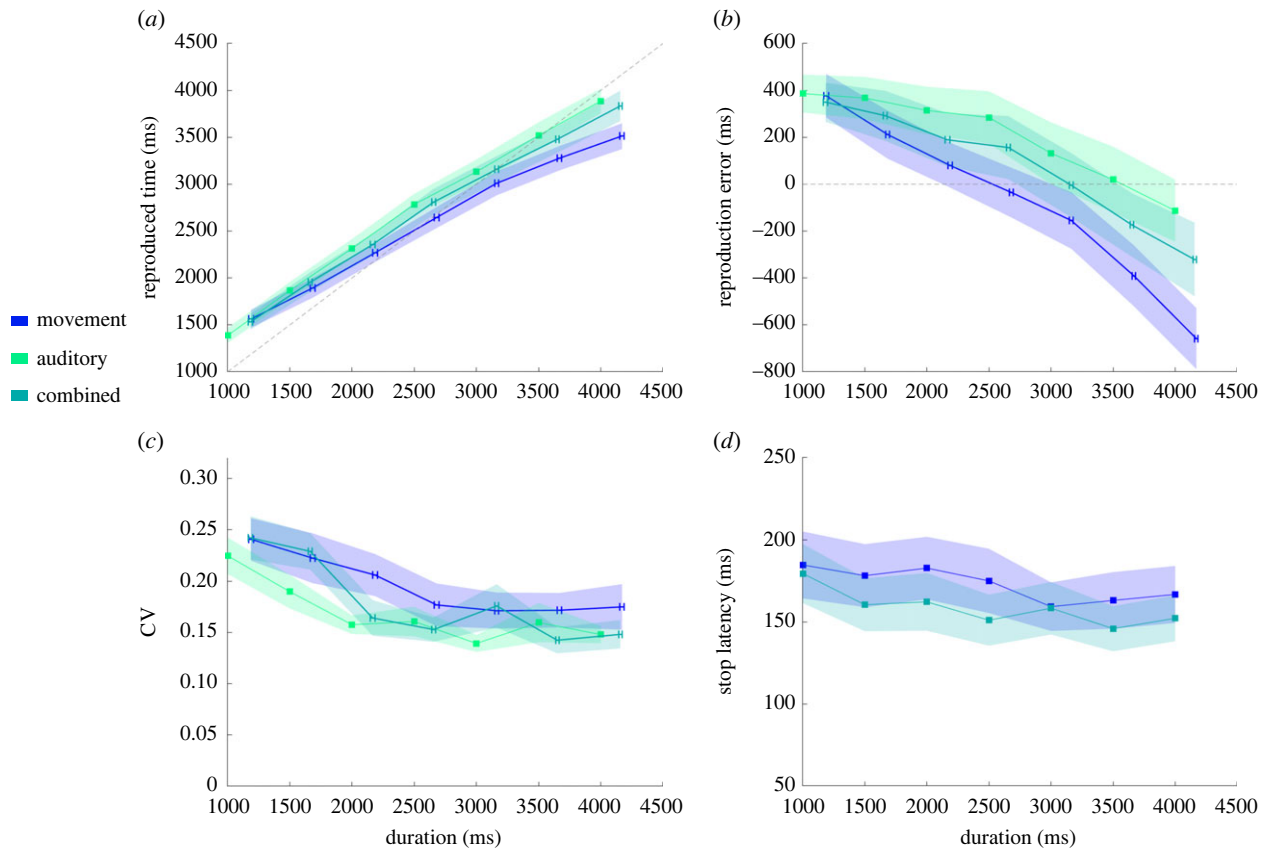
Figure 2. Temporal reproduction results. (*a*) Reproduction performance plotted as a function of target duration and trial type. Shaded vertical error bars represent reproduction standard error, and horizontal bars represent standard error of time spent moving for movement and combined trials. (*b*) Constant error values (reproduced duration − target duration) as a function of target duration and trial type. (*c*) Coefficient of variation (std/mean) of reproduction time as a function of target duration and trial type. (*d*) Stop latencies of movement and combined trials when the movement brake was applied. These values plus the fixed target durations (*x*-axis) determined target durations for trials with movement.

target durations for movement and combined trials include these stop latencies, which are shown with horizontal error bars to denote the standard error (figure 2*a*–*c*).

We next employed a linear mixed effects model with trial type and duration as fixed effects and subject as a random effect to characterize how these variables affected reproduced duration. We did this analysis using constant error values (reproduced duration − target duration) rather than raw reproduction values as the dependent variable, as they represent the same underlying data but provide information about the direction and magnitude of errors (figure 2*b*). Model terms were nested using the Satterthwaite method. The model revealed a significant interaction of duration and trial type [$F(2, 4045.04) = 30.577$, $p < 0.001$], and main effects of trial type [$F(2, 4045.05) = 5.512$, $p = 00.004$] and duration [$F(1, 4045.15) = 821.134$, $p < 0.001$]. We examined the estimated marginal means and contrasts of trial types with Holm-adjusted *p*-values, which revealed a significant difference between auditory (EMM = 180.185, CI[−25.003, 385.374]) and movement (EMM = −51.534, CI[−256.860, 153.791]) trials (contrast = 231.720, SE = 20.915, $p < 0.001$), combined (EMM = 81.551, CI[−123.649, 286.751]) and auditory trials (contrast = −98.635, SE = 20.592, $p < 0.001$), and finally, between combined and movement trials (contrast = 133.085, SE = 20.940, $p < 0.001$). We also compared the estimated marginal means to zero and did not find a significant result for any of the trial types ($p_{auditory} = 0.085$, $p_{movement} = 0.623$, $p_{combined} = 0.436$).

Our next analysis focused on the error slopes produced by the different trial types. This analysis reveals the degree

of central tendency (i.e. attraction of estimates towards the mean). A significant slope difference was found between all trial types pairs (auditory − movement = 0.161, SE = 0.021, $p < 0.001$; combined − auditory = −0.061, SE = 0.021, $p = 0.003$; combined − movement = 0.101, SE = 0.021, $p < 0.001$). Movement trials exhibited the lowest slope, and auditory trials exhibited the highest slope.

Next, we examined the CV as a measure of reproduction precision across durations and trial types. We employed a LMM with trial type and duration as fixed effects and subject as a random effect and found that the CV varied as a function of duration [$F(1, 395.18) = 64.227$, $p < 0.001$] but not the trial type [$F(2, 395.05) = 2.575$, $p = 0.077$]. The interaction of the trial type and duration was not significant [$F(2, 395.05) = 0.893$, $p = 0.410$].

We were additionally interested in the relationship between movement parameters and reproduced time, particularly for the combined trial type. We assessed this effect by performing subject-level partial correlations between Euclidean movement distance during duration encoding and the subsequent reproduced duration, controlling for target duration. The distribution of individual correlation coefficients is displayed in figure 3. By using a one-sample t-test, we found that the values were distributed significantly above zero for both motor [$t(19) = 2.974$, $p = 0.008$, Cohen's $D = 0.665$] and combined [$t(19) = 4.278$, $p < 0.001$, $D = 0.957$] conditions, indicating a positive relationship between movement distance and reproduced duration, and replicating prior work that movement distances are associated with longer estimated durations [11,30].
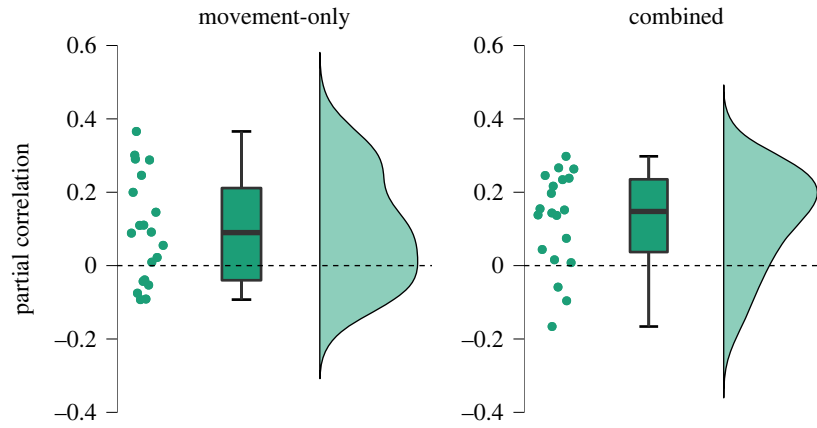
Figure 3. Effects of movement on duration estimates. Spearman partial correlation coefficients are displayed for each subject as raincloud plots in both the unisensory, movement-only condition and the multisensory, combined condition. For each subject, partial correlations were calculated as the association between single-trial reproduced durations and the movement length during the encoding phase, while controlling for duration. On average, the correlation coefficients in both groups were significantly greater than zero.
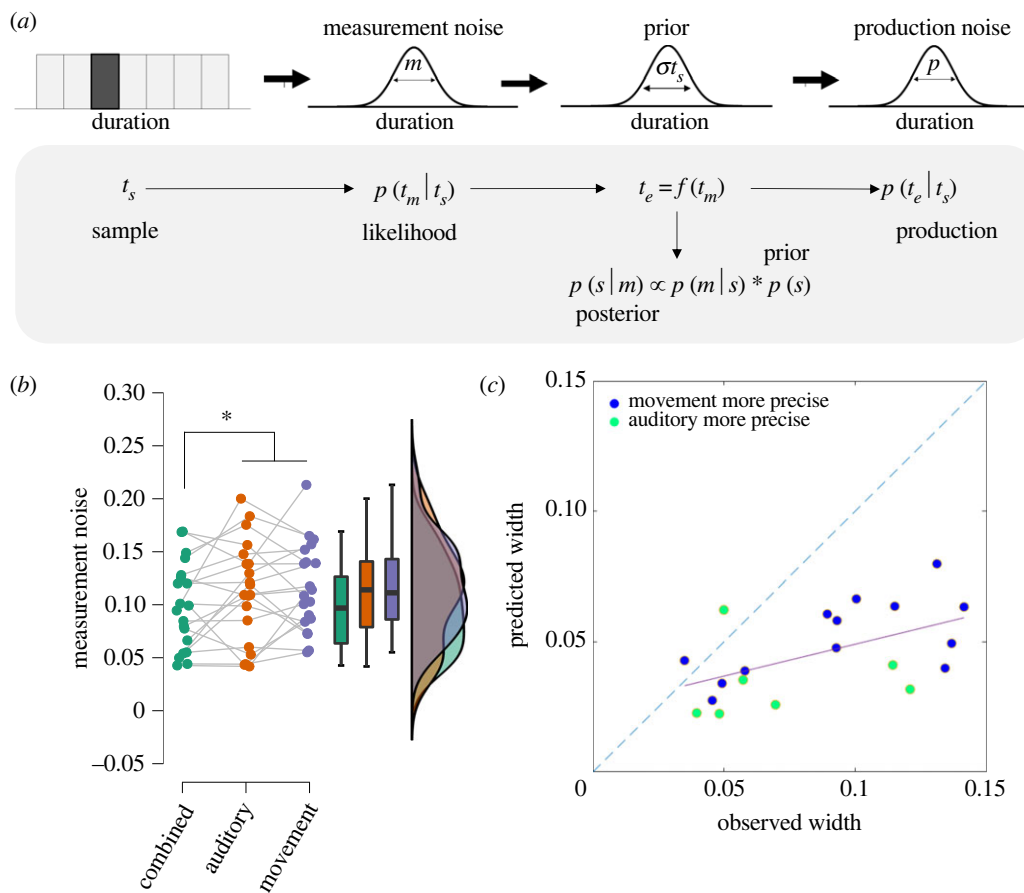

Figure 4. Bayesian observer–actor model for time estimates. (a) Schematic of the model, in which a sample interval presented on a given trial t_s is perceived with some measurement noise from a Gaussian distribution with width m. This estimate is then combined with a prior distribution of previously experienced intervals, also characterized as a Gaussian, with width σt_s to form a posterior estimate t_e. During reproduction, the posterior estimate is further corrupted by motor production noise as a Gaussian distribution with width p. For model σt_s, m and p were set as free parameters, whereas σt_s was measured directly from the experienced sample intervals for each subject. (b) Model fits for the measurement noise width (m) for all three modality conditions. We observed that the multisensory, combined measurement noise was significantly lower than both the unisensory auditory and movement conditions, together. (c) Scatterplot comparing the observed multisensory width to the predicted width given by combining both of the unisensory widths together via cue combination. Here, we observed that the observed widths fell on average below the predicted width, yet were also significantly correlated. Individual points are coloured based on which of the unisensory widths was smaller (more precise); no differences were found between either of these groups.

(a) Cue combination

To examine the impact of modality on time estimates, we further fit reproduced durations with a Bayesian observer–actor model. The results of our model fits provided estimates of the measurement noise (m), production noise (p), and offset (b) for each modality. Repeated-measure analyses of variance for each parameter revealed only a main effect of condition for measurement noise [$F(2, 38) = 2.133$, $p = 0.039$],

with all other parameters being non-significant (all $p < 0.05$). For the noise parameter, measurement noise scores were significantly lower for the combined condition compared to movement [$t(19) = -2.118$, $p = 0.048$], but not auditory intervals [$t(19) = -1.682$, $p = 0.109$] (figure 4). Due to our *a priori* hypothesis that the combined multisensory estimates would be better than both unisensory estimates, we averaged auditory and movement measurement widths and compared them to the combined measurement noise, where a significant difference was observed [$t(19) = -2.15$, $p = 0.045$].

To compare with the predictions of the cue combination model, we used the unisensory measurement noise widths to generate a predicted width for their optimal combination. That is, the product of the unisensory widths divided by their sum. Here, we observed that the predicted widths were on average lower than the observed widths from the combined multisensory condition [$t(19) = 6.111$, $p < 0.001$] (figure 4). In addition, we observed a strong correlation between observed and predicted widths [$r(18) = 0.654$, $p = 0.002$], indicating that the cue combination equation provided a good prediction of multisensory noise, even if the observed estimates were sub-optimal. As a further observation, we noted that subjects with lower multisensory noise estimates were closer to the optimal prediction, a finding we confirmed quantitatively through the correlation of multisensory widths with the difference between those widths and the predicted estimate [$r(18) = -0.883$, $p < 0.001$]. As a final check, we compared subjects who exhibited smaller unisensory widths for one modality (e.g. auditory) over the other (e.g. movement). No differences were observed between subjects with greater precision in either modality, for either the multisensory combined estimates [$t(18) = 0.765$, $p = 0.454$] or the difference from optimality [$t(18) = -0.528$, $p = 0.604$], indicating that improved estimates for neither unisensory modality conveyed a special benefit in cue combination.

4. Discussion

We administered a temporal reproduction experiment in which we tested timing performance for auditory and motor timing, and both simultaneously. Measuring these two sources of information allowed us to assess how self-movements are judged in comparison to auditory tones and shed light on the computational mechanisms that drive movement-related improvements previously observed during perceptual timing tasks [9,11]. We found that modality significantly impacted reproduction performance, such that motor trials resulted in shorter estimates than auditory trials. Critically, combined trials were estimated most accurately, suggesting that the natural biases introduced by auditory and motor intervals oppose each other, but 'work together' to form the most veridical estimate based on the information available. In addition, our results fit into a framework of Bayesian cue combination in which multimodal interval measurements are more precise than unisensory measurements [14]. We also found that the reproduction slope was lowest for the movement condition, indicating a greater degree of central tendency compared to the other conditions. An unanswered question is the extent to which this reflects intrinsic properties of movement interval timing, such as a greater reliance on an internalized prior distribution [26] or a general susceptibility towards underestimation as target durations increase.

These results corroborate precious accounts of movement enhancing cross-modal timing [11–13,36,37]. However, it is of interest whether this is accomplished in an optimal fashion. According to the Bayesian cue combination framework, we predicted that the measurement noise (i.e. distribution width) of the combined condition would be lower than for unimodal conditions. The model results indicated that the measurement noise of the combined condition was significantly lower than for movement and auditory conditions together. While we did not observe meaningful differences in coefficient of variation (an index of variability that is often tied to precision [38,39]) between trial types, we suggest that the measurement noise parameter is a more useful indicator of precision in our experiment given that the target durations were not fixed for trials that involved movement, and further that this measure attempts to remove motor production noise. We next compared the observed measurement noise to the model prediction that would indicate optimally combined estimates. The observed and predicted values were significantly positively correlated, although we note that participants generally combined sub-optimally. This pattern was not dependent on which unimodal condition was more precise for individual participants. The model also predicts that during optimal multimodal timing, the mean reproduction estimate should gravitate towards the more precise modality; however, given that we did not find overall differences in unimodal timing precision across participants and performance was generally sub-optimal, this prediction was challenging to test in the current paradigm. However, we note that sub-optimality in human perception is a commonly observed feature, which may depend on multiple other factors beyond movement [35]. Remarkably, this lack of difference highlights that movement timing is at least as precise as auditory timing, which until now has been documented as the most precisely timed modality [40–42]. Future work may assess this prediction more closely where larger differences between unimodal conditions exist (e.g. visual timing).

We have described some computational principles by which motor and sensory information may be combined for a more precise multimodal estimate. These perspectives are strengthened by discussing their relation to neural mechanisms. There is a great degree of functional overlap between motor and timing activity in brain regions considered vital to motor control, with greater representation of supra-second intervals in cortical regions and sub-second intervals in subcortical regions like the cerebellum [43,44]. The supplementary motor area (SMA) stands out as a region of interest, as it is activated across a wide range of timing tasks and encodes time intervals in neurons organized along a rostrocaudal 'chronotopic' gradient [45]. In addition, the SMA exists within a larger cortico-thalamic-basal ganglia timing circuit [20] that encodes intervals, integrates multiple sensory inputs [46], and sends predictive signals to sensory areas.

By using these insights, we have outlined two possibilities to describe the neural implementation of movement-related timing effects [14]. The first possibility, feedforward enhancement, posits that these effects are instantiated in motor circuits and sharpen duration measurements via corollary feedback to motor regions like the SMA, thereby sharpening the tuning of duration-selective neurons. The SMA is equipped to respond to these signals via white matter connections with the primary motor cortex, basal ganglia and spinal cord [47],

all of which require a high degree of temporal coordination during movement control. In addition, deep reinforcement learning agents with feedforward modules (and no recurrence) successfully learn to produce temporal intervals by generating stereotyped trajectories in their environments [48].

The second possibility, active sensing [5], proposes that motor activity acts on earlier sensory regions to enhance cross-modal temporal measurements. Outside of the timing domain, this is a well-established process. For example, auditory perception is enhanced when a sound is triggered by an action [49], and accordingly, motor preparation has been found to elicit responses in the auditory cortex [50]. There are also several examples tying motor activity to changes in visual processing [51–55]. The active sensing hypothesis is compatible with prior research on multisensory integration, as it essentially describes the convergence of multiple signals to shape neural computations. A classic example is the superior colliculus, which integrates visual, auditory, and sensorimotor signals (among others) to guide eye and head movements [56,57]. These signals are not localized to dedicated hubs, but rather occur throughout the neocortex [58], and have been proposed to reflect generalizable 'canonical operations' (e.g. divisive normalization and oscillatory phase resets) when integrating a diverse range of inputs—including from motor circuits [59]. For example, saccade onsets elicit time-locked local field potential changes in primary visual cortex [60]. The active sensing hypothesis has empirical support from several lines of research and can build upon known neural integration mechanisms—more generalizable than previously thought—to shed light on how movement can improve timing.

We have primarily discussed our results with a focus on the ability of motor activity to shape auditory timing. However, it can also be said that the enhancement is bidirectional (i.e. movement timing improves with added auditory information). In a related fashion, one study found that auditory intervals trained in a temporal discrimination task selectively improved timing precision for the same intervals in a subsequent motor task [61]. These behavioural effects are paralleled by brain circuitry, particularly in the dorsal auditory stream with bidirectional connections between hubs in the auditory cortex, inferior parietal lobule, and premotor cortex [62]. This pathway has been highlighted as essential to speech and music perception and production, especially for auditory predictive processes [62–64]. Further, the dorsal auditory pathway uses predictive mechanisms to support beat perception as described in the action simulation for auditory prediction (ASAP) hypothesis [65,66]. Although our study's focus was on interval timing rather than beat timing, the dorsal pathway's important role in auditory perception may have implications for our observed results [67].

Beyond basic neuroscience, this work has implications in clinical disorders associated with motor and timing deficits such as Parkinson's disease and Huntington's disease [68–71]. These parallel deficits are not restricted to movement disorders, but occur in psychiatric or neurodevelopmental conditions such as schizophrenia and attention deficit hyperactivity disorder [72,73]. Motor training has shown some usefulness in rehabilitation and symptom management; Parkinson's patients have been found to reduce their gait variability when exposed to rhythmic auditory stimuli that can adaptively synchronize with their steps [74]. Critically, this improvement may depend on the optimality with which patients incorporate these two sources of temporal information, and so by adopting a baseline measurement, one could track improvement over time. In stroke patients, fine motor skills are re-learned more effectively with musical motor training than functional motor training [75]. Thus, many benefits of movement training rely strongly on integrating relevant sensory information, and based on this evidence, interactions between motor and auditory interval timing may be a promising avenue to explore in the treatment and diagnosis of movement disorders.

In conclusion, while converging evidence suggests a powerful role of movement in shaping time perception [14,76], studies have focused primarily on perceptual timing with movement as an added component without isolating how movements are consciously timed on their own (i.e. how movements are incorporated when subjects are explicitly engaged in estimating time [22]). Our experiment provided the distinct advantage of isolating and comparing timing in movement and auditory modalities, in addition to testing predictions about their integration under a Bayesian cue combination framework. We found that multisensory timing was superior to unisensory timing as reflected in the higher accuracy of combined estimates, and in addition, measurement noise of combined estimates reflected at least some degree of optimal cue combination (with the caveat that participants often perform sub-optimally compared to computational models). This study thus addresses a prior gap in knowledge where consciously timed self-movement were not well understood, especially in relation to timing in other modalities. We also expanded on a growing body of research on movement-timing effects, first by describing potential computational mechanisms that drive them, and ways they may be instantiated in neural circuits.

References

1. Cisek P, Kalaska JF. 2010 Neural mechanisms for interacting with a world full of action choices. *Annu. Rev. Neurosci.* **33**, 269–298. (doi:10.1146/annurev. neuro.051508.135409)

2. Alhussein L, Smith MA. 2021 Motor planning under uncertainty. *Elife* **10**, e67019. (doi:10.7554/eLife.67019)

3. Kaplan HS, Thula OS, Khoss N, Zimmer M. 2020 Nested neuronal dynamics orchestrate a behavioral hierarchy across timescales. *Neuron* **105**, 562–576. e9. (doi:10.1016/j.neuron.2019.10.037)

4. Franklin DW, Wolpert DM. 2011 Computational mechanisms of sensorimotor control. *Neuron* **72**, 425–442. (doi:10.1016/j.neuron.2011.10.006)

5. Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P. 2010 Dynamics of active sensing and perceptual selection. *Curr. Opin. Neurobiol.* **20**, 172–176. (doi:10.1016/j.conb.2010.02.010)

6. Tomassini A, Morrone MC. 2016 Perceived visual time depends on motor preparation and direction of hand movements. *Sci. Rep.* **6**, 27947. (doi:10.1038/srep27947)

7. Yokosaka T, Kuroki S, Nishida SY, Watanabe J. 2015 Apparent time interval of visual stimuli is compressed during fast hand movement. *PLoS ONE* **10**, e0124901. (doi:10.1371/journal.pone.0124901)

8. Yon D, Edey R, Ivry RB, Press C. 2017 Time on your hands: perceived duration of sensory events is biased toward concurrent actions. *J. Exp. Psychol. Gen.* **146**, 182–193. (doi:10.1037/xge0000254)

9. De KR, Zhou W, Joiner WM, Wiener M. 2021 Slowing the body slows down time perception. *Elife* **10**, e63607. (doi:10.7554/eLife.63607)

10. Manning FC, Schutz M. 2016 Trained to keep a beat: movement-related enhancements to timing perception in percussionists and non-percussionists. *Psychol Res.* **80**, 532–542. (doi:10.1007/s00426-015-0678-5)

11. Wiener M, Zhou W, Bader F, Joiner WM. 2019 Movement improves the quality of temporal perception and decision-making. *ENeuro* **6**, ENEURO.0042-19.2019. (doi:10.1523/ENEURO.0042-19.2019)

12. Iordanescu L, Grabowecky M, Suzuki S. 2013 Action enhances auditory but not visual temporal sensitivity. *Psychon. Bull. Rev.* **20**, 108–114. (doi:10.3758/s13423-012-0330-y)

13. Safaie M, Jurado-Parras MT, Sarno S, Louis J, Karoutchi C, Petit LF, Pasquet MO, Eloy C, Robbe D. 2020 Turning the body into a clock: accurate timing is facilitated by simple stereotyped interactions with the environment. *Proc. Natl Acad. Sci. USA* **117**, 13 084–13 093. (doi:10.1073/pnas.1921226117)

14. De Kock R, Gladhill KA, Ali MN, Joiner WM, Wiener M. 2021 How movements shape the perception of time. *Trends Cogn. Sci.* **25**, 950–963. (doi:10.1016/j.tics.2021.08.002)

15. Seilheimer RL, Rosenberg A, Angelaki DE. 2014 Models and processes of multisensory cue combination. *Curr. Opin. Neurobiol.* **25**, 38–46. (doi:10.1016/j.conb.2013.11.008)

16. Ma WJ, Beck JM, Latham PE, Pouget A. 2006 Bayesian inference with probabilistic population codes. *Nat. Neurosci.* **9**, 1432–1438. (doi:10.1038/nn1790)

17. Alais D, Burr D. 2019 Cue combination within a Bayesian framework. In *Multisensory processes: the auditory perspective* (eds AKC Lee, MT Wallace, AB Coffin, AN Popper, RR Fay), pp. 9–31. New York, NY: Springer International Publishing.

18. Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE. 2011 Neural correlates of reliability-based cue weighting during multisensory integration. *Nat. Neurosci.* **15**, 146–154. (doi:10.1038/nn.2983)

19. Gu Y, Angelaki DE, DeAngelis GC. 2008 Neural correlates of multisensory cue integration in macaque MSTd. *Nat. Neurosci.* **11**, 1201–1210. (doi:10.1038/nn.2191)

20. Merchant H, Yarrow K. 2016 How the motor system both encodes and influences our sense of time. *Curr. Opin. Behav. Sci.* **8**, 22–27. (doi:10.1016/j.cobeha.2016.01.006)

21. Repp BH, Su YH. 2013 Sensorimotor synchronization: a review of recent research (2006-2012). *Psychon. Bull. Rev.* **20**, 403–452. (doi:10.3758/s13423-012-0371-2)

22. Coull JT, Nobre AC. 2008 Dissociating explicit timing from temporal expectation with fMRI. *Curr. Opin. Neurobiol.* **18**, 137–144. (doi:10.1016/j.conb.2008.07.011)

23. Walsh V. 2003 A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* **7**, 483–488. (doi:10.1016/j.tics.2003.09.002)

24. Guo J, Zhang Z, Sternad D, Song JH. 2019 Improved motor timing enhances time perception. *J. Vis.* **19**, 218b. (doi:10.1167/19.10.218b)

25. Chen YH, Cesari P. 2015 Elite athletes refine their internal clocks. *Motor Control* **19**, 90–101. (doi:10.1123/mc.2013-0081)

26. Cicchini GM, Arrighi R, Cecchetti L, Giusti M, Burr DC. 2012 Optimal encoding of interval timing in expert percussionists. *J. Neurosci.* **32**, 1056–1060. (doi:10.1523/JNEUROSCI.3411-11.2012)

27. Oldfield RC. 1971 The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* **9**, 97–113. (doi:10.1016/0028-3932(71)90067-4)

28. Nguyen KP, Zhou W, McKenna E, Colucci-Chang K, Bray LC, Hosseini EA, Alhussein L, Rezazad M, Joiner WM. 2019 The 24-h savings of adaptation to novel movement dynamics initially reflects the recall of previous performance. *J. Neurophysiol.* **122**, 933–946. (doi:10.1152/jn.00569.2018)

29. Hosseini EA, Nguyen KP, Joiner WM. 2017 The decay of motor adaptation to novel movement dynamics reveals an asymmetry in the stability of motion state-dependent learning. *PLoS Comput. Biol.* **13**, e1005492. (doi:10.1371/journal.pcbi.1005492)

30. De Kock R, Zhou W, Joiner WM, Wiener M. 2021 Slowing the body slows down time perception. *Elife* **10**, e63607. (doi:10.7554/eLife.63607)

31. Remington ED, Parks TV, Jazayeri M. 2018 Late Bayesian inference in mental transformations. *Nature Comm.* **9**, 1–13.

32. Jazayeri M, Shadlen MN. 2010 Temporal context calibrates interval timing. *Nature Neurosci.* **13**, 1020–1026.

33. Petzschner FH, 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.

34. Hartcher-O'Brien J, Di Luca M, Ernst MO. 2014 The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. *PLoS ONE* **9**, e89339.

35. Rahnev D, Denison RN. 2018 Suboptimality in perceptual decision making. *Behav. Brain Sci.* **41**, e223. (doi:10.1017/S0140525X18000936)

36. Carlini A, French R. 2014 Visual tracking combined with hand-tracking improves time perception of moving stimuli. *Sci. Rep.* **4**, 5363. (doi:10.1038/srep05363)

37. Manning F, Schutz M. 2013 Moving to the beat improves timing perception. *Psychon. Bull. Rev.* **20**, 1133–1139. (doi:10.3758/s13423-013-0439-7)

38. Brown SW. 1997 Attentional resources in timing: interference effects in concurrent temporal and nontemporal working memory tasks. *Percept. Psychophys.* **59**, 1118–1140. (doi:10.3758/BF03205526)

39. Lewis PA, Miall RC. 2009 The precision of temporal judgement: milliseconds, many minutes, and beyond. *Phil. Trans. R. Soc. B* **364**, 1897–1905. (doi:10.1098/rstb.2009.0020)

40. Wiener M, Thompson JC, Coslett HB. 2014 Continuous carryover of temporal context dissociates response bias from perceptual influence for duration. *PLoS ONE* **9**, e100803. (doi:10.1371/journal.pone.0100803)

41. Jones LA, Poliakoff E, Wells J. 2009 Good vibrations: human interval timing in the vibrotactile modality. *Q. J. Exp. Psychol. (Hove)* **62**, 2171–2186. (doi:10.1080/17470210902782200)

42. Burr D, Banks MS, Morrone MC. 2009 Auditory dominance over vision in the perception of interval duration. *Exp. Brain Res.* **198**, 49–57. (doi:10.1007/s00221-009-1933-z)

43. Nani A, Manuello J, Liloia D, Duca S, Costa T, Cauda F. 2019 The neural correlates of time: a meta-analysis of neuroimaging studies. *J. Cogn. Neurosci.* **31**, 1796–1826. (doi:10.1162/jocn_a_01459)

44. Wiener M, Turkeltaub P, Coslett HB. 2010 The image of time: a voxel-wise meta-analysis. *Neuroimage* **49**, 1728–1740. (doi:10.1016/j.neuroimage.2009.09.064)

45. Protopapa F, Hayashi MJ, Kulashekhar S, van der Zwaag W, Battistella G, Murray MM, Kanai R, Bueti D. 2019 Chronotopic maps in human supplementary motor area. *PLoS Biol.* **17**, e3000026. (doi:10.1371/journal.pbio.3000026)

46. Nagy A, Eördegh G, Paréczy Z, Márkus Z, Benedek G. 2006 Multisensory integration in the basal ganglia. *Eur. J. Neurosci.* **24**, 917–924. (doi:10.1111/j.1460-9568.2006.04942.x)

47. Vergani F, Lacerda L, Martino J, Attems J, Morris C, Mitchell P, De Schotten MT, Dell'Acqua F. 2014 White matter connections of the supplementary motor area in humans. *J. Neurol. Neurosurg. Psychiatry.* **85**, 1377–1385. (doi:10.1136/jnnp-2013-307492)

48. Deverett B, Faulkner R, Fortunato M, Wayne G, Leibo JZ. 2019 *Interval timing in deep reinforcement learning agents.* See https://proceedings.neurips.cc/

paper/2019/hash/2bf283c05b601f21364d052ca0ec798d-Abstract.html.

49. Myers JC, Mock JR, Golob EJ. 2020 Sensorimotor integration can enhance auditory perception. *Sci. Rep.* **10**, 1496. (doi:10.1038/s41598-020-58447-z)

50. Gale DJ, Areshenkoff CN, Honda C, Johnsrude IS, Flanagan JR, Gallivan JP. 2021 Motor planning modulates neural activity patterns in early human auditory cortex. *Cereb Cortex* **31**, 2952–2967. (doi:10.1093/cercor/bhaa403)

51. Yon D, Gilbert SJ, de Lange FP, Press C. 2018 Action sharpens sensory representations of expected outcomes. *Nat. Commun.* **9**, 4288. (doi:10.1038/s41467-018-06752-7)

52. Tomassini A, Maris E, Hilt P, Fadiga L, D'Ausilio A. 2020 Visual detection is locked to the internal dynamics of cortico-motor control. *PLoS Biol.* **18**, e3000898. (doi:10.1371/journal.pbio.3000898)

53. Saleem AB, Ayaz A, Jeffery KJ, Harris KD, Carandini M. 2013 Integration of visual motion and locomotion in mouse visual cortex. *Nat. Neurosci.* **16**, 1864–1869. (doi:10.1038/nn.3567)

54. Niell CM, Stryker MP. 2010 Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* **65**, 472–479. (doi:10.1016/j.neuron.2010.01.033)

55. Cao L, Händel B. 2019 Walking enhances peripheral visual processing in humans. *PLoS Biol.* **17**, e3000511. (doi:10.1371/journal.pbio.3000511)

56. Stein BE. 1998 Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Exp. Brain Res.* **123**, 124–135. (doi:10.1007/s002210050553)

57. Distler C, Hoffmann KP. 2015 Direct projections from the dorsal premotor cortex to the superior colliculus in the macaque (*Macaca mulatta*). *J. Comp. Neurol.* **523**, 2390–2408. (doi:10.1002/cne.23794)

58. Ghazanfar AA, Schroeder CE. 2006 Is neocortex essentially multisensory? *Trends Cogn. Sci.* **10**, 278–285. (doi:10.1016/j.tics.2006.04.008)

59. Van Atteveldt N, Murray MM, Thut G, Schroeder CE. 2014 Multisensory integration: flexible use of

general operations. *Neuron* **81**, 1240–1253. (doi:10.1016/j.neuron.2014.02.044)

60. Ito J, Maldonado P, Singer W, Grün S. 2011 Saccade-related modulations of neuronal excitability support synchrony of visually elicited spikes. *Cereb Cortex* **21**, 2482–2497. (doi:10.1093/cercor/bhr020)

61. Meegan DV, Aslin RN, Jacobs RA. 2000 Motor timing learned without motor training. *Nat. Neurosci.* **3**, 860–862. (doi:10.1038/78757)

62. Rauschecker JP. 2011 An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hear Res* **271**, 16–25. (doi:10.1016/j.heares.2010.09.001)

63. Friederici AD. 2012 The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cogn. Sci.* **16**, 262–268. (doi:10.1016/j.tics.2012.04.001)

64. Zatorre RJ, Chen JL, Penhune VB. 2007 When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* **8**, 547–558. (doi:10.1038/nrn2152)

65. Patel AD, Iversen JR. 2014 The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Front. Syst. Neurosci.* **8**, 57. (doi:10.3389/fnsys.2014.00057)

66. Cannon JJ, Patel AD. 2021 How beat perception co-opts motor neurophysiology. *Trends Cogn. Sci.* **25**, 137–150. (doi:10.1016/j.tics.2020.11.002)

67. Teki S, Grube M, Griffiths TD. 2011 A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Front. Integr. Neurosci.* **5**, 90. (doi:10.3389/fnint.2011.00090)

68. Avanzino L, Pelosin E, Vicario CM, Lagravinese G, Abbruzzese G, Martino D. 2016 Time processing and motor control in movement disorders. *Front. Hum. Neurosci.* **10**, 631. (doi:10.3389/fnhum.2016.00631)

69. Singh A, Cole RC, Espinoza AI, Evans A, Cao S, Cavanagh JF, Narayanan NS. 2021 Timing variability and midfrontal 4 Hz rhythms correlate

with cognition in Parkinson's disease. *NPJ Parkinsons Dis.* **7**, 14. (doi:10.1038/s41531-021-00158-x)

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

71. Cope TE, Grube M, Singh B, Burn DJ, Griffiths TD. 2014 The basal ganglia in perceptual timing: timing performance in multiple system atrophy and Huntington's disease. *Neuropsychologia* **52**, 73–81. (doi:10.1016/j.neuropsychologia.2013.09.039)

72. Walther S, Strik W. 2012 Motor symptoms and schizophrenia. *Neuropsychobiology* **66**, 77–92. (doi:10.1159/000339456)

73. Yang B, Chan RC, Zou X, Jing J, Mai J, Li J. 2007 Time perception deficit in children with ADHD. *Brain Res.* **1170**, 90–96. (doi:10.1016/j.brainres.2007.07.021)

74. Miyake Y. 2009 Interpersonal synchronization of body motion and the walk-mate walking support robot. *IEEE Trans. Rob.* **25**, 638–644. (doi:10.1109/TRO.2009.2020350)

75. Schneider S, Münte T, Rodriguez-Fornells A, Sailer M, Altenmüller E. 2010 Music-supported training is more efficient than functional motor training for recovery of fine motor skills in stroke patients. *Music Percept.* **27**, 271–280. (doi:10.1525/mp.2010.27.4.271)

76. Balasubramaniam R, Haegens S, Jazayeri M, Merchant H, Sternad D, Song JH. 2021 Neural encoding and representation of time for sensorimotor control and learning. *J. Neurosci.* **41**, 866–872. (doi:10.1523/JNEUROSCI.1652-20.2020)

77. De Kock R, Zhou W, Datta P, Mychal Joiner W, Wiener M. 2023 The role of consciously timed movements in shaping and improving auditory timing. Dryad Digital Repository. (doi:10.25338/B8FD0N)