scientific reports

Check for updates

OPEN

Perturbation context in paced finger tapping tunes the error-correction mechanism

Ariel D. Silva[1,2] & Rodrigo Laje[1,2,3]✉

Sensorimotor synchronization (SMS) is the mainly specifically human ability to move in sync with a periodic external stimulus, as in keeping pace with music. The most common experimental paradigm to study its largely unknown underlying mechanism is the paced finger-tapping task, where a participant taps to a periodic sequence of brief stimuli. Contrary to reaction time, this task involves temporal prediction because the participant needs to trigger the motor action in advance for the tap and the stimulus to occur simultaneously, then an error-correction mechanism takes past performance as input to adjust the following prediction. In a different, simpler task, it has been shown that exposure to a distribution of individual temporal intervals creates a "temporal context" that can bias the estimation/production of a single target interval. As temporal estimation and production are also involved in SMS, we asked whether a paced finger-tapping task with period perturbations would show any time-related context effect. In this work we show that a perturbation context can indeed be generated by exposure to period perturbations during paced finger tapping, affecting the shape and size of the resynchronization curve. Response asymmetry is also affected, thus evidencing an interplay between context and intrinsic nonlinearities of the correction mechanism. We conclude that perturbation context calibrates the underlying error-correction mechanism in SMS.

In the study of time processing in the scale of hundreds of milliseconds, traditionally known as millisecond timing, the description of behavior and the understanding of its neural mechanisms have shown great advances in the last two decades[1–4]. The need for considering the processing of time in the multiple time scales involved in motor planning and control is being recognized as a fundamental step towards understanding the interaction between brain and body[5].

Sensorimotor synchronization (SMS), the ability to move in sync with a periodic external stimulus, belongs to this time scale and underlies many human-specific activities like music and dance[6,7]. The simplest task to study SMS is paced finger tapping where a participant is instructed to tap in pace with an external metronome (usually a periodic sequence of brief tones). The stimulus sequence may be perturbed by an unexpected change of interstimulus period. This task allowed researchers to make the first steps towards identifying the neural correlates of SMS, like finding a perceptual and a sensorimotor component during resynchronization after a period perturbation[8]and an asymmetry in the processing of positive vs negative errors that reproduces a known behavioral asymmetry[9]. Advances in a related experimental paradigm with monkeys are to be noted, where single-neuron recordings showed clearly temporal profiles[10]and metronomical neural activity was found along the whole hierarchy of sensory, associative, and motor areas[11]. In a related paradigm where monkeys had to synchronize saccades to periodic visual stimuli[12], activity of individual neurons in the cerebellar dentate gyrus showed correlation with either the current synchronization error or the timing of the next saccade.

Despite these recent important advances, many fundamental issues remain unsolved. In this work we focus on the underlying error-correction mechanism that allows a person to recover average synchrony after an unexpected period perturbation in a paced finger-tapping task. In SMS research it is universally accepted that both keeping average synchrony and recovering it after a perturbation are achieved thanks to an error-correction mechanism[6,9,13,14]. Despite a long history of using linear models to describe the resynchronization behavior[7], mounting evidence from analyses of the resynchronization time series under different perturbation types and signs suggests that the error-correction mechanism is intrinsically nonlinear[9,13–15]. Other experimental findings also call for nonlinear models[16–21]. This knowledge could influence the way in which we design experiments and analyze the data.

[1]Sensorimotor Dynamics Lab, Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Bernal, Argentina. [2]CONICET, Buenos Aires, Argentina. [3]Departamento de Computación, Universidad de Buenos Aires, Buenos Aires, Argentina. ✉email: rlaje@unq.edu.ar

In temporal experimental paradigms other than paced finger tapping, it has been demonstrated that time estimation and production are biased by many factors[22]. We are interested in the effect of temporal context. In a traditional interval timing paradigm where participants are asked to estimate a single time interval demarcated by two flashed stimuli and reproduce it with a key press (a time reproduction task), it was shown that time estimates for a particular duration depended on the distribution of intervals presented to the participant[23], with a bias towards the mean of the distribution. The processing of time in this range is thus "calibrated" by exposure to different time interval distributions that create a temporal context biasing estimations.

In recent years the effect of temporal context was found also in SMS-related experimental paradigms, as the following examples show. During the synchronization phase in a synchronization-continuation paradigm[18,24] participants showed the "central tendency effect", i.e. a bias towards the mean as in the temporal context effect discussed above. Participants biased their estimation of the interstimulus interval depending on the distribution it was drawn from (towards the mean of the distribution), and the magnitude of the bias decreased along the trial as more stimuli were presented. In a different work, participants in a synchronization-continuation experiment[25] showed a similar central tendency effect. The difference between the presented interstimulus interval and the produced interresponse intervals in a trial (i.e. the tempo-matching error) showed a negative slope as a function of the interstimulus interval. That is, participants tapped too slow when stimuli had a fast rate, and too fast when stimuli had a slow rate (although the researchers pooled synchronization and continuation data to compute the tempo-matching error, so the actual result for synchronization specifically may be obscured).

Although the evidence points to the existence of a time-related context in time processing both for single time intervals and for synchronization to a periodic sequence, up to our knowledge there is no direct evidence to date that the error-correction mechanism underlying SMS is influenced or calibrated by it. In this work we first analyze a compilation of quantitative evidence from different sources suggesting that the timing of the first tap after a perturbation depends on context, despite having the same timing of previous responses and the same timing of stimuli up to that point. This novel time-related context appears to be generated by exposure to either a single perturbation type during the whole experiment or two perturbation types randomly alternated. Motivated by this, we then show results from our specific experiment proving the existence of such a perturbation context in SMS and its effect during the resynchronization phase.

Results

Figure 1a shows a schematic of the task for the two perturbation types considered in this work. The period of the stimuli sequence is T, also referred to as the interstimulus interval (ISI). The asynchrony at step n is defined as $e_n = R_n - S_n$, that is the difference between the occurrence times of response R_n and stimulus S_n. In

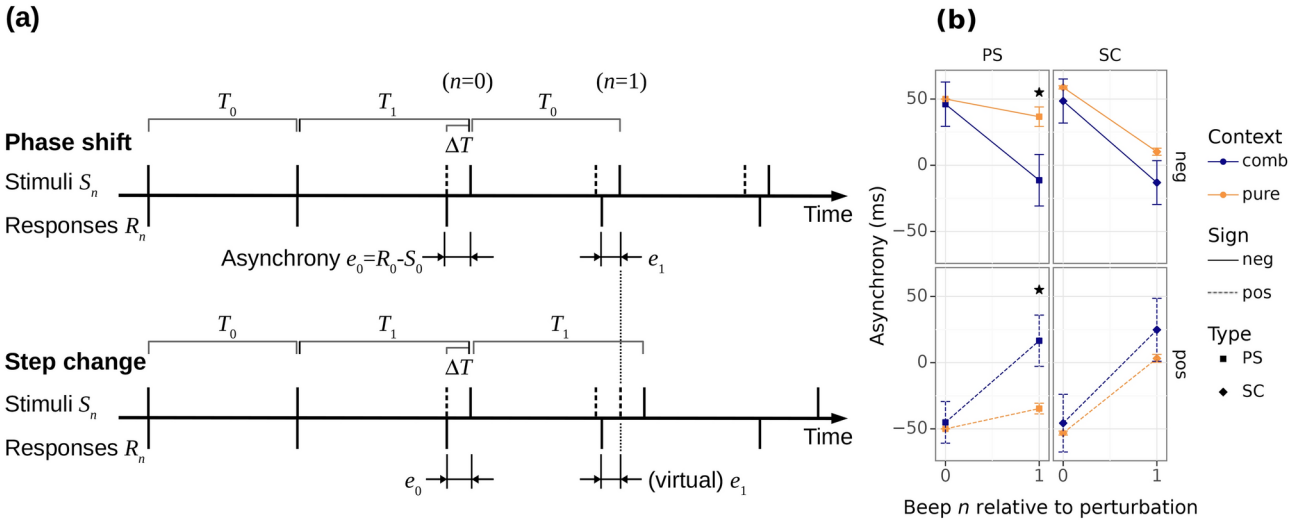

Fig. 1. **(a)** Schematic of a paced finger-tapping task with period perturbations. For simplicity the pre-perturbation asynchrony values are zero. As the time of perturbation is unexpected, at $n = 0$ there is a forced error for both perturbation types and resynchronization ensues. At $n = 1$ the underlying error-correction mechanism is already at work based on past performance; if there were no effect of context, the occurrence time of the response (in absolute time) should be the same for both perturbation types as shown. **(b)** Prior evidence from the paced finger tapping literature. Asynchrony after perturbation estimated from published works for the two main types of perturbations (SC: step change; PS: phase shift) and the two perturbation signs (pos: $\Delta T > 0$, period lengthens; neg: $\Delta T < 0$, period shortens) (mean $\pm 95\%$ confidence interval). For SC perturbations the virtual asynchrony is used as defined in panel a. The PS perturbations show diverging values at $n = 1$ for pure versus combined contexts, suggesting a potential effect of context (asterisks indicate significant differences after Bonferroni correction; see Methods). Data digitized from published works as described in Methods, subsection Sample Size Justification.

step-change (SC) perturbations the baseline period T_0 of the sequence is unexpectedly changed at $n = 0$ by an amount ΔT; the asynchrony at $n = 0$ thus shows a forced error equal to $-\Delta T$ (on average). In phase-shift (PS) perturbations the change at $n = 0$ is followed by an opposite shift at $n = 1$, thus making the sequence recover its original period. In both cases resynchronization follows as a convergence of asynchrony values to the post-perturbation baseline[14].

Prior evidence of perturbation context influencing the error-correction mechanism

In this subsection we compile prior evidence gathered from the paced finger tapping literature to show that exposure to either type of perturbation may generate a perturbation context effect that modifies the resynchronization response. The evidence is the following and it is based on comparing the response timing after SC and PS perturbations. We digitized data from published papers with paced finger tapping (Figure 1b; see Methods, subsection Sample Size Justification) where participants were exposed to either a single perturbation type ("pure") or two different perturbation types at random during the same experiment ("combined").

In experiments where the perturbations are unexpected (i.e., they occur at a random stimulus in a trial) the asynchrony at the perturbed beep $n = 0$ represents a forced error and, on average, it equals the time interval by which the corresponding stimulus was shifted, with opposite sign (Figure 1a). The response at the following beep ($n = 1$) is already affected by the correction mechanism attempting to decrease the forced error. Comparing asynchrony values from different perturbations, however, should be done carefully because the timing of the post-perturbation stimuli S_n is not the same between SC and PS[15]—that is, the time references of each asynchrony at $n \geq 1$ are different for different perturbations as shown in Figure 1a. In order to correctly compare and determine whether the post-perturbation response timing differs between different perturbation types, for the SC perturbations we show the "virtual" asynchrony between response and the extrapolated stimulus (as defined in panel a) thus fixing the time reference across perturbation types.

Let's assume for a moment that there is no context effect; that is, assume that the correction mechanism after a specific perturbation type has the same calibration whether it is a "pure" or a "combined" experiment. In this case, the responses at $n = 1$ should coincide in absolute time (same virtual asynchrony) because both the past performance and the stimulus sequence were identical up to that point. What is observed in the PS perturbations, instead, is that they differ (Figure 1b).

Isolating the effect of perturbation context

The departure between pure and combined PS at $n = 1$ shown in Figure 1b suggests there is an effect of perturbation context produced by exposure to either one or the two types of perturbation. However, we gathered data from different papers where the experiments were not designed to specifically test this. In addition, the traditional way of comparing in the literature is one perturbation type against the other which involves a potential confounding because the two perturbation types have an effect by themselves, both at the perceptual level (different saliencies) and at the behavioral level (different shifting of stimuli from $n = 1$ on, thus of time references)[6]. In order to isolate the potential effect of perturbation context and make the correct comparison to determine a causal relationship, in this work we exposed 74 participants to different perturbation types according to the following four groups: Group 1 was exposed to SC perturbations only; Group 2 was exposed to PS perturbations only; Group 3 was exposed to SC and PS perturbations in random order (larger perturbation size); Group 4 was exposed to SC and PS perturbations in random order (smaller perturbation size). Groups 1 and 2 belong to the "pure" context while Groups 3 and 4 belong to the "combined" context. In this way we can compare responses between different contexts in the same experiment while keeping the perturbation type constant.

The experiment is paced finger tapping with auditory stimuli and period perturbations to determine whether perturbation context has any effect during the resynchronization phase after a perturbation. The participant is instructed to keep average synchrony as well as possible and, in case a perturbation arises, to recover synchrony without stopping tapping. Perturbations can be of two different types: step-change perturbations ("SC", where the stimuli period abruptly changes by an amount ΔT at a random beep in the sequence) or phase-shift perturbations ("PS", where the period changes at two consecutive beeps by opposite amounts ΔT and $-\Delta T$ such that it goes back to its original value). These perturbation types are two of the most traditionally used in the paced finger tapping literature[14]. A perturbation can have two different signs: positive, where the period increases at the perturbed beep ($\Delta T > 0$, "pos"), or negative, where the period decreases at the perturbed beep ($\Delta T < 0$, "neg"); and two different sizes: $|\Delta T| = 50$ ms (larger perturbation size) and $|\Delta T| = 20$ ms (smaller perturbation size). Importantly, as described above in our experiment the participants are also exposed to either context: "pure" (Groups 1 and 2) or "combined" (Groups 3 and 4). The final experimental design is a fully factorial combination of conditions Context x Perturbation Type x Perturbation Sign x Perturbation Size; see Methods for a detailed description. The number of participants was determined by an a priori statistical power analysis using published data as shown in Figure 1b (see Methods, subsection Sample Size Justification). We first present results from the larger perturbation size (the one used for power analysis); we use the smaller perturbation size to show that the findings persist for different conditions.

Perturbation context affects the resynchronization response

Figure 2 shows the averaged time series of asynchrony for every condition, as a function of beep number along the trial (renumbered such that the perturbation occurs at $n = 0$) for the larger perturbation size. The resynchronization behavior is qualitatively similar across all conditions. As the perturbations are unexpected due to the actual perturbation occurring at a random beep in the trial, at $n = 0$ the asynchrony shows a forced error that on average is opposite to the perturbation size. Then the participant gradually recovers synchrony and reaches a new baseline.

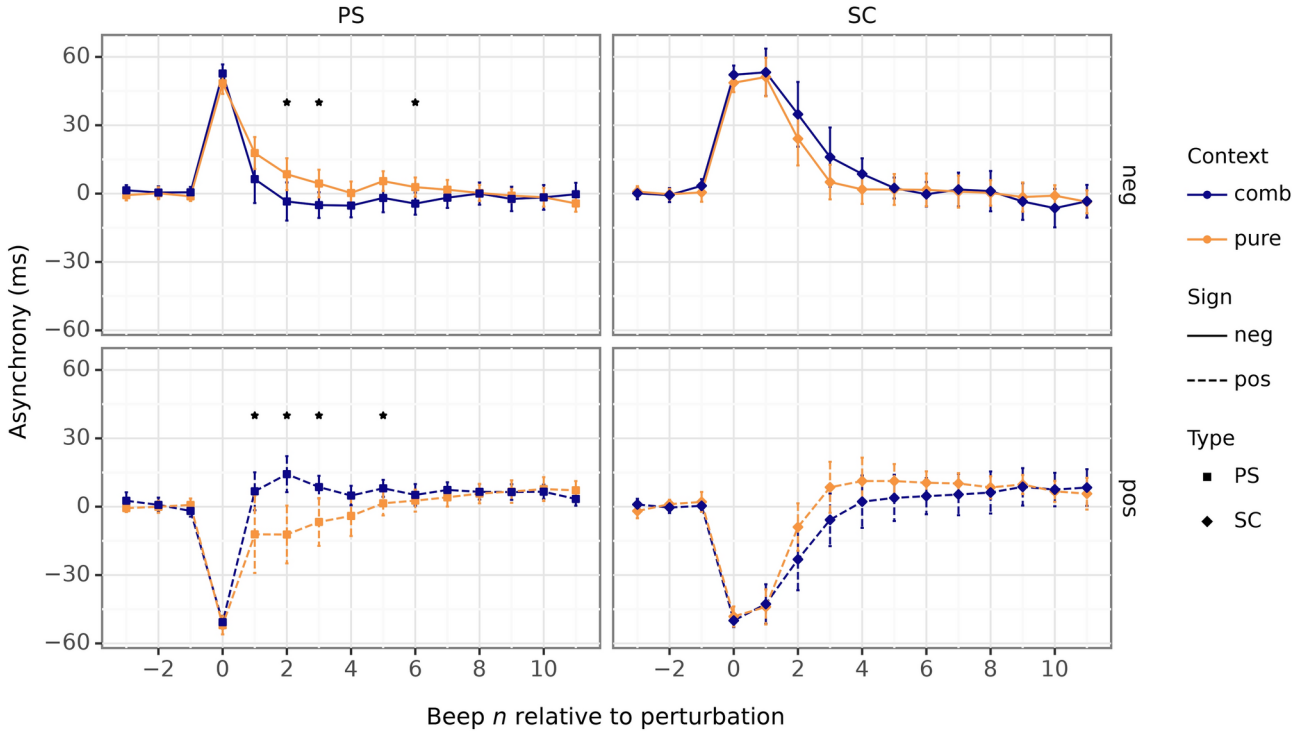

Fig. 2. Effect of Context on resynchronization after a perturbation. Asynchrony as a function of stimulus number along a trial ($n = 0$ is the perturbation beep). There are systematic differences between contexts (blue and orange curves): switching from combined to pure context shifts the asynchrony curve during resynchronization upwards or downwards depending on the condition, which is represented by a significant triple interaction Context x Perturbation Type x Perturbation Sign. Mean across participants $\pm 95\%$ confidence interval. Asterisks mark significant differences after permutation testing with FDR correction for multiple comparisons (see Methods).

Quantitatively, however, the time series show systematic differences between pure and combined context conditions during the resynchronization phase. Expanding on the findings from prior literature, the differences in our experiment go well beyond $n = 1$ and consequently differences in SC perturbations also appear that were not evident in the data from the literature. An exploratory analysis reveals that some of the observed differences between contexts remain significant after adjusting for multiple comparisons (asterisks from permutation testing across participants and trials, with FDR correction across beeps; see Methods).

In order to properly test the effect of Context we fitted a Linear Mixed Model (LMM) with Asynchrony as dependent variable (restricted to the resynchronization phase, $n = 1$ through 6, pooled) and factors Context (levels pure/combined), Perturbation Type (levels PS/SC), and Perturbation Sign (levels pos/neg). As we are interested in the effect of Context—either as a main effect or any of its interactions—we included the two-way interactions of Context and the three-way interaction. Participant was included as a random effect; see detailed model specification in Methods. Results indicate that there is a significant three-way interaction effect, meaning that the effect of Context depends on the type and sign of the perturbation (significant three-way interaction: $\chi^2(2) = 87.3$, $p = 2 \times 10^{-16}$; see full model results in Methods). This can be interpreted by observing Figure 2 where the three-way interaction shows as a neat disordinal, nearly crossover interaction: in the negative PS perturbations, "combined" Context produces lower Asynchrony values than "pure" (difference estimate between combined and pure: 8.8 ms; 95% CI: [1.9; 15.7] ms, corrected $p = 0.025$), while in the positive PS perturbations it produces higher Asynchrony values (difference: -14.5 ms, 95% CI: $[-22.1; -6.9]$ ms, corrected $p = 0.001$); whereas the opposite is true for the SC perturbations (SCneg difference: -4.8 ms, 95% CI: $[-11.9; 2.3]$ ms, corrected $p = 0.19$; SCpos difference: 7.8 ms, 95% CI: [0.6; 15.1] ms, corrected $p = 0.048$; see Methods for a summary of post-hoc comparisons). Effect sizes are shown in Supplementary Fig. S1 to graphically support the estimated marginal means comparisons.

An alternative model (the same three factors and all two-way interactions, but no three-way interaction) performs significantly worse, meaning the three-way interaction represents an important aspect of the data (see Methods).

It is important to note that all comparisons in Figure 2 are made between perturbations of the same type and sign but coming from different contexts. For example, in the upper left panel both time series are the response to a negative PS perturbation—the only difference being that one corresponds to Group 2 (pure context) and the other to Group 3 (combined context). That is, all comparisons involve the same stimuli sequence and thus any observed difference can be unequivocally attributed to different psychological states (context).

Response asymmetry during resynchronization depends on perturbation context

Response asymmetry, a hallmark of nonlinearity, quantifies the degree of mirror-image similarity between time series corresponding to perturbations of opposite signs[13–15]. Figure 3 shows the same data as before but arranged such that now the asymmetry between positive and negative perturbations can be more easily analyzed. In this figure we inverted the sign of the Asynchrony of the positive perturbations, so if the original responses to positive and negative perturbations are symmetric then the two plotted time series will be mostly coincident or overlapping, whereas if they are asymmetric they will show as non-overlapping. An exploratory analysis, again, reveals that some conditions are significantly asymmetric after adjusting for multiple comparisons (asterisks from permutation testing with FDR correction across beeps; see Methods).

In order to test whether there is significant asymmetry, we fitted an LMM to the data shown in Figure 3. Thanks to the sign inversion of the positive conditions, any asymmetry would be represented by a difference between negative and inverted-positive conditions, thus appropriate for linear regression and ANOVA. That is, we look for any effect of Perturbation Sign (either main effect or any of its interactions). The LMM is similar as before with Asynchrony as the dependent variable (beeps $n = 1$ through 6, pooled), factors Context (levels pure/combined), Perturbation Type (levels PS/SC), and Perturbation Sign (levels neg/inv-pos). Participant was included as a random effect; see detailed model specification in Methods.

Model results reflect what can be seen in Figure 3: there are varying degrees of asymmetry, i.e., not all conditions show the same difference between negative and inverted-positive series, and it depends on both Context and Perturbation Type levels (significant three-way interaction: $\chi^2(2) = 14.5$; $p = 0.0007$). An alternative model with all two-way interactions and no three-way interaction shows comparable performance (non significant difference in model comparison: $\chi^2(1) = 1.3$; $p = 0.25$), but see the three-way interaction is preferred because it offers a theoretical interpretation compatible with previous results as discussed in the next section.

Gradual increase of asymmetry

The previous result tells us that some conditions are more asymmetric than others. In this subsection we show that there is a specific ordering that can be interpreted in terms of our novel results regarding the effect of Context and previous literature. In Figure 4a we show the actual asymmetry, that is the difference between the corresponding time series from Figure 3. In the resynchronization phase ($n = 1$ through 6) the conditions have different asymmetry time evolutions, with some conditions closer to zero (less asymmetric) and some others farther away from zero (more asymmetric). The specific order is, from less to more asymmetric: PS pure < PS comb < SC comb <SC pure, and it can be seen more clearly in Figure 4b where we show the average asymmetry across the resynchronization phase.

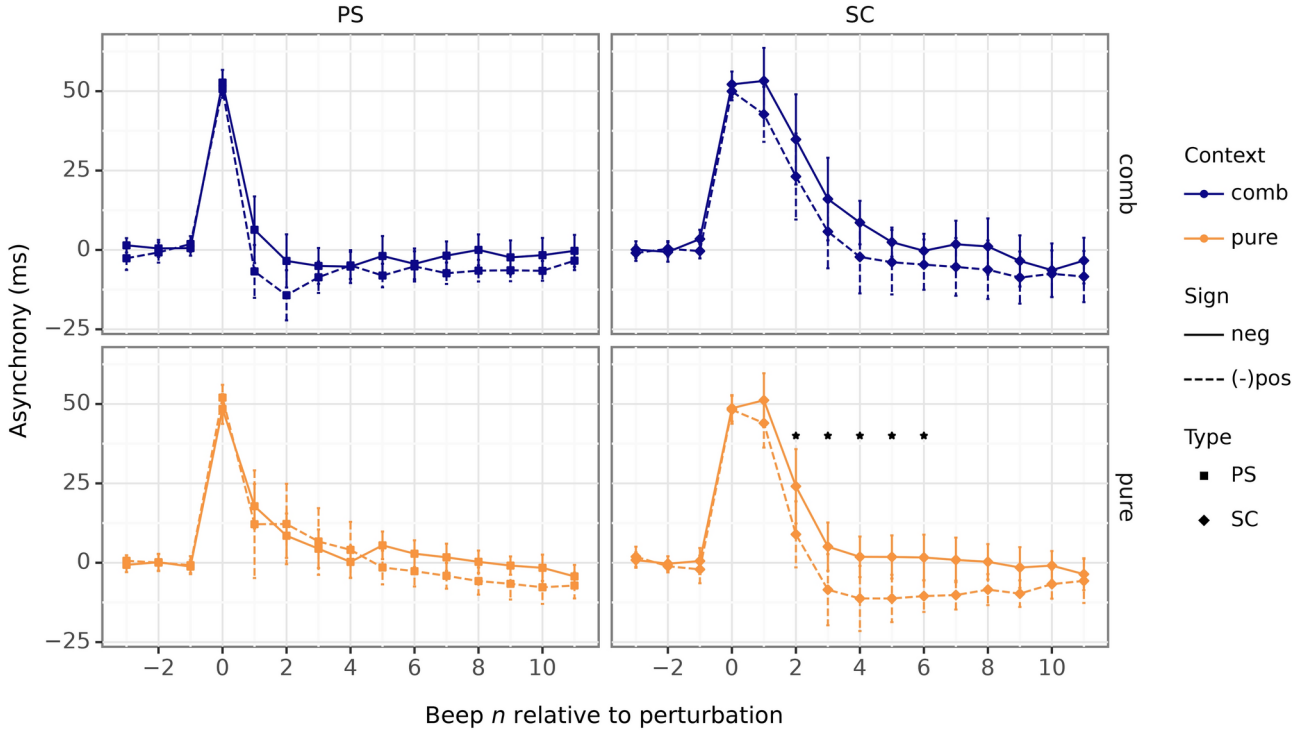

Fig. 3. Asymmetry between responses to perturbations of opposite signs. Same data as above (asynchrony as a function of stimulus number along a trial), here after inverting the sign of the asynchrony from positive perturbations. In this way, any asymmetry will show as a difference between time series within each panel (negative response minus inverted-positive response). Mean across participants $\pm 95\%$ confidence interval. Asterisks mark significant differences after permutation testing with FDR correction for multiple comparisons (see Methods).

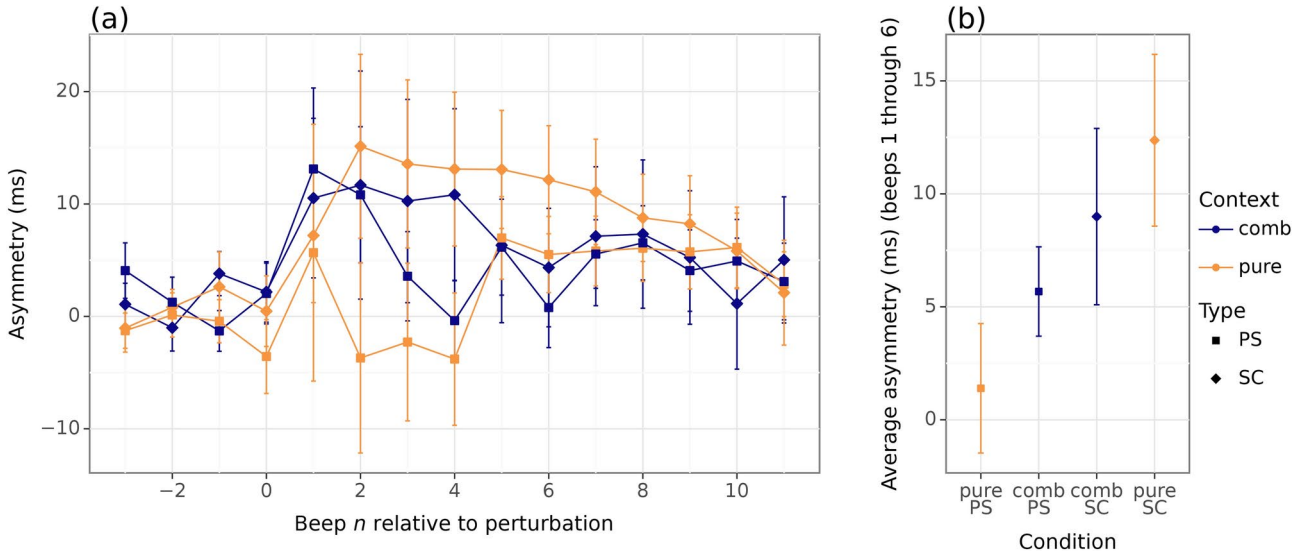

Fig. 4. Asymmetry increases in a specific way depending on both Context and Perturbation Type. **(a)** Asymmetry between responses from perturbations of opposite signs (negative minus inverted-positive). Difference between corresponding time series from previous figure $\pm 95\%$ confidence interval. **(b)** Mean asymmetry during the resynchronization phase (beeps $n = 1$ through 6) from panel a. Mean across beeps $\pm 95\%$ confidence interval.

Although the differences between consecutive conditions in Figure 4b are not significant after correcting for multiple comparisons (see Methods), their ordering has a very clear interpretation based on two facts. First, it is known that pure SC perturbations are more asymmetric than pure PS perturbations[13,15], which in this work is additionally supported by a significant difference between them (asymmetry difference estimate between SC and PS: 10.8 ms, 95% CI: [2.3; 19.3] ms, $t(718) = 2.49$, $p = 0.013$). Second, note that the "combined" conditions have intermediate asymmetry values—the crossover interaction effect of Context we showed above makes the less asymmetric responses more asymmetric and vice versa. We will interpret this in a broader scope in the Discussion.

Context and asymmetry findings hold for different perturbation sizes

In order to determine whether our findings hold in different conditions, we designed our experiment to include a complete set of data with a smaller perturbation size ($\Delta T = \pm 20$ ms) which we analyze here. This perturbation size is particularly interesting because it is just above the detection threshold for step changes (2% of ISI) and right at it for phase shifts (4% of ISI)[14,26,27]. Figure 5 shows a very similar qualitative behavior to that observed for the larger perturbation size. Specifically, conditions SCpos, SCneg, and PSpos in Figure 5a shows the same pattern as in Figure 2 (PSneg, however, shows no clear difference). This is supported by an LMM with a significant three-way interaction as before, meaning that the effect of Context depends on both Perturbation Type and Sign (Supplementary Table S1); and differences between contexts keeping the same signs as before, meaning the direction of the effect is preserved (Supplementary Table S2; note that the difference for PSneg is not significantly different from zero).

Analogously, Figure 5b shows a very similar pattern of asymmetries as in Figure 3: mostly overlapping curves in PSpure, highest neg-pos difference in SCpure, and intermediate differences in PScomb and SCcomb. This is quantitatively confirmed by computing the differences in Figure 6b where the same effect of context is observed: "comb" context makes the PS perturbations more asymmetric and the SC perturbations less asymmetric as in Figure 4b. See Supplementary Table S3 for LMM-ANOVA results (significant three-way interaction) and Supplementary Table S4 for asymmetry differences.

Some other effects in the linear regression for the smaller perturbation size are not significant, however, probably due to lack of statistical power—since the statistical power analysis of our original experiment was performed based on the expected effect size for the larger perturbations, an even larger sample size would be needed for $\Delta T = \pm 20$ as it shows a smaller effect size (see effect sizes for the smaller perturbation size in Supplementary Fig. S1). Despite this, both the context effect and the asymmetry results are replicated in the $\Delta T = \pm 20$ data.

Context is established during the first few trials

As a first approach to determine how rapidly perturbation context is established we analyzed the size of the context effect across the three experimental blocks for the larger perturbation size (presumably larger effect). We considered two different, non-mutually exclusive hypotheses: 1) If exposure to a context were gradually integrated along the experiment, then we would expect no difference between contexts at the beginning of the experiment and an increasing difference as the experiment proceeds; 2) As participants in the "pure" context face

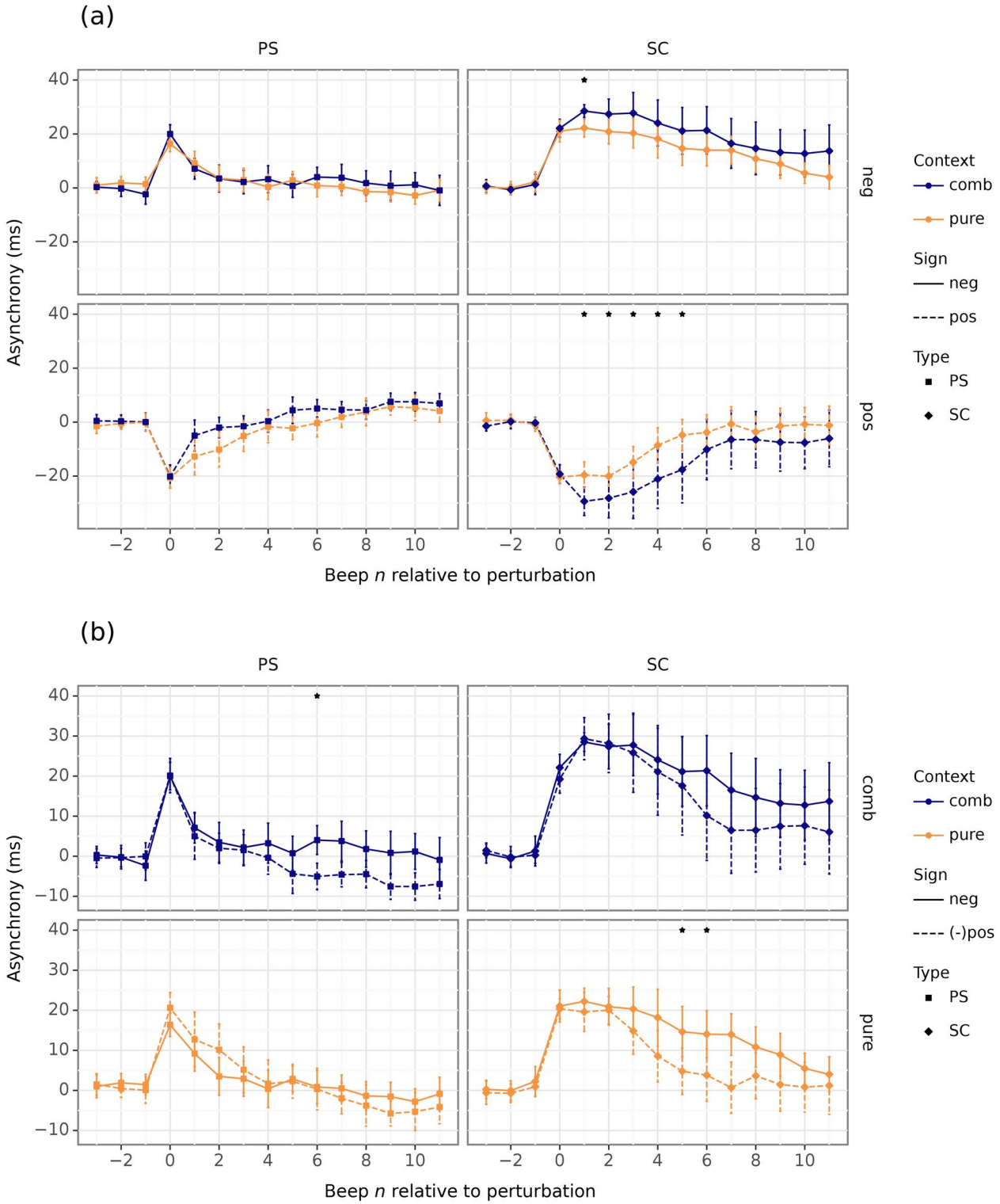

Fig. 5. Generalizability of context and asymmetry findings across different perturbation sizes. **(a)** Effect of Context for the smaller perturbation size ($\Delta T = \pm 20$ ms; compare to Figure 2). Asynchrony as a function of stimulus number along a trial. The data shows the same effect of context as the larger perturbation size. Mean across participants $\pm 95\%$ confidence interval. Asterisks mark significant differences after permutation testing with FDR correction for multiple comparisons (see Methods). **(b)** Asymmetry between responses to perturbations of opposite signs for the smaller perturbation size ($\Delta T = \pm 20$ ms; compare to Figure 3). Same data as in panel (a), here after inverting the sign of the asynchrony from positive perturbations. Mean across participants ±95% confidence interval. Asterisks mark significant differences after permutation testing with FDR correction for multiple comparisons (see Methods).

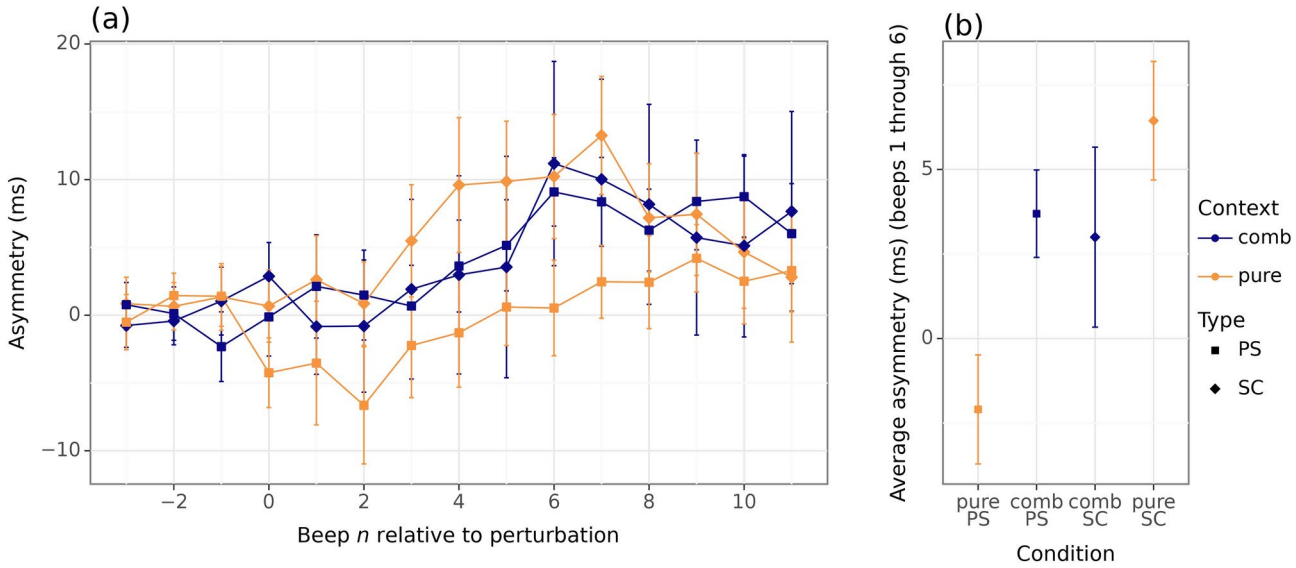

Fig. 6. Asymmetry for the smaller perturbation size. PS perturbations increase their asymmetry when in "comb" context, while SC perturbation decrease it, as for the larger perturbation size. **(a)** Asymmetry between responses from perturbations of opposite signs (negative minus inverted-positive). Difference between corresponding time series from Figure 5b $\pm 95\%$ confidence interval. **(b)** Mean asymmetry during the resynchronization phase (beeps $n = 1$ through 6) from panel a; compare to Figure 4b. Mean across beeps $\pm 95\%$ confidence interval.

less uncertainty (single perturbation type) than participants in the "combined" context (two perturbation types), then we would observe a different rate of adaptation to either context.

Regarding the first hypothesis, Supplementary Fig. S2 shows the difference between contexts as a function of block number for all conditions. Averaged data in right panel shows no clear trend as a function of block number. We conclude that context might be established early on within the first block and thus we plotted data from the first block only in Supplementary Fig. S2. In this case there is a clear difference between the first and second halves of the first block, although there is no easily interpretable pattern—for some conditions the size of the context effect decreases along the first block (PSpos, SCneg) while for other it switches sign (SCpos) or is constant (PSneg).

Regarding the second hypothesis, Supplementary Fig. S3 shows the difference between blocks as a function of context. The averaged data in the right panel shows no clear differences between pure and combined context, so we conclude that any potential difference in adaptation rate between contexts should appear within the first block. In order to confirm this we plotted data from the first block only (difference between second and first halves) in Supplementary Fig. S3. Although there is a clear difference between contexts, as before the result is not easily interpretable—for some conditions context shows no evolution within the first block (PSpos comb, SCneg pure) while for others it increases (PSpos pure) or decreases (SCpos comb).

In summary, it seems to be the case that context takes a few trials to be established and participants in different contexts have potentially different adaptation rates. Notice that invalid trials in our experiment were discarded from analysis and repeated at the end of the block but nonetheless might contribute to the setting of context; this implies that context might be taking slightly more trials to be established. A specifically designed experiment would be needed focusing on the first trials only, where the statistical power would come from a large number of participants to counter the small number of trials analyzed per participant.

Discussion

The error-correction mechanism depends on perturbation context

In this work we showed that exposure to different types of period perturbations in a paced finger-tapping task creates a novel time-related context that modifies the resynchronization response. In other words, the response after a specific perturbation type is different depending on whether the participant was exposed to that perturbation type only during the experiment or to two perturbation types at random. We interpret this in terms of the underlying error-correction mechanism that produces the resynchronization response, suggesting that the mechanism is calibrated by the perturbation context. In terms of a potential mathematical model of the correction mechanism[14,15], calibration by perturbation context may be implemented for instance as an adjustment of the values of correction coefficients depending on whether the participant was exposed to a single or multiple perturbation types during an experiment. Once a context is set, the calibrated mechanism produces responses to all perturbation types and signs[13,15]. The fact that we replicated our results at the smaller perturbation size (near the detection threshold for both SC and PS perturbations) gives us a hint that perturbation context might be set independently of conscious detection of the perturbation, or at least that a few detected perturbations would

suffice for setting context. Additional experiments with clearly subliminal perturbation sizes would be needed with an even larger sample size due to its likely even smaller effect size.

Our results challenge the idea of several different mechanisms, namely the two proposed independent processes of "phase correction" and "period correction"[7](see discussion in Bavassi et al., 2013[14]), where the first one only is active in PS perturbations but both are active in SC perturbations. If the participant doesn't know in advance which perturbation type comes next (combined context), then it is difficult to accommodate the idea of setting the appropriate mechanism beforehand, particularly after considering that PS and SC perturbations are identical up to and including $n = 0$.

The error-correction mechanism is intrinsically nonlinear

Our results regarding the symmetry of pure PS and pure SC perturbations are in agreement with previous literature, and also allow us to interpret the behavior after different perturbation types in a broader scope. On one hand, responses to pure PS perturbations of 10% of the period or less like the ones used in this work are mostly symmetric[6,28] (Figure 4b), leading to mathematical models of behavior that were traditionally linear[29–31]. On the other hand, responses to pure SC perturbations are asymmetric[13,14] (Figure 4b), which lead to nonlinear mathematical models[13–15]. When set in a combined context, however, responses to PS perturbations are more asymmetric than traditionally reported with a degree of asymmetry closer to that of SC perturbations. Thus traditional usage of linear models to represent responses to PS perturbations might not be the best choice.

We interpret this again in terms of the underlying correction mechanism, adding support to the parsimonious hypothesis of a single, nonlinear correction mechanism across perturbation types and signs. Our results add to mounting evidence supporting an intrinsically nonlinear correction mechanism even for perturbation magnitudes as small as 10% of the period[9,13–15]—that nonetheless might be calibrated according to context. The combined context is particularly important for this conclusion because participants don't know in advance which perturbation type comes next since at every trial the stimuli sequences up to $n = 0$ are identical for both perturbation types. We conceptually model the context-dependent calibration of the correction mechanism as an adjustment of the values of the mechanism coefficients[15] such that in pure contexts the coefficients are adjusted to the specific perturbation type, but in a combined context a compromise is reached between the two perturbation types because the participant doesn't know which one comes next.

Methodological implications

We would like to emphasize the importance of measuring and analyzing the full resynchronization phase ($n = 1$ through 6) instead of either just the first response after perturbation ($n = 1$, traditionally reported as the phase correction response or PCR in the tapping literature) or just the post-perturbation baseline (average of $n = 7$ through end of trial). Had we analyzed either $n = 1$ only or $n \geq 7$ only, where the SC perturbations show no effect of context at all (Figures 1b and 2), we would have missed their critical role in both the determination of the context effect and the interpretation of asymmetry.

The existence of a time-related, perturbation-generated context effect in paced finger tapping forces us to make more careful comparisons when the data come from experiments with different perturbation types. The response size and its time evolution during resynchronization will depend on the whole set of perturbation types the participants were exposed to, even when the comparison is made between data of a single perturbation type.

Potential connection between context, neural correlates, and computational models of SMS

As perturbation context is a novel finding, there is no established neural correlate. However, research on neural correlates of time processing in general and of sensorimotor synchronization in particular has seen a surge in the last decade[32] and can help us speculate what a correlate of context might look like. It has been shown that motor timing of individual intervals is related to the speed of a trajectory in the neural space spanned by the activity of every neuron in a population, with longer(shorter) intervals being represented by slower(faster) trajectories in the monkey medial frontal cortex[1,4,33]. Synchronization to a periodic sequence, however, seems to have a different neural representation in that responses to different stimulus periods are represented by neural trajectories of different amplitudes, at least in monkeys[34]. A change in stimulus period might be seen as a switch between trajectories of different amplitudes, and context might thus be related to the speed of the switching—in this way a pure SC would switch more quickly than a combined SC leading to a slightly faster approach to baseline as seen in Fig. 2. It is worth noting, on the other hand, that changes in the speed of the neural trajectory—rather than its amplitude—were observed in a related task that didn't involve synchronization but used periodic stimuli[35] (see in addition[18,36]).

Regarding biologically-plausible computational models of motor timing, similar interpretations and caveats might be valid. Recurrent neural network models propose that the brain encodes time in the varying activity patterns of neuronal populations[37,38]. Advances with such models show how temporal scaling of a motor pattern is achieved by a "speed knob" represented by a constant, low amplitude input to the neurons in the network[39,40]. That is, a global switch can affect the activity of a recurrent neural network in a dynamic way without resorting to a change in connectivity parameter values. Since the effect of context on our time series can be interpreted as a faster or slower convergence to the post-perturbation baseline, we propose that context might be represented by such a global parameter. A word of caution is in order, however, as it must also be taken into account that not only the correction of asynchronies seems to be faster or slower but firstly positive and negative perturbations imply that the whole motor pattern is temporally stretched or compressed.

In terms of a coarser-grain computational model of neural activity, context might also be best represented by different values of a parameter instead of a dedicated module. Take for instance a model of motor planning and sensory anticipation[18]. The model has a nonlinear motor planning module (MPM) for periodic interval production, consisting of three units representing the average activity of three subpopulations of neurons, and a

sensory anticipation module (SAM) that updates the input to the MPM depending on the error signal between its activity and the expected sensory feedback. The model neatly accommodates prior information about the distribution of intervals. Context in this model could be represented, for instance, by constants K (which quantifies the updating of input information) and α (which controls correction of phase difference between sensory and motor signals). It remains to be seen whether the model would perform well when exposed to different context-defining perturbation types in random order.

The computational modeling of context is perhaps more easily envisaged at the behavioural level. Computational models of behaviour can be classified into three main approaches: algorithmic, phase oscillator, and neuro-mechanistic[19]. For the sake of discussion here we the consider the algorithmic class of models where the error-correction mechanism for the asynchrony e_n is represented as an iteration over previous values—i.e., a particular form of the correction function f in the difference equation or map $e_{n+1} = f(e_n; x_n, T_n)$. The correction function $f$is usually polynomial, and linear and nonlinear terms are chosen to best reproduce features of the asynchrony time series[14,15]. These models can provide a unified description of many different conditions: they fit and predict time series from different perturbation types and sizes with a single set of parameter values, a surprisingly rare feat in the experimental psychology literature[15]. The error-correction mechanism in these models is represented by the particular set of coefficient values for the chosen linear and nonlinear terms. We propose that context, being a global characteristic of the task, is assimilated to the set of coefficient values and changing context would mean that a different tuning of the coefficients would be needed. Modelling work will be needed to determine whether a change of context would be represented by a new value of every parameter or only some of them, potentially leading to hints about neural correlates and behavioral interpretation of the involved parameters.

A broader scope: perturbation context, transients, and dynamical systems

The effect of perturbation context is evident in the resynchronization region only; that is, in the transient, as opposed to stationary or equilibrium measures like the pre- or post-perturbation baselines and single data like the PCR which are some of the most common behavioral measures in SMS[6]. This is probably why this effect was hidden in plain sight in the decades-long tapping literature, and it took the specific plot shown in Figure 1b with data from various sources to begin to unveil it.

The transient vs. stationary argument should be seen instead as mutually complementing pieces of evidence within a broader perspective that is rapidly gaining ground: the dynamical systems perspective on flexible timing[41]. According to this framework, the behavioral dynamics results from the interplay among initial conditions, external inputs and internal parameters of the underlying error-correction mechanism. This level of abstraction aims at inferring computational principles with a certain degree of invariance across conditions, that is understanding for example the apparent variety of responses to different perturbation types as particular cases resulting from the rich dynamics of a single mechanism[9,13–15]. Under this light, perturbation context affects the internal parameters of the correction mechanism, usually represented by the coefficient values of a mathematical model[15], where the difference between time series during the transient allows us to access solutions and parameter values that are commonly hidden from analysis.

Methods

Experiment

Ethical considerations

Experimental protocols were designed according to national and international guidelines and were approved by the Ethics Committee of Universidad Nacional de Quilmes. All participants signed a written informed consent.

Participants

Recruited participants were 80 volunteers; six of them could not complete the experiment due to difficulties in synchronizing and three were excluded according to the outlier criteria (see below). The final number of participants was $N = 71$ (ages 18-63 yr, mean 28.6 yr; 33 women; 74 right-handed).

Task and perturbations

The task was auditorily-paced finger tapping with unexpected period perturbations. Participants were instructed to maintain average synchrony as well as possible, using the index finger of the dominant hand. In the event of a perturbation, the participant had to get back to average synchrony without stopping tapping. A single trial consisted of synchronization to a sequence of 35 brief tones with a baseline interstimulus period $T_0 = 500$ ms and a single perturbation. The perturbations occurred at a beep in the sequence randomly chosen in the range 17-22 (beeps were renumbered afterwards such that $n = 0$ corresponds to the perturbation). The perturbations could be either a step-change (SC) perturbation where the sequence period T changes once by an amount ΔT, or a phase-shift (PS) perturbation where the period changes twice at consecutive beeps (first by an amount ΔT and then $-\Delta T$ at the following beep). Each perturbation type could be either positive ($\Delta T = 50$ ms or 20 ms, pos) or negative ($\Delta T = -50$ ms or -20 ms, neg) or neutral ($\Delta T = 0$, isochronous sequence, used as a control, data not included in analysis). A participant exposed to either SC only or PS only perturbations belonged to the "pure" context (Groups 1 and 2, respectively), while a participant exposed to both perturbation types (randomly interspersed) belonged to the "combined" context (Group 3 larger perturbation size, and Group 4 smaller perturbation size). See Table 1 for a summary of experimental conditions.

Experimental setup

The device consisted of an Arduino Mega development board operating as a slave to a master controller program in Python. Board-controller communication was performed via the serial port. At each trial the board received

Group	Context	Type	Sign	Size (ms)	Exp. length (trials)	N particip
1	pure	SC	pos, neg	50, 20	60	17
2	pure	PS	pos, neg	50, 20	60	16
3	combined	SC, PS	pos, neg	50	72	18
4	combined	SC, PS	pos, neg	20	72	20

Table 1. Summary of experimental conditions. Number of participants per group is after outlier removal.

parameters from the controller (interstimulus period, number of stimuli in the trial, whether there should be a perturbation, beep number and size of perturbation, etc) and generated the auditory stimuli, registered tapping, and sent the recorded data back to the controller. The board was complemented with a custom-made shield for interfacing with the user[42] (force sensor to detect taps, audio signals, etc.). Each time the board detected a tap on the force sensor, auditory feedback was sent to the participant as a means to reduce timing variability[14] and to more directly compare to the literature. Both stimulus and feedback tones were 50 ms-long sinusoidal sounds of 440 Hz (A4) and 660 Hz (around E5), respectively. To prevent the participant from having visual feedback, a screen was placed to hide the hand. The sound was played diotically through Sennheiser HD419 headphones. Participants were able to set the sound volume to a comfortable level. A detailed description of the experimental device with open software and hardware is available[42].

Experimental design

We randomly assigned participants to either of the following four groups: Group 1 was exposed to SC perturbations only (18 participants); Group 2 was exposed to PS perturbations only (18 participants); Groups 3 and 4 were exposed to randomly interspersed SC trials and PS trials (18 participants and 20 participants respectively; total 74 participants before outlier removal). Groups 1 and 2 belong to the "pure" context while Groups 3 and 4 belong to the "combined" context. Every perturbation type was presented in two possible sizes and signs: $\Delta T = \pm 50$ ms and $\Delta T = \pm 20$ ms ($\Delta T = 0$ "isochronous" was used as control; data not included in analysis but available in the repository). The final dataset (excluding the isochronous trials) is represented by a fully factorial combination of conditions Context (levels pure/comb) x Perturbation Type (levels PS/SC) x Perturbation Sign (levels pos/neg) x Perturbation Size (levels 50/20 ms). A summary of the experimental conditions is displayed in Table 1.

Every participant performed the experiment in a single session with two stages: Demo and Test. The Demo allowed the participant to become familiar with the task (one trial for every condition). The Test consisted of 12 trials for every experimental condition (including isochronous trials, experiment length was 2x2x12 + 1x12 = 60 trials for Groups 1 and 2, and 2x2x12 + 2x12 = 72 trials for Groups 3 and 4). Trials were equally distributed in three blocks and presented in random order (within-block randomization), with a random intertrial interval between 0.5 and 1 s. Participants moved to the next trial with a key press. Participants were asked to take a short rest between blocks to prevent fatigue.

Each stage ended when all trials were successfully completed. A trial was considered valid if the participant started tapping before the sixth stimulus, did not miss any response thereafter, and did not produce excess responses (more than one per stimulus). Each invalid trial was repeated at the end of the block, until the participant was able to complete all trials successfully.

Outlier criteria

We used a robust definition of outliers (Tukey's fences) and applied uniform criteria at the trial, participant, and condition levels.

- Trial level
  - Trial mean asynchrony. For every participant and condition, we computed the mean asynchrony of every trial, and found the lower (Q1) and upper (Q3) quartiles of the distribution. Any trial with a mean asynchrony outside the interval [Q1-1.5 IQR; Q3+1.5 IQR] (Tukey's fences, where IQR=Q3-Q1 is the interquartile range) was flagged as outlier and removed from the dataset.
  - Trial standard deviation of asynchronies. For every participant and condition, we computed the standard deviation of the asynchronies of every trial, and lower and upper quartiles of the distribution. Any trial with a standard deviation outside Tukey's fences was flagged as outlier and removed from the dataset.
- Participant level
  - Participant in a condition. If more than 50% of trials from the same participant were removed as outliers in a given condition, then the whole participant was removed from the condition.
  - Participant mean asynchrony. We computed the mean asynchrony of every participant in a condition, and the lower and upper quartiles of the distribution. Any participant with a mean outside Tukey's fences was flagged as outlier and removed from the condition.
  - Participant standard deviation of asynchronies. We computed the standard deviation of asynchronies of every participant in a condition. Any participant with a standard deviation outside Tukey's fences was flagged as outlier and removed from the condition.

- Condition level
  - Condition in the experiment. If a participant was removed as an outlier in more than 50% of the conditions, then he/she was removed from the whole experiment. One participant was removed from Group 1, and two were removed from Group 2.The highest number of outlier trials per participant was 11, representing 15% of his/her total trials (at least 1 outlier trial was detected for every participant). The percentage of outlier trials was 8.5% in Group 1, 9.3% in Group 2, 8.1% in Group 3, and 7.9% in Group 4, suggesting that the different experiment durations had a negligible effect on performance.

Data preprocessing
All trials were time-shifted so that they were aligned at the perturbation beep. This beep was renamed $n = 0$ and an analysis range was defined from $n = -6$ through $n = 11$ (range where all participants responded to all stimuli). For every trial we defined three time regions: pre-perturbation (beeps $n = -6$ through -1); resynchronization (beeps $n = 1$ through 6 for SC, beeps $n = 2$ through 6 for PS); and post-perturbation (beeps $n = 7$ through 11); see the regions defined for an exemplary time series in Versaci and Laje, 2021, Figure 1c[43]. For every participant and condition, we defined the pre-perturbation baseline as the average of asynchronies across trials during the pre-perturbation region and then subtracted it from every trial.

Sample size justification

Based on previous literature, we estimated the number of participants needed to observe a significant difference between pure and combined contexts at the first beep after perturbation ($n = 1$) for the larger (50 ms) PS perturbations by power analysis[44]. Code for reproducing the estimation and plotting Figure 1b can be found in the following GitHub repository: https://github.com/SMDynamicsLab/Context2024, file "/analysis/power_analysis.py"

Search of published data
As the specific experimental manipulation of context had not been done before, we searched the paced finger tapping literature for data from published experiments with a single perturbation type ("pure" condition) and with two perturbation types in random order ("combined" condition). The inclusion criteria were the following:

- Paced finger tapping with auditory stimuli;
- PS and/or SC perturbation types;
- Perturbation size of ± 50 ms, baseline period of 500 ms;
- Enough statistical information (sample size, standard error / standard deviation, etc);
- Comparable experimental conditions, e.g. unexpected perturbation.The data, shown in Table 2 and plotted in Figure 1b, was digitized from the published figures using g3data in Ubuntu. Examples of potentially relevant publications that were excluded for specific reasons:
-
- (Repp, 2010)[45] Unusual results from musically trained participants, acknowledged by the author ("...a pattern that has not been observed previously in numerous studies with musically trained participants").
- (López & Laje, 2019)[13] Data available for SC perturbations but ΔT is smaller than 50 ms and not enough data to interpolate.

Condition in Figure 1b	Reference	Original figure	Original condition
Pure SC (pos and neg)	(Thaut et al., 1998)	Fig. 2A	Step change:
		Fig. 2B	$ISI = 500$ ms, $\Delta T = \pm 50$ ms
	(Bavassi et al., 2013)	Fig. 5B	Step change:
			$ISI = 500$ ms, $\Delta T = \pm 50$ ms
Pure PS (pos and neg)	(Repp, 2002a)	Fig. 9	Phase Correction Response (PCR):
			$ISI = 500$ ms, $\Delta T = \pm 50$ ms
	(Repp, 2002b)	Fig. 2A	Phase Correction Response (PCR):
			$ISI = 500$ ms, $\Delta T = \pm 50$ ms
	(Repp, 2010)	Fig. 1A	Phase Correction Response (PCR):
			$ISI = 500$ ms, $\Delta T = \pm 50$ ms
Combined SC (pos and neg)	(Large et al., 2002)	Fig. 2A	Tempo perturbation: Linear interpolation between $ISI = 400$ ms, $\Delta T = \pm 32$ ms and $ISI = 600$ ms, $\Delta T = \pm 48$ ms
Combined PS (pos and neg)	(Large et al., 2002)	Fig. 2A	Tempo perturbation: Linear interpolation between $ISI = 400$ ms, $\Delta T = \pm 32$ ms and $ISI = 600$ ms, $\Delta T = \pm 48$ ms

Table 2. Literature sources for every condition.

- (Repp et al., 2012; Fig. 1)[46] Data must be interpolated; discarded because the Pure PS condition is well represented by three other papers without interpolation.

Power analysis

We want to estimate the minimum number of participants N_{estim} such that the probability of rejecting the null hypothesis given that the alternative hypothesis is true is greater than or equal to the desired power $(1-\beta)$ for an effect size as shown in Figure 1b at $n=1$ (see for instance Suresh & Chandrashekara 2012[47], section "Sample size estimation with two means"):

$$1-\Phi(t_{1-\alpha}-t)\geq 1-\beta$$

that is

$$\Phi(t_{1-\alpha}-t)\geq \beta \tag{1}$$

where Φ is the cumulative distribution function (CDF) of the normal distribution, $t_{1-\alpha}=1.64$ is the critical value of Student's *t* (upper one-tailed, corresponding to $\alpha=0.05$ under the approximation of infinite degrees of freedom), and *t* is the value of Student's *t* corresponding to the observed difference:

$$t=\frac{SES}{\sqrt{2/N_{\text{estim}}}}$$

where *SES* is the observed standardized effect size, that is the standardized difference between conditions from previous literature:

$$SES=\frac{\left(e^{(n=1)}_{\text{pure}}-e^{(n=1)}_{\text{comb}}\right)}{SD_{\text{pooled}}}$$

where $e^{(n=1)}_{\text{pure,comb}}$ is the estimated asynchrony of the response at $n=1$ (averaged across experiments with context *pure*, *comb*) and SD_{pooled} is the pooled standard deviation of the asynchrony estimates:

$$SD_{\text{pooled}}=\sqrt{\frac{(SD^2_{\text{pure}}+SD^2_{\text{comb}})}{2}}$$

where the standard deviation of each condition is the square root of the average variance across experiments, weighted by the corresponding degrees of freedom *df*:

$$SD_{(\text{pure,comb})}=\sqrt{\frac{df_{(\text{pure,comb})\text{exp1}}SD^2_{(\text{pure,comb})\text{exp1}}+df_{(\text{pure,comb})\text{exp2}}SD^2_{(\text{pure,comb})\text{exp2}}+\ldots}{df_{(\text{pure,comb})\text{exp1}}+df_{(\text{pure,comb})\text{exp2}}+\ldots}}$$
$$df_{(\text{pure,comb})\text{expn}}=N_{(\text{pure,comb})\text{expn}}-1$$

Putting all this together, Eq. 1 can be expressed as

$$N_{\text{estim}}\geq 2\left(\frac{t_{1-\alpha}-z_{\beta}}{SES}\right)^2 \tag{2}$$

where $z_{\beta}=-1.28$ (corresponding to $\beta=0.1$ or a power of 0.9 under the approximation of normal distribution). By applying this to the PSneg and PSpos data, we obtained $N_{\text{estim}}\geq 12$. We obtained a similar estimation by using the function solve_power from statsmodel/TTestIndPower[48] in Python.

We obtained a slightly higher estimation by dropping the usual approximations for $t_{1-\alpha}$ and z_{β}, that is by using instead:

$$t_{1-\alpha}=\text{iCDF}(1-\alpha;\ 2N_{\text{estim}}-2)$$
$$z_{\beta}=\text{iCDF}(\beta;\ 2N_{\text{estim}}-2)$$

(where iCDF is Student's inverse CDF or quantile function with $2N_{\text{estim}}-2$ degrees of freedom) and then finding the value of N_{estim} that fulfills the condition in a consistent way. Without the approximations we obtained $N_{\text{estim}}\geq 13$, so we kept this larger estimate instead.

Testing of published data

In order to test for the statistical significance of the differences observed in Figure 1b at $n=1$, we computed a t-test for every comparison between contexts (PSpos, PSneg, SCpos, SCneg, all at $n=1$) and adjusted the

Factor	χ^2	df	p
Context	6.38	1	0.011
Perturb_type	4.01	1	0.045
Perturb_sign	12.66	1	3.7×10^{-4}
Context:Perturb_type	8.47	1	0.0036
Context:Perturb_sign	27.41	1	1.6×10^{-7}
Context:Perturb_type:Perturb_sign	85.28	2	2.2×10^{-16}

Table 3. Anova results, Context testing.

Condition	pure-comb diff estimate (ms)	SE (ms)	95% CI (ms)	df	t	p
PS neg	8.84	3.50	[1.98; 15.68]	125.9	2.53	0.025
PS pos	−14.49	3.87	[-22.08; -6.89]	155.1	−3.74	0.001
SC neg	−4.79	3.63	[-11.92; 2.32]	140.9	−1.32	0.189
SC pos	7.81	3.70	[0.56; 15.06]	143.5	2.11	0.048

Table 4. Estimated marginal means and contrasts for every condition in Figure S1a (FDR-corrected p-values and confidence intervals).

p-values via Bonferroni for four comparisons. The adjusted p-values are 0.00023, 1.7×10^{-6}, 0.059, and 0.59, respectively.

Exploratory analysis: permutation testing

Even an exploratory analysis (less statistical power) prior to fitting any statistical model leads to partially significant results. We tested the differences seen in each panel of Figures 2 and 3 via permutation testing and correcting for multiple comparisons. We first defined the variables DIFF and ASYM. For Figure 2, variable DIFF is the difference between contexts for every combination of Perturbation Type and Perturbation Sign (larger DIFF values represent larger effect of context). For Figure 3, variable ASYM is the difference between signs for every combination of Context and Perturbation Type (larger ASYM values represent larger degree of asymmetry; responses for positive perturbations have their signs inverted; see main text). Then for DIFF and ASYM separately we computed p-values for every beep n in the resynchronization region (see next) and fed them to the FDR Benjamini/Hochberg (positively correlated) algorithm with $\alpha = 0.05$. Beeps with significant differences after correction are indicated by asterisks in Figures 2 and 3.

The p-values to be fed to FDR were estimated via permutation testing as follows. We illustrate the procedure with the PSneg condition in the upper left panel of Figure 2 where the variable DIFF is the difference between the two curves (analogously for ASYM in Figure 3). To generate the distribution of DIFF under the null hypothesis we first pooled the time series from all trials from all participants from the PSneg condition (analogously for the other conditions), then randomly assigned trials to surrogate participants in surrogate contexts "pure" and "combined" according to the number of data in the original dataset, and then performed the same computation that was applied to obtain the actual result, that is average across surrogate trials then across surrogate participants and then take the difference between surrogate contexts, to get a surrogate DIFF time series. This procedure was repeated 5000 times to get a null distribution of 5000 surrogate DIFF time series. We finally compared the true DIFF value at each beep $n = 0$ through 6 to the null distribution at the corresponding beep and computed a p-value as the proportion of null DIFF values above the true DIFF value or below its opposite value (i.e. two-tailed).

Hypothesis testing: effect of context

Code for reproducing all statistical hypothesis testing is available in the following GitHub repository: https://github.com/SMDynamicsLab/Context2024, file "/analysis/hypothesis_testing.Rmd"

Linear regression

Main model: Linear Mixed Model with Asynchrony as dependent variable (pooled beeps $n = 1$ through 6 for SC, $n = 2$ through 6 for PS), fixed-effect factor Context (levels pure/combined), fixed-effect factor Perturbation Type (levels PS/SC), fixed-effect factor Perturbation Sign (levels pos/neg), two-way interactions Context x Perturbation Type and Context x Perturbation Sign, three-way interaction, and random-effect factor Subject. Functions lmer (from library lme4[49]) and Anova (from library car[50]) in R.

Analysis of Deviance Table: Anova table of model parameters; testing the effect of Context in Figure 2; see Table 3 (Type III Wald χ^2 tests)

Post-hoc comparisons: Context differences for every condition in Figure S1a; see Table 4 (FDR-corrected p-values and confidence intervals; functions emmeans, pairs and test from library emmeans[51] in R)

Alternative model: Same as above but no three-way interaction and all two-way interactions. Linear Mixed Model with Asynchrony as dependent variable (pooled beeps $n = 1$ through 6 for SC, $n = 2$ through 6 for PS),

	npar	AIC	BIC	logLik	deviance	χ^2	Df	p-value
Alt. model (context)	9	6600.5	6641.9	-3291.3	6582.5			
Main model (context)	10	6570.4	6616.4	-3275.2	6550.4	32.08	1	1.5×10^{-8}

Table 5. Model comparison, Context testing. Main model is preferred.

Factor	χ^2	df	p
Context	4.07	1	0.044
Perturb_type	3.44	1	0.063
Perturb_sign	0.095	1	0.758
Perturb_sign:Context	0.38	1	0.537
Perturb_sign:Perturb_type	6.25	1	0.012
Perturb_sign:Context:Perturb_type	14.54	2	6.9×10^{-4}

Table 6. Anova results, asymmetry testing.

Asymmetry difference	Estimate (ms)	SE (ms)	95% CI (ms)	df	t	p
PS comb - PS pure	2.57	4.16	[-5.59; 10.73]	700.7	0.62	0.901
SC comb - PS comb	4.13	3.97	[-3.65; 11.91]	680.5	1.04	0.655
SC pure - SC comb	4.11	4.19	[-4.12; 12.33]	714.4	0.98	0.697

Table 7. Estimated marginal means and contrasts for asymmetry differences between consecutive conditions in Figure 4b (Sidak-corrected p-values and confidence intervals).

fixed-effect factor Context (levels pure/combined), fixed-effect factor Perturbation Type (levels PS/SC), fixed-effect factor Perturbation Sign (levels pos/neg), all two-way interactions Context x Perturbation Type, Context x Perturbation Sign, and Perturbation Type x Perturbation Sign, no three-way interaction, and random-effect factor Subject. Model comparison is shown in Table 5 (function anova from stats[52] library in R).

Hypothesis testing: effect of perturbation sign (asymmetry)

Code for reproducing all statistical hypothesis testing is available in the following GitHub repository: https://github.com/SMDynamicsLab/Context2024, file "/analysis/hypothesis_testing.Rmd"

We quantify the degree of response asymmetry in a Context x PerturbationType combination as the mirror-image difference between responses to perturbations of opposite signs (Figure 3). In order to use linear regression and ANOVA we implemented this by inverting the sign of the asynchrony of the positive perturbations, so that asymmetry can be interpreted as the difference between levels of the factor Perturbation Sign (negative minus inverted-positive).

Linear regression

Main model: Linear Mixed Model with Asynchrony as dependent variable (pooled beeps $n = 1$ through 6 for SC, $n = 2$ through 6 for PS) after inverting sign of the asynchronies from positive perturbations as described above. Fixed-effect factor Context (levels pure/combined), fixed-effect factor Perturbation Type (levels PS/SC), fixed-effect factor Perturbation Sign (levels neg/inv-pos), two-way interactions Perturbation Sign x Context and Perturbation Sign x Perturbation Type, three-way interaction, and random-effect factor Subject. Functions lmer (from library lme4) and Anova (from library car) in R.

Analysis of Deviance Table: Anova table of model parameters; testing the effect of Perturbation Sign (asymmetry) in Figure 3; see Table 6 (Type III Wald χ^2 tests).

Post-hoc comparisons: Asymmetry differences between consecutive conditions in Figure 4b; see Table 7 (Sidak-corrected p-values and confidence intervals; functions emmeans, pairs and test from library emmeans in R).

Alternative model: Same as above but no three-way interaction and all two-way interactions. Linear Mixed Model with Asynchrony as dependent variable (pooled beeps $n = 1$ through 6 for SC, $n = 2$ through 6 for PS) after inverting sign of the positive perturbations. Fixed-effect factor Context (levels pure/combined), fixed-effect factor Perturbation Type (levels PS/SC), fixed-effect factor Perturbation Sign (levels neg/inv-pos), all two-way interactions Context x Perturbation Type, Context x Perturbation Sign, and Perturbation Type x Perturbation Sign, no three-way interaction, and random-effect factor Subject. Model comparison is shown in Table 8 (function anova from stats[52] library in R).

	npar	AIC	BIC	logLik	deviance	χ^2	Df	p-value
Alt. model (asymmetry)	9	6495.3	6536.6	-3238.6	6477.3			
Main model (asymmetry)	10	6496.0	6541.9	-3238.0	6476.0	1.309	1	0.253

Table 8. Model comparison, Asymmetry testing. Main model is preferred due to theoretical interpretation.

Color-blind and grayscale friendly plots

We used the plasma colormap.

Data and code availability

All data and code to reproduce our results (experiment, analysis, figures) are available at the Sensorimotor Dynamics Lab's website http://www.ldsm.web.unq.edu.ar/context2024 and Github repository: https://github.com/SMDynamicsLab/Context2024. Reproduction of Figure 1b: /analysis/power_analysis.py. Reproduction of Figures 2- 6 and S1-S3: /analysis/analysis.py. Hypothesis testing: /analysis/hypothesis_testing.Rmd.

Received: 29 April 2024; Accepted: 4 November 2024
Published online: 11 November 2024

References

1. Tsao, A., Yousefzadeh, S. A., Meck, W. H., Moser, M. B. & Moser, E. I. The neural bases for timing of durations. *Nat Rev Neurosci* **23**, 646–665 (2022).
2. Hogendoorn, H. Perception in real-time: predicting the present, reconstructing the past. *Trends Cogn Sci* **26**, 128–141 (2022).
3. Monteiro, T. et al. Using temperature to analyze the neural basis of a time-based decision. *Nat Neurosci* **26**, 1407–1416 (2023).
4. Paton, J. J. & Buonomano, D. V. The Neural Basis of Timing: Distributed Mechanisms for Diverse Functions. *Neuron* **98**, 687–705 (2018).
5. Inagaki, H. K. et al. Neural Algorithms and Circuits for Motor Planning. *Annu Rev Neurosci* **45**, 249–271 (2022).
6. Repp, B. H. Sensorimotor synchronization: a review of the tapping literature. *Psychon Bull Rev* **12**, 969–992 (2005).
7. Repp, B. H. & Su, Y. H. Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon Bull Rev* **20**, 403–452 (2013).
8. Rosso, M., Moens, B., Leman, M. & Moumdjian, L. Neural entrainment underpins sensorimotor synchronization to dynamic rhythmic stimuli. *Neuroimage* **277**, 120226 (2023).
9. Bavassi, L., Kamienkowski, J. E., Sigman, M. & Laje, R. Sensorimotor synchronization: neurophysiological markers of the asynchrony in a finger-tapping task. *Psychol Res* **81**, 143–156 (2017).
10. Merchant, H., Zarco, W., rez, O., Prado, L. & Bartolo, R. Measuring time with different neural chronometers during a synchronization-continuation task. *Proc Natl Acad Sci U S A* **108**, 19784–19789 (2011).
11. de Lafuente, V. et al. Keeping time and rhythm by internal simulation of sensory stimuli and behavioral actions. *Sci Adv* **10**, eadh8185 (2024).
12. Okada, K.-I., Takeya, R. & Tanaka, M. Neural signals regulating motor synchronization in the primate deep cerebellar nuclei. *Nature Communications* **13**, 2504 (2022).
13. López, S. L. & Laje, R. Spatiotemporal perturbations in paced finger tapping suggest a common mechanism for the processing of time errors. *Sci Rep* **9**, 17814 (2019).
14. Bavassi, M. L., Tagliazucchi, E. & Laje, R. Small perturbations in a finger-tapping task reveal inherent nonlinearities of the underlying error correction mechanism. *Hum Mov Sci* **32**, 21–47 (2013).
15. González, C. R., Bavassi, M. L. & Laje, R. Response to perturbations as a built-in feature in a mathematical model for paced finger tapping. *Phys Rev E* **100**, 062412 (2019).
16. Haken, H., Kelso, J. A. & Bunz, H. A theoretical model of phase transitions in human hand movements. *Biol Cybern* **51**, 347–356 (1985).
17. Loehr, J. D., Large, E. W. & Palmer, C. Temporal coordination and adaptation to rate change in music performance. *J Exp Psychol Hum Percept Perform* **37**, 1292–1309 (2011).
18. Egger, S. W., Le, N. M. & Jazayeri, M. A neural circuit model for human sensorimotor timing. *Nat Commun* **11**, 3933 (2020).
19. Large, E. W. et al. Dynamic models for musical rhythm perception and coordination. *Front Comput Neurosci* **17**, 1151895 (2023).
20. Lainscsek, C. et al. Finger tapping movements of Parkinson's disease patients automatically rated using nonlinear delay differential equations. *Chaos* **22**, 013119 (2012).
21. Roman, I. R., Washburn, A., Large, E. W., Chafe, C. & Fujioka, T. Delayed feedback embedded in perception-action coordination cycles results in anticipation behavior during synchronized rhythmic action: A dynamical systems approach. *PLoS Comput Biol* **15**, e1007371 (2019).
22. Rhodes, D. On the Distinction Between Perceived Duration and Event Timing: Towards a Unified Model of Time Perception. *Timing & Time Perception* **6**, 90–123 (2018).
23. Jazayeri, M. & Shadlen, M. N. Temporal context calibrates interval timing. *Nat Neurosci* **13**, 1020–1026 (2010).
24. Narain, D., Remington, E. D., Zeeuw, C. I. D. & Jazayeri, M. A cerebellar mechanism for learning prior distributions of time intervals. *Nature communications* **9**, 469 (2018).
25. Kaya, E. & Henry, M. J. Reliable estimation of internal oscillator properties from a novel, fast-paced tapping paradigm. *Sci Rep* **12**, 20466 (2022).
26. Repp, B. H. Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychological research* **63**, 106–128 (2000).
27. Repp, B. H. Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance* **27**, 600 (2001).
28. Repp, B. H. Phase correction in sensorimotor synchronization: nonlinearities in voluntary and involuntary responses to perturbations. *Hum Mov Sci* **21**, 1–37 (2002).
29. Repp, B. H. & Keller, P. E. Adaptation to tempo changes in sensorimotor synchronization: effects of intention, attention, and awareness. *Q J Exp Psychol A* **57**, 499–521 (2004).
30. Schulze, H., Cordes, A. & Vorberg, D. Keeping Synchrony While Tempo Changes: Accelerando and Ritardando. *Music Perception* **22**, 461–477 (2005).

31. Van der Steen, M. C. & Keller, P. E. The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Front Hum Neurosci* **7**, 253 (2013).
32. Merchant, H. & de Lafuente, V. A second introduction to the neurobiology of interval timing. *Neurobiology of Interval Timing* 3–23 (2024).
33. Sohn, H., Narain, D., Meirhaeghe, N. & Jazayeri, M. Bayesian computation through cortical latent dynamics. *Neuron* **103**, 934–947 (2019).
34. Merchant, H. *et al.* Diverse time encoding strategies within the medial premotor areas of the primate. *Neurobiology of Interval Timing* 117–140 (2024).
35. Egger, S. W., Remington, E. D., Chang, C.-J. & Jazayeri, M. Internal models of sensorimotor integration regulate cortical dynamics. *Nature neuroscience* **22**, 1871–1882 (2019).
36. Egger, S. W. & Jazayeri, M. A nonlinear updating algorithm captures suboptimal inference in the presence of signal-dependent noise. *Scientific Reports* **8**, 12597 (2018).
37. Laje, R. & Buonomano, D. V. Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nature neuroscience* **16**, 925–933 (2013).
38. Buonomano, D. V. & Laje, R. Population clocks: motor timing with neural dynamics. *Trends in cognitive sciences* **14**, 520–527 (2010).
39. Hardy, N. F., Goudar, V., Romero-Sosa, J. L. & Buonomano, D. V. A model of temporal scaling correctly predicts that motor timing improves with speed. *Nature communications* **9**, 4732 (2018).
40. Wang, J., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible timing by temporal scaling of cortical responses. *Nature neuroscience* **21**, 102–110 (2018).
41. Remington, E. D., Egger, S. W., Narain, D., Wang, J. & Jazayeri, M. A Dynamical Systems Perspective on Flexible Motor Timing. *Trends Cogn Sci* **22**, 938–952 (2018).
42. Caral, P. *et al.* Tappingduino: A versatile shield for paced finger tapping experiments with arduino. *OSF*[SPACE]https://doi.org/10.31234/osf.io/wa9j3 (2023).
43. Versaci, L. & Laje, R. Time-oriented attention improves accuracy in a paced finger-tapping task. *Eur J Neurosci* (2021).
44. Lakens, D. Sample size justification. Collabra. *Psychology* **8**, 33267 (2022).
45. Repp, B. H. Sensorimotor synchronization and perception of timing: effects of music training and task experience. *Hum Mov Sci* **29**, 200–213 (2010).
46. Repp, B. H., Keller, P. E. & Jacoby, N. Quantifying phase correction in sensorimotor synchronization: empirical comparison of three paradigms. *Acta Psychol (Amst)* **139**, 281–290 (2012).
47. Suresh, K. & Chandrashekara, S. Sample size estimation and power analysis for clinical research studies. *J Hum Reprod Sci* **5**, 7–13 (2012).
48. Seabold, S. & Perktold, J. Statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference* (2010).
49. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48. https://doi.org/10.18637/jss.v067.i01 (2015).
50. Fox, J. & Weisberg, S. *An R Companion to Applied Regression* (Sage, Thousand Oaks CA, 2019), third edn.
51. Lenth, R. V. *emmeans: Estimated Marginal Means, aka Least-Squares Means* (2023). R package version 1.8.7.
52. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2024).

Acknowledgements

This work was supported by Universidad Nacional de Quilmes (Argentina, grant PUNQ #53/1028) and CONICET (Argentina). We thank Laura C. Estrada and Diego Fernández Slezak for help during data collection.

Author contributions

A.D.S. and R.L. designed the experiments. R.L. conceived of the approach. A.D.S. wrote the code and performed the experiments. A.D.S. and R.L. analyzed the data. R.L. wrote the manuscript draft. A.D.S. and R.L. edited and reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-78786-5.

Correspondence and requests for materials should be addressed to R.L.