RESEARCH REPORT

EJN European Journal of Neuroscience FENS WILEY

Time-oriented attention improves accuracy in a paced finger-tapping task

Leonardo Versaci[1] (ID) | Rodrigo Laje[2] (ID)

[1]Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Bernal, Argentina

[2]CONICET, Buenos Aires, Argentina

Correspondence
Rodrigo Laje, Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, R. S. Peña 352, Bernal, Buenos Aires, B1876BXD, Argentina.
Emil: rlaje@unq.edu.ar

Abstract

Finger tapping is a task widely used in a variety of experimental paradigms, in particular to understand sensorimotor synchronization and time processing in the range of hundreds of milliseconds (millisecond timing). Normally, subjects do not receive any instruction about what to attend to and the results are seldom interpreted taking into account the possible effects of attention. In this work we show that attention can be oriented to the purely temporal aspects of a paced finger-tapping task and that it affects performance. Specifically, time-oriented attention improves the accuracy in paced finger tapping and it also increases the resynchronization efficiency after a period perturbation. We use two markers of the attention level: auditory ERPs and subjective report of the mental workload. In addition, we propose a novel algorithm to separate the auditory, stimulus-related components from the somatosensory, response-related ones, which are naturally overlapped in the recorded EEG.

KEYWORDS

attention, event-related potentials, finger tapping, sensorimotor synchronization, time processing

1 | INTRODUCTION

In this work we address the question of whether sensorimotor synchronization (SMS) is affected by the level of attention during a paced finger-tapping task. The potential effects of attention on SMS began to be considered only recently; a significant effect of attention on the subject's performance would potentially mean that future experiments should consider attention as a variable to be controlled for, or at least take into account its contribution to the overall variability. In the same sense, it might lead us to reconsider or reinterpret many published results in the literature of paced finger tapping given the ample variety of experimental conditions where attention would be expected to vary. In addition, establishing the role of attention in paced finger tapping and thus deciding whether the task is strongly automatic (Kahneman & Chajczyk, 1983) might help us identify which brain regions

are involved, a still unanswered question due to the multitude of hypothesized processes underlying this behaviour (López & Laje, 2019).

1.1 | Paced finger tapping and the error correction mechanism

In a paced finger-tapping task the subject is instructed to tap in synchrony with a periodic external stimulus, as in keeping pace with music. This very simple task allows us to study the processing of time in the brain (Bavassi et al., 2017), multisensory integration (Mates & Aschersleben, 2000), interpersonal coordination (Konvalinka et al., 2010), timing mechanisms in non-human primates (Merchant et al., 2013) and interspecific comparisons (Zarco et al., 2009), differences between

musicians and non-musicians (Franěk et al., 1991), metrical perception (Iversen et al., 2015), motor control (Dione & Delevoye-Turrell, 2015) and higher cognitive functions like attention (Miyake et al., 2004). It is important to note that paced finger tapping cannot be sustained without an error correction mechanism—otherwise the intrinsic variability of motor responses and the smallest detuning between actual and perceived interstimulus interval would accumulate leading to diverging differences between stimuli and taps (Repp, 2005).

One of the most commonly used observables to study the error correction mechanism is the asynchrony, that is, the difference between the occurrence time of the n-th response R_n (tap) and the occurrence time of the associated stimulus S_n (beep): $e_n = R_n - S_n$. In a paced finger-tapping task the mean asynchrony (MA) is usually negative, meaning that on average the response precedes the stimulus. Normally, non-musician subjects have an MA of a few tens of milliseconds depending on the experimental condition (Mates, 1994). For instance, Aschersleben and Prinz (1995) observed that additional auditory feedback from the taps (i.e. a tone of a different pitch added at time of finger contact) produces a decrease in absolute value of the MA (closer to zero). It is also observed that MA depends on the interstimulus interval (ISI), with MA values getting more negative (i.e. larger in absolute value) as the ISI is increased (Kolers & Brewster, 1985; Repp, 2003). The fact that in general taps precede beeps is one of the most successfully replicated findings in SMS, and it is called 'negative MA' (NMA; Repp, 2005). Surprisingly, there is no clear consensus yet about its mechanisms and the absence of an informed theoretical framework makes its interpretation difficult. On the other hand, the MA has shown to be sensitive to experimental manipulation, which allows us to investigate the factors the error correction mechanism depends on (Repp, 2005).

1.2 | The error correction mechanism and the role of perturbations

The error correction mechanism can be probed by performing perturbations to the stimuli sequence. One of the most frequently used temporal perturbations is an abrupt change in the interstimulus interval, known as a step change (an abrupt change of tempo in musical terms). When unexpected, the temporal perturbation induces a forced error in the asynchrony at the perturbation beep, after which the subject has to recover average synchronization, usually achieved in a few taps. There can be differences between the pre- and post-perturbation steady states. (Bavassi et al., 2013) reported a change between pre- and post-perturbation MA opposite to the one expected by a non-perturbative change in ISI of the same size (i.e. between constant ISI conditions).

Praamstra et al. (2003) reported a similar effect with phase-shift perturbations.

Between the pre- and post-perturbation steady states there is a resynchronization phase. Resynchronization is the transient that begins with the forced error in the perturbation beep and ends when the subject reaches the post-perturbation baseline. Several measures have been proposed to quantify the resynchronization phase, for instance, the phase correction response (PCR) defined as the difference between the occurrence time of the first response after perturbation and its expected occurrence time if no perturbation was in place (Repp, 2011; Repp & Keller, 2004). In this work we introduce a new measure that takes into account the pre-post change in MA called *resynchronization efficiency*.

1.3 | Temporal perturbations and temporal attention

The effect of attention on a finger-tapping task has been studied under a few conditions. Repp and Keller (2004) analysed the effect of attention on the resynchronization phase after step-change perturbations by using a dual-task paradigm, with arithmetic operations as secondary tasks. He observed that subjects corrected more quickly (greater PCR) in the single- versus dual-task condition. In other words, they concluded that diverting attention from the main (finger tapping) task makes the resynchronization phase slower (smaller PCR) after a step change.

The effect of attention on the steady-state phase in a finger-tapping task was analysed by Miyake et al. (2004) in a dual-task paradigm with word recall as secondary task. He did not find any difference in asynchrony variability between single- and dual-task conditions across a range of ISI from 450 to 1,500 ms. Regarding the steady state phase, Miyake concludes that finger tapping synchronization is an automatic behaviour not mediated by attention. A different work by Caspi (2002) attempted to modify the MA in the steady state by instructing the subject to focus the attention on the tactile sensation from the tap. In this work, Caspi made the tapping surface shift downwards at some point to prevent the finger from making contact and instructed the subjects to stop tapping immediately after perceiving it. His hypothesis was that this instruction would orient attention towards the tactile sensation, probably making it have better temporal resolution and thus decreasing the time difference between tap and beep. Difficulties with the experimental design hampered reaching a clear conclusion.

It is worth noting the purely temporal nature of the paced finger-tapping task. It is reasonable thus to expect effects on the performance when the attention is focused specifically on temporal aspects of the task. However, to the best of our knowledge there are no published works on the effects of

time-oriented attention on paced finger tapping. In this work we developed a novel experimental paradigm according to this idea.

1.4 | Attention and electrophysiology

Our experimental paradigm allows us to manipulate the level of attention, and we record auditory ERPs in order to have an electrophysiological support for it. It is known that attention modulates early components of stimulus-induced ERPs (Alho et al., 1994a; Eimer & Forster, 2003a; Hansen & Hillyard, 1980; Teder et al., 1993), and particularly P2. Although its functional significance is not thoroughly understood, there is mounting evidence that the auditory- and somatosensory-related P2 is related to attention and, specifically, that its amplitude is reduced (i.e. less positive) under high attention conditions (Crowley & Colrain, 2004; Eimer & Forster, 2003a; Mueller et al., 2008; Näätänen, 1990). In a paced finger-tapping task, however, it is difficult to isolate stimulus-induced ERPs because of contamination with sensorimotor activity from the taps. Due to the very nature of the task, one could expect temporal overlapping of both tap-related components (finger movement, and proprioceptive and tactile perception) and stimulus-related components (auditory stimuli). To address this difficulty we turn to a work by Woldorff (1993) who developed the Adjar algorithm to isolate adjacent ERPs. The algorithm was originally designed for an experimental paradigm with dichotic auditory stimuli of very short ISIs, but it is powerful enough to be generalized. In this work we developed a novel algorithm to isolate the auditory stimulus-induced ERPs in a paced finger-tapping task by adapting Woldorff's algorithm (see details in the Appendix A).

1.5 | This work

We set out to analyse the synchronization performance in a paced finger-tapping task with perturbations to study the potential effects of time-oriented attention on both the steady-state and resynchronization phases. To this we set two levels of attention: NORMAL and HIGH. Unlike other works (Caspi, 2002; Miyake et al., 2004; Repp & Keller, 2004), our HIGH condition makes the subject direct the attention to the temporal aspects of the stimuli sequence. We test the hypothesis of whether attention improves performance in two potential ways: that a higher level of attention leads to better accuracy and precision in the steady state phase, and to a higher efficiency in the resynchronization phase. In addition, two more factors are included in the experimental design: auditory feedback and perturbation sign. Finally, to have a quantitative support for our claim of an actually higher level of attention in the HIGH condition, we report two measures commonly used as evidence of a change in attention level: auditory ERPs and mental workload.

2 | METHODS

2.1 | Subjects and ethical considerations

Forty-four right-handed musicians (11 women) participated as paid volunteers ($5, roughly 150 Argentine pesos) after signing an informed consent. We chose right-handed subjects to diminish potential confounding factors due to lateralization that would wash out the EEG recordings. We decided to study musicians only (with at least 3 years of practice) because they usually show smaller variability (Repp, 2010). Mean age was 27.5 years (range 19–44) and mean practice was 5.3 years (range 3–10). It is known from the literature that the effect of gender on measures of synchronization is not significant (Crozier et al., 2019; Krause et al., 2010). Also the effect of age is not significant for ages above around 15 years (Drewing et al., 2006). Similarly, musical training has little effect for training durations above a few years (Law & Zentner, 2012). Despite this, to make sure we get similar distributions across conditions, we chose to randomly assign subjects as described below. We recorded behavioural data (response occurrence times) from all 44 subjects; in addition, we recorded EEG and mental workload from 22 of them. Our experimental protocols were designed in accordance with national and international guidelines and were approved by the Ethics Committee of the University of Quilmes.

2.2 | Experimental design

The task is paced finger tapping with unexpected step-change period perturbations (baseline ISI = 600 ms, perturbations +60 ms and −60 ms). We designed a full factorial experiment, with factors Attention (levels NORMAL and HIGH, between-subjects), Auditory Feedback (levels No FBK and With FBK, within-subjects), and Perturbation Sign (levels POS and NEG, within-subjects). Regarding the Attention factor, subjects were randomly assigned to either one of the levels. The rationale for this design choice was to avoid carryover effects and was based on a pilot experiment where we found that mixing attention levels made all data be contaminated by the HIGH level.

2.3 | Procedure

Subjects were instructed to tap with the right index finger on a force resistive sensor (FSR) while keeping pace with an

auditory metronome. In order to avoid potential differences arising from the effector, all subjects were instructed to use the wrist joint and keep the forearm at rest on the desk. They were told not to move any other body part during a trial and look at a fixation cross in the centre of the screen. Before starting the experiment, subjects adapted to the task during a 10-min-long practice phase. We created the two levels of attention (NORMAL and HIGH) by using two separate force sensors located side-to-side, in the following way. In the NORMAL condition subjects tapped on the left sensor (see Figure 1a) without any specific instruction about what to pay attention to. In the HIGH condition subjects started tapping on the left sensor, and they had to switch to the right sensor as soon as they perceived the step-change perturbation and keep tapping on the right sensor until the end of the trial. Auditory feedback from the taps was added (With FBK) or not (No FBK) as a short tone of a different pitch.

Every subject was randomly assigned to either NORMAL or HIGH attention. The experiment consisted of two blocks (one block for each feedback condition, in random order and counterbalanced across subjects) with a 3-min rest between them. Within each block, the subject had to perform 12 valid trials (6 trials from each perturbation sign, random order) with a 1-min rest after the first 6 valid trials. Each trial was a sequence of 32 stimuli; the pre-perturbation value of the ISI was 600 ms; the perturbation was a tempo step change of ±60 ms, occurring between beeps 11 and 15 at random. Stimulus tones were 50-ms long, 500-Hz sinusoidal waveform sounds; feedback tones were 50-ms long, 1800-Hz sinusoidal waveform sounds. The experiment's total duration was approximately 45–60 min including the practice phase.

We consider a trial valid if: (1) none of the asynchronies are larger (in absolute value) than 250 ms; (2) there are equal numbers of stimuli and responses since the occurrence of the first recorded response; (3) the first recorded response occurs at most at the third beep; (4) for HIGH attention only: the subject must switch from left sensor to right sensor right after the perturbation (exactly one beep after). If a trial is considered invalid, the subject receives feedback about the type of error and an additional trial is added at the end of the block. The EEG subjects were asked to quantify the mental workload in completing the task at the end of the experiment. According to the Cognitive Load Theory (Sweller, 2011), the level of automatism in performing a task and associated
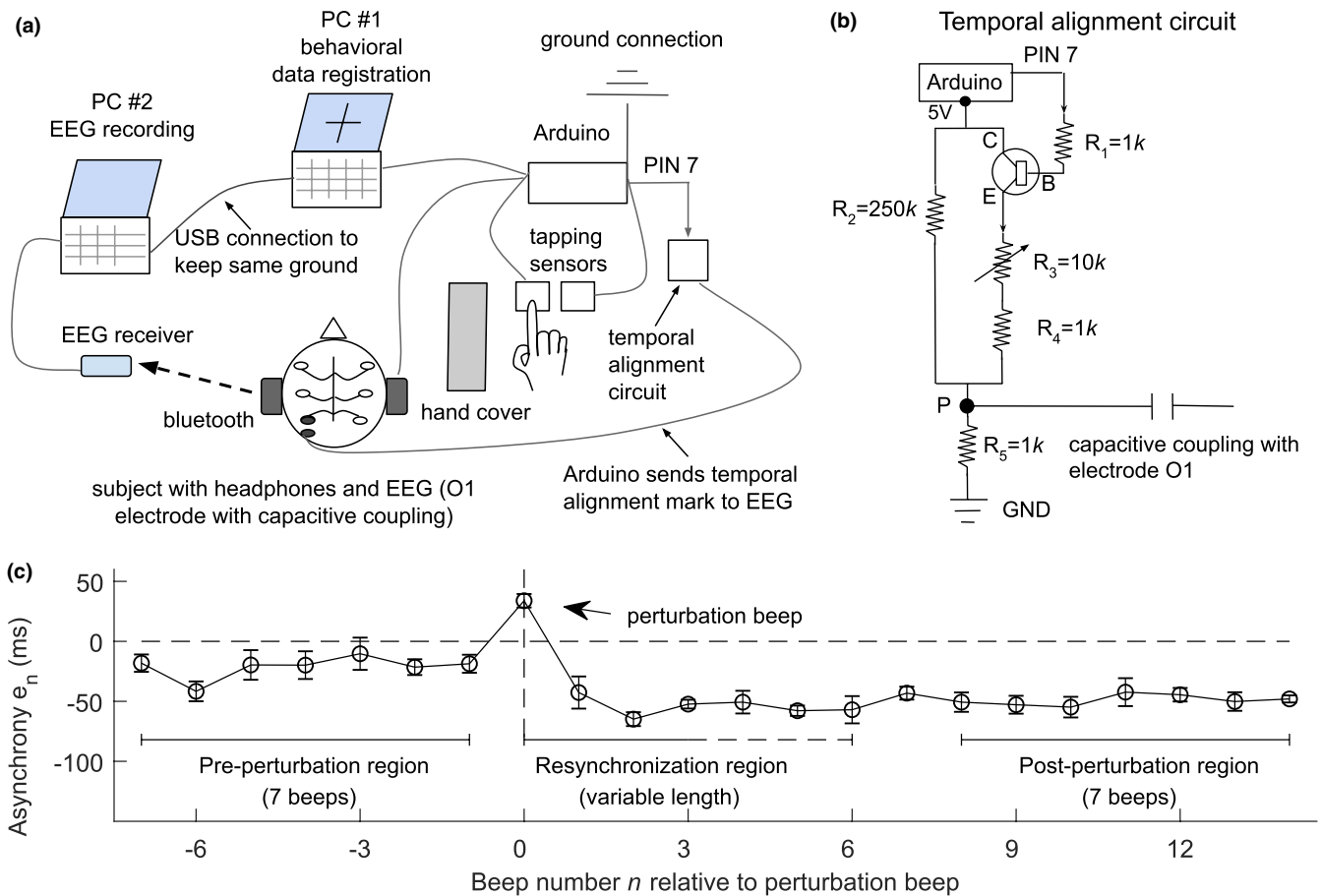

FIGURE 1 (a) Experimental setup. (b) Circuit for the temporal alignment between Arduino and EEG. (c) Average trial from one subject in one condition (mean asynchrony ± standard error across trials). $ISI_{pre} = 600$ ms, $ISI_{post} = 540$ ms (negative perturbation)

mental workload is inversely related. We used a 9-point Likert scale, from 'very, very low mental effort' (1) through 'very, very high mental effort' (9) (Paas et al., 2003). Each subject reported the mental workload for each feedback condition; the mental workload corresponding to the attention level which the subject was assigned to was the average between the two reports.

2.4 | Behavioural data recording

Stimuli and feedback presentation and response recording were performed with an Arduino Mega 2560 and two force-sensitive resistors (FSR) separated 0.5 cm from each other, connected to PC #1 (the EEG data were recorded by PC #2, see below; see Figure 1a). We used custom-written C code and adapted code from a program originally designed for Arduino Uno (Schultz & van Vugt, 2016) in order to present sounds with sinusoidal waveform. According to Schultz, the combination Arduino-FSR has the shortest latency between the time of contact and the time of occurrence of the feedback (0.6 ms) and also the smallest latency variability (0.3 ms), among other features. Non-EEG subjects used Sennheiser HD 280 Pro headphones; EEG subjects used Sennheiser CX 300-II ear-canal phones. To avoid visual feedback from their own finger movement, the subjects were asked to fix gaze at a cross in the screen and a vertical cardboard screen was put to block peripheral vision. Subjects set sound volume to a comfortable level. The code for controlling the experiment and the code for data analysis were written in MATLAB$^{®}$.

2.5 | Behavioural data processing

Every subject performed six valid trials for each condition (feedback and perturbation sign). After the experiment we took all valid trials from all subjects and aligned them at the perturbation beep (labelled $n = 0$), and selected the range $n = -7$ through $n = 14$ where all trials from all subjects had responses. We discarded any response outside that range, leaving all trials with the same number of responses (22) and the perturbation at the same position ($n = 0$). For each subject and condition we averaged across the six trials to obtain an average trial in which we defined three regions (Figure 1c):

- Pre-perturbation region (the seven responses before perturbation: $n = -7$ through $n = -1$);
- Post-perturbation region (the last seven responses: $n = 8$ through $n = 14$);
- Resynchronization region (variable length: $n = 0$ through the resynchronization beep which is the first of two consecutive beeps whose asynchronies are statistically indistinguishable from the post-perturbation baseline that is

the mean asynchrony of all trials in the post-perturbation region).

As described above, we recorded two groups of subjects:

- Year 2018 (22 subjects; behaviour only; headphones);
- Year 2019 (22 subjects; behaviour + EEG + mental workload; ear-canal phones).

The *pre-perturbation mean asynchrony* (MA_{pre}) was computed as follows: for each subject and condition, we first averaged the asynchronies in the pre-perturbation region (along the trial), then averaged across trials. Analogously, we defined the *post-perturbation mean asynchrony* (MA_{post}) in the post-perturbation region. We defined the pre-post change in MA as $\Delta MA = MA_{post} - MA_{pre}$.

The *pre-perturbation intratrial SD* (SD_{pre}) was computed as follows: for each subject and condition, we first computed the standard deviation of asynchronies in the pre-perturbation region (along the trial), then averaged across trials. The *post-perturbation intratrial SD* (SD_{post}) follows an analogous definition in the post-perturbation region. We defined the pre-post change in *SD* as $\Delta SD = SD_{post} - SD_{pre}$.

2.5.1 | Data pooling

The experimental design is the same for both subject groups (full factorial experiment, three-factor design: attention, feedback and perturbation sign). To determine whether the groups were comparable we pooled all subjects and performed two ANOVAs, one for MA_{pre} and one for SD_{pre} with Year as an additional factor (which represents all differences: year, measurement and headphone/earphone; we dropped the factor perturbation sign as the two tested values were pre-perturbation only). We did not find any significant main or interaction effects of year, either for MA_{pre} (three-way ANOVA, with year and attention as between-subject factors, and feedback as within-subject factor; significant effect of attention $F(1, 40) = 11.99$, $p = 0.0013$; significant effect of feedback $F(1, 40) = 5.91$, $p = 0.019$; non-significant effect of year $F(1, 40) = 0.05$, $p = 0.81$; non-significant att x fbk interaction $F(1, 40) = 0.51$, $p = 0.47$; non-significant att × year interaction $F(1, 40) = 0.27$, $p = 0.60$; non-significant fbk × year interaction $F(1, 40) = 0.81$, $p = 0.37$; non-significant att × fbk × year interaction $F(1, 40) = 0.51$, $p = 0.48$) or SD_{pre} (three-way ANOVA, with att and year as between-subject factors, and feedback as within-subject factor; non-significant effect of attention $F(1, 40) = 3.53$, $p = 0.06$; non-significant effect of feedback $F(1, 40) = 1.39$, $p = 0.24$; non-significant effect of year $F(1, 40) = 1.72$,

$p = 0.19$; non-significant att×fbk interaction $F(1, 40) = 0.02$, $p = 0.87$; non-significant att×year interaction $F(1, 40) = 0.01$, $p = 0.91$; non-significant fbk×year interaction $F(1, 40) = 2.82$, $p = 0.10$; non-significant att × fbk × year interaction $F(1, 40) = 0.40$, $p = 0.52$). Based on these results, we decided to pool all subjects into a single group of $N = 44$.

2.5.2 | Heteroscedasticity in MA_{pre}

A Levene test applied on MA in the pre-perturbation region revealed a significant heteroscedasticity: inter-subject variability in NORMAL was larger than in HIGH ($p = 0.019$; feedback conditions were averaged). This fact can be observed in Figure 2 where temporal series in HIGH are distributed in a narrower band of asynchronies than in NORMAL. To remove heteroscedasticity a $\log_{10}(-MA + \text{constant})$ transformation was applied. All subsequent analyses and statistical tests were performed on the transformed data.

2.5.3 | Expected ΔMA and ΔSD values from previous behavioural works

The hatched bars in Figure 4 represent the expected values based on a work by Repp, 2003 (Figure 2a,b, 'Slow auditory' condition) in which the ISI was fixed during each trial, and changed across trials (non-perturbative period

manipulation). We digitized and replotted the data and performed a robust linear regression; see Figure B1. The obtained slope for MA was -0.0402 ± 0.0075, confidence interval $[-0.0545; -0.0259]$, $p = 3 \times 10^{-8}$; therefore the expected change in MA due to a non-perturbative ISI change of ± 60 ms is ± 2.41 ms (hatched bars in Figure 4a). The obtained slope for SD was 0.0255 ± 0.0015, confidence interval $[0.0256; 0.0284]$, $p = 8 \times 10^{-65}$; thus the expected change in SD due to a non-perturbative ISI change of ± 60 ms is ± 1.53 ms (hatched bars in Figure 4b).

2.6 | Resynchronization efficiency

We define resynchronization efficiency γ as:

$$\gamma = \frac{\left|E_0 - E_{post}\right|}{E_{sum}} \tag{1}$$

where E_0 is the asynchrony at the perturbation beep, averaged across trials (i.e. the mean forced error at $n = 0$); E_{post} is the post-perturbation mean asynchrony, averaged across trials; and E_{sum} is the sum of all asynchrony values (relative to E_{post}) from the perturbation beep ($n = 0$) through the resynchronization beep (n_R):

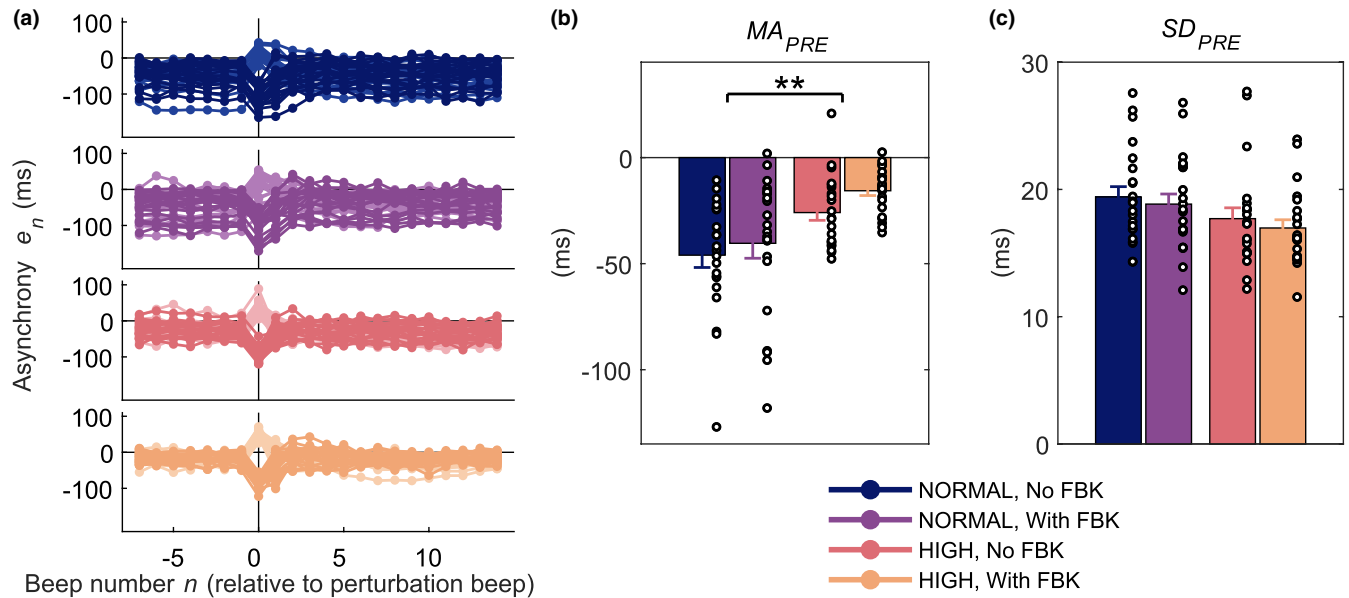$$E_{sum} = \sum_{n=0}^{n=n_R} \left|E_n - E_{post}\right| \tag{2}$$

FIGURE 2 (a) Average time series for each subject and each condition (mean across trials; in each subpanel, POS perturbations are darker, NEG perturbations are lighter). (b) Attention significantly improves accuracy. Auditory feedback shows a similar behaviour but its effect is not significant. Mean asynchrony in the pre-perturbation region MA_{pre} (bars: mean $\pm$ standard error across subjects; each dot represents a subject). Attention factor: $p = 0.0035$. (c) Attention and feedback seem to improve precision too but their effects are not significant. Intra-trial standard deviation in the pre-perturbation region SD_{pre} (bars: mean $\pm$ standard error across subjects; each dot represents a subject)

where E_n is the mean asynchrony of the beep n across trials in the resynchronization region. See Figure B2 for a graphical depiction of E_0, E_{post}, E_{sum} and n_R.

Conceptually, the resynchronization efficiency is the ratio between the amount of asynchrony the subject must correct during resynchronization and the sum of actual asynchronies until resynchronized. The resynchronization efficiency takes values ranging from 0 to 1, and depends on three factors: size of forced error at the perturbation bip, duration of the resynchronization phase and shape of resynchronization phase. At a fixed value of the forced error $E_0 - E_{post}$, the shorter the resynchronization phase the greater the efficiency; while at a fixed duration of the resynchronization phase, the larger the forced error $E_0 - E_{post}$ the greater the efficiency. Efficiency also depends on the shape of the asynchrony time series: at fixed forced error and duration of the resynchronization phase, a subject with a shallower asynchrony time series will have a greater efficiency than a subject with constantly high asynchrony values until resynchronized.

2.7 | EEG recording

EEG recording was performed with an Emotiv EPOC+ system with a sampling rate of 128 Hz. The system has 14 channels placed in locations consistent with the 10–20 montage, and two other additional electrodes located in the mastoids. The Emotiv EPOC+ was specifically validated for auditory ERP recording (Badcock et al., 2013, 2015) and also for a wide range of measurements (Pietto et al., 2018). In order to refer the EEG data and the behavioural data to a common time axis, we adapted the Emotiv EPOC+ to receive temporal alignment marks. These marks were sent to one electrode exclusively used for that purpose, at least 500 ms before the beginning of each trial. The O1 electrode was chosen to accomplish this function because it is located outside the area of study. The 13 remaining electrodes were used to collect the EEG data.

The temporal alignment marks were delivered with a simple, custom-made circuit (Figure 1b). Each mark consists of a sudden variation of the voltage at node P (placed near O1) and is registered by the Emotiv as a well-defined oscillation in the O1 channel by means of capacitive coupling. The amplitude of the mark registered in the O1 channel is proportional to the variation of voltage in node P. The circuit consists of a voltage divider in which one of the resistors varies its value according to the state of the kn2222 transistor. When the transistor is in cutoff mode the whole current flows through R2, and the voltage at point P reaches its minimum. When the transistor is in saturation mode most of the current passes through R3 and R4 (as the sum of both resistors is much lower than R2) and the voltage in node P reaches its maximum. The resistor values were selected in such a way that the voltage in node P has minimum and maximum values of 20 mV and 1.5 V respectively.

R3 is a variable resistor used to control the maximum voltage in point P (mark amplitude). The Arduino was programmed to alternate the transistor mode between cutoff (PIN 7 = LOW) and saturation (PIN 7 = HIGH) every 100 ms during 400 ms. This specific mark is detected by software at a later stage (i.e. offline, after the experiment).

When the finger makes contact with the sensor an electrostatic charge is generated which is observed as a tiny voltage variation in the O1 electrode. To eliminate this source of noise it is necessary to equip the circuit with an adequate ground connection and also have the subjects rest their forearms on the desk surface.

2.8 | EEG signal processing

The collected EEG data were band-pass filtered with a non-causal, eighth-order Butterworth filter with half-power frequencies of 0.6 and 30 Hz. Response-locked epochs (-180 to 710 ms from each response occurrence) were obtained for every response between the first response and the $n = -1$ response, for every trial. After that, a baseline correction was performed on each epoch by subtracting its mean value. We used three criteria for discarding outlier data. First, epochs were discarded when a voltage sample was greater than ± 100 µV (after baseline correction). Second, epochs containing eyes blinks and eyes movements artefacts were discarded by implementing the Step Function algorithm (Luck, 2014) with a 200-ms moving window in steps of 7 ms. In each window, the mean value of the first 100 ms was compared with the mean value of the last 100 ms; when the difference between both values was greater than ± 35 µV the whole epoch was discarded. Third, an epoch was discarded if any of its behavioural asynchronies fell outside the range -170 to $+80$ ms (values greatly exceeding the 1.5 IQR rule). In addition, a whole electrode was discarded if the number of its valid epochs was lower than 50% of its total number of epochs. This criterion was applied to all electrodes placed within the area of interest (AF3, F3, F7, FC5, FC6, F8, F4, AF4). According to this criterion, three subjects presented one invalid electrode; none of them presented more than one invalid electrode per condition. After the discarding process, there remained an average of 107 valid epochs per subject, condition and electrode (equivalent to an average of discarded epochs of 10.3% per subject, condition and electrode).

Stimulus-locked epochs were obtained from the valid resp-locked epochs (-100 ms to 350 ms from stimulus occurrence) in the pre-perturbation region ($n = -7$ through $n = -1$). On average, 73 stim-locked epochs were obtained per subject, condition and electrode. Then, an adaptation of the Adjar algorithm (see Appendix A) was applied over the stim-locked epochs to remove the ERP components not related to the auditory stimuli.

2.9 | Permutation testing for ERPs

We implemented a whole-subject permutation approach (Luck, 2014) to test the difference between NORMAL and HIGH ERPs.

We first defined a temporal window in which the ERPs were compared. Previous works consistently suggest a specific window location and also point to a directional difference between the attended versus non-attended ERPs (Hansen & Hillyard, 1980; Näätänen et al., 1978; Pereira et al., 2014; Teder et al., 1993). In a selective attention paradigm with auditory dichotic stimuli, Teder finds that in latencies going from N1 to P2 the stimuli-attended ERP is negatively displaced with respect to the ERP of the non-attended stimuli, with the highest difference occurring around 180 ms (for an ISI of 480). Based on this, we defined a window centred in 180 ms and going from 110 to 250 ms. Next, we created a null distribution by random permutation of the data as follows. Eleven subjects were randomly selected from the group of 22 subjects. We computed an average ERP within this subgroup and called it 'surrogate NORMAL'. We averaged the ERPs from the remaining 11 subjects and called it 'surrogate HIGH'. We performed a non-paired t test between the 'surrogate NORMAL' and 'surrogate HIGH' ERPs for every time sample in the entire window of comparison. We found the maximum of all t values in the window and called it t_{max}. We repeated this procedure 1,000 times to obtain a distribution of 1,000 t_{max} values. Lastly, we performed a non-paired t test between the actual NORMAL and HIGH ERPs (i.e. without permutation) for every time sample in the comparison window, and labelled the time bin as significant if the actual t value was higher than the 95 percentile of the t_{max} distribution (i.e. one-tailed). The region where we found significant differences is displayed as a shaded rectangle in Figure 5a.

3 | RESULTS

Figure 2a shows the average time series for each subject and each condition. In the pre-perturbation region ($n = -7$ to $n = -1$, both inclusive) the asynchrony is relatively constant, meaning the subject is synchronized. When perturbation occurs ($n = 0$) a forced error is induced and thus the asynchrony changes abruptly. After that, the subject resynchronizes to a new, slightly different baseline in a few taps.

3.1 | Attention improves accuracy

Consider first the pre-perturbation region. In this region no effect of perturbation sign is expected so we averaged across POS and NEG levels. The pre-perturbation mean asynchrony MA_{pre} shifts towards less negative values (greater accuracy) when either feedback is added or attention is increased (Figure 2b), but only the attention factor is significant. That is, attention improves accuracy, according to what can be expected intuitively. The additional auditory feedback, also in favour to what is expected as it adds information, shows a similar tendency to produce a less negative MA but the effect is not significant (two-way ANOVA, with attention as between-subject factor and feedback as within-subject factor; attention $F(1, 42) = 9.84$, $p = 0.0035$, Cohen's $d = 1.08$; feedback $F(1, 42) = 1.74$, $p = 0.19$, Cohen's $d = 0.59$; interaction $F(1, 42) = 0.03$, $p = 0.86$).

Figure 2c shows the standard deviation of asynchronies in the pre-perturbation region SD_{pre}. Both increasing attention and adding auditory feedback make SD_{pre} decrease, that is, both improve precision, although their effects are not significant (two-way ANOVA, with attention as between-subject factor, feedback as within-subject factor, and subject as random factor; attention $F(1, 42) = 3.56$, $p = 0.061$, Cohen's $d = 0.57$; feedback $F(1, 42) = 1.35$, $p = 0.25$, Cohen's $d = 0.23$; interaction $F(1, 42) = 0.02$, $p = 0.88$).

We did not find any clear evolution of the performance of the subjects along the experiment that could be construed as tiredness/adaptation/learning, despite being a long and repetitive task. See Figures B3 and B4.

In summary, in the pre-perturbation region both attention and auditory feedback behave as intuitively expected with regard to accuracy and precision (i.e. both make them improve, albeit some effects are non-significant in our study).

3.2 | Attention and feedback have opposite effects on resynchronization

We now show the subject's behaviour in the resynchronization phase. Perhaps not surprisingly, we found that a high level of attention improves the resynchronization behaviour of the subject, measured as a significant increase in the efficiency to reach the post-perturbation baseline (Figure 3). On the other hand, the resynchronization efficiency diminishes when auditory feedback is added, which is an unexpected result. Furthermore, the resynchronization efficiency is greater after a positive perturbation (increased period) than after a negative one (decreased period), suggesting asymmetric behaviour in line with previous reports (Bavassi et al., 2013; López & Laje, 2019; three-way ANOVA, with attention as between-subject factor and feedback and perturbation sign as within-subject factors; attention $F(1, 42) = 5.75$, $p = 0.021$, Cohen's $d = 0.68$; feedback $F(1, 42) = 11.01$, $p = 0.0019$, Cohen's $d = 0.65$; perturbation sign $F(1, 42) = 6.56$, $p = 0.014$, Cohen's $d = 0.61$; att x fbk interaction $F(1, 42) = 0.005$, $p = 0.94$;

att × sign interaction $F(1, 42) = 0.23$, $p = 0.62$; fbk × sign interaction $F(1, 42) = 1.09$, $p = 0.30$; att × fbk × sign interaction $F(1, 42) = 2.65$, $p = 0.11$). Notably, none of the interaction effects is significant, which suggests that either adding or removing the auditory feedback does not modify the subject's attention and allows us to discard a potential confounding (i.e. attention and feedback are independent variables).
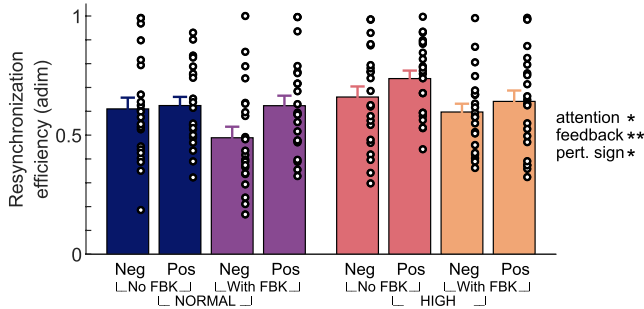
3.3 | Opposite effects of perturbative versus non-perturbative manipulations of the period

We found that a step-change perturbation produces a change in the accuracy (Figure 4a), in line with our previous reports (Bavassi et al., 2013).

In this work we extend and generalize our previous results. Figure 4a shows the MA change between the pre- and post-perturbation regions, ΔMA, for every experimental condition. All three factors have significant effects on ΔMA: on average, high attention leads to a more negative ΔMA (i.e. subject less accurate after perturbation), whereas the addition of auditory feedback shifts the post-perturbation MA towards more positive values (subject more accurate
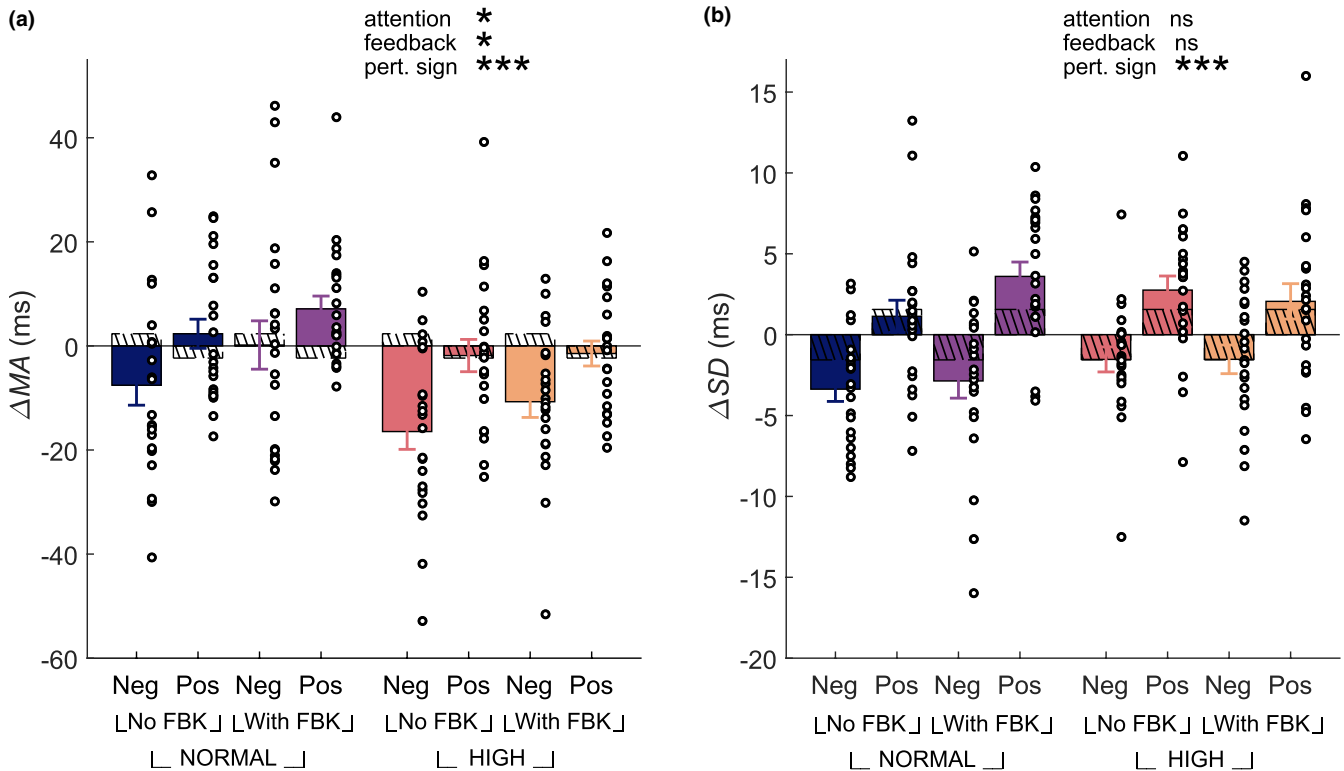

FIGURE 3 Resynchronization efficiency (bars: mean ± standard error across subjects; each dot represents a subject). Increased attention makes the resynchronization efficiency increase; counterintuitively, the additional feedback makes it decrease. Positive perturbations resynchronize with higher efficiency than negative ones. Attention $p = 0.021$; feedback $p = 0.0019$; perturbation sign $p = 0.014$

FIGURE 4 (a) An ISI perturbation induces a pre-post change of accuracy (ΔMA), and attention, feedback and perturbation sign have all significant effects (coloured bars: mean ± standard error across subjects; each dot represents a subject; attention $p = 0.010$; feedback $p = 0.010$; perturbation sign $p = 0.00041$). The direction of the change is in general not consistent with the known dependence of MA on the ISI when ISI is varied across trials (hatched bars represent expected values based on the literature, see Figure B1a). (b) An ISI perturbation induces a pre-post change of precision (ΔSD), with a significant effect of perturbation sign. The NEG and POS perturbations make the subjects more and less precise, respectively (coloured bars: mean ± standard error across subjects; each dot represents a subject; attention and feedback not significant; perturbation sign $p = 10^{-9}$). This result roughly agrees with the expected values (hatched bars) from the known dependence of the SD on the ISI (see Figure B1b)
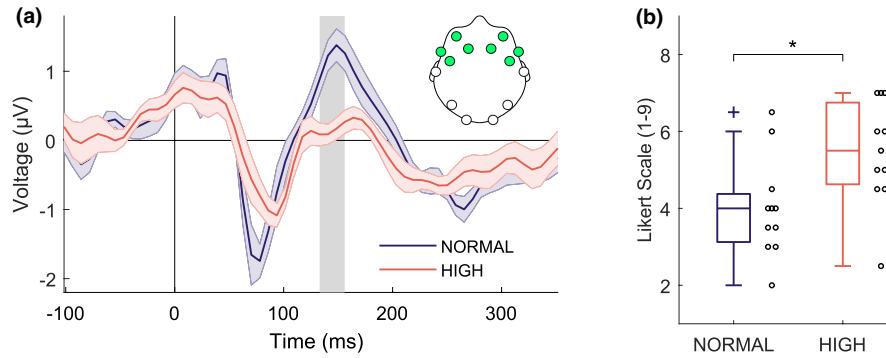
FIGURE 5 (a) Auditory-related ERPs for NORMAL and HIGH attention in the fronto-central region (stimulus-aligned data; $t = 0$ is the occurrence of the stimuli). The shaded rectangle represents the region where the difference is significant ($t = 133$ ms through 156 ms). Mean ± standard error across subjects; NORMAL and HIGH levels were averaged across feedback conditions and perturbation signs. (b) Mental workload (subjective report of the mental effort required during the task); attention $p = 0.015$

after perturbation); the positive perturbations make the MA change less than the negative ones (three-way ANOVA, with attention as between-subject factor and feedback and perturbation sign as within-subject factors; attention $F(1, 42) = 7.19$, $p = 0.010$, Cohen's $d = 0.80$; feedback $F(1, 42) = 6.96$, $p = 0.011$, Cohen's $d = 0.58$; perturbation sign $F(1, 42) = 15.26$, $p = 0.00041$, Cohen's $d = 1.13$; att x fbk interaction $F(1, 42) = 0.73$, $p = 0.39$; att × sign interaction $F(1, 42) = 0.38$, $p = 0.53$; fbk × sign interaction $F(1, 42) = 1.66$, $p = 0.20$; att × fbk × sign interaction $F(1, 42) = 0.22$, $p = 0.63$).

The results in the previous paragraph are not consistent with the expected difference $\text{MA}\left(\text{ISI}_{\text{post}}\right) - \text{MA}\left(\text{ISI}_{\text{pre}}\right)$ based on the known dependence of the mean asynchrony on the ISI for isochronous trials (i.e. non-perturbative ISI manipulations; Repp, 2003). As shown in our Figure B1a, MA is a decreasing function of ISI when the ISI is kept constant during a trial (varied across trials as a control parameter). The predicted values according to this known dependence are shown as hatched bars in Figure 4a: a negative ΔMA for perturbations that increase the period (i.e. positive perturbations), and a positive ΔMA for perturbations that decrease the period (i.e. negative perturbations as illustrated by hatched bars in Figure 4a). That prediction is almost always opposite to what we get in our experiment. This suggests that a period perturbation has an effect of its own, different from that obtained by non-perturbative manipulations of the period.

Precision is also affected by a step-change perturbation. Figure 4b shows the pre-post change in the intra-trial SD (ΔSD) for every condition. The perturbation sign is the only parameter with a significant effect: negative perturbations produce a decrease in SD (subjects become more precise after perturbation), while the positive perturbations lead to an increase in SD (subjects less precise after perturbation; three-way ANOVA, with

attention as between-subject factor and feedback and perturbation sign as within-subject factors; attention $F(1, 42) = 1.43$, $p = 0.23$, Cohen's $d = 0.36$; feedback $F(1, 42) = 0.80$, $p = 0.37$, Cohen's $d = 0.25$; perturbation sign $F(1, 42) = 61.1$, $p = 9.9 \times 10^{-10}$, Cohen's $d = 2.13$; att × fbk interaction $F(1, 42) = 2.08$, $p = 0.15$; att × sign interaction $F(1, 42) = 1.65$, $p = 0.20$; fbk × sign interaction $F(1, 42) = 0.20$, $p = 0.65$; att × fbk × sign interaction $F(1, 42) = 0.92$, $p = 0.34$). These results are roughly consistent with the known dependence of SD on the ISI from isochronous trials. Previous works with non-perturbative manipulation of the period (e.g. Repp, 2003) showed that the SD is an increasing function of ISI (see Figure B1b). In our work, the observed ΔSD values are consistent with such dependence, with expected values represented by hatched bars in Figure 4b.

3.4 | Electrophysiological correlates of attention level

In order to show that our task makes the subjects reliably change the attention level, we measured two different markers of the level of attention: auditory stimulus-related ERPs (neural correlate of attention), and mental workload (subjective report of mental effort).

Figure 5a shows the auditory stimulus-related ERPs corresponding to each level of attention in the pre-perturbation region (feedback conditions and perturbation signs averaged). As described in the Introduction, in a paced finger-tapping task an overlap is expected between auditory (stimulus-related) components and sensorimotor (tap-related) components. According to the literature, the neural correlates of attention rely on the auditory, stimulus-related components so we developed a novel algorithm called TapAdjar (see Appendix A) to isolate the auditory components from the

whole EEG signal. Significant differences can be observed between attention levels in the P2 component according to what is described in the literature (amplitude of P2 diminishes in high attention conditions; Hansen & Hillyard, 1980; Näätänen et al., 1978; Pereira et al., 2014; Teder et al., 1993); specifically, the HIGH condition shows a large deflection from NORMAL around P2, with significant differences in the shaded rectangle going from 133 to 156 ms after the onset of the stimuli (bootstrapping by t_{max} permutation approach; see Section 2).

The mental workload results show an effect of attention that is consistent with the ERP results: subjects in the HIGH attention group report significantly higher mental effort compared to the NORMAL attention subjects (Figure 5b); Likert scale, 1 = very very low effort, 9 = very very high effort; HIGH: 5.45 ± 1.36, mean $\pm$ SD; 5.50 ± 1.04, median $\pm$ MAD. NORMAL: 4.00 ± 1.30, mean $\pm$ SD; 4 ± 0.90, median $\pm$ MAD; Kruskal–Wallis test, $p = 0.016$, Cohen's $d = 1.08$, Hedge's $d = 1.05$. Scores for each single condition are shown in Figure B5.

4 | DISCUSSION

4.1 | SMS is vulnerable to temporal attention

We found significant effects of attention on several measures of SMS in auditorily paced finger tapping. It seems natural to think that higher attention leads to a more accurate behaviour, but it should also be taken into account that attention can be oriented to different aspects of a task, for instance to different sensory modalities or semantic structures stored in memory (Posner, 1980). The significant effect we found may be related to the fact that attention is oriented specifically to a temporal aspect of the task (tempo change). It is worth considering whether we would observe the same results if attention was oriented to other aspects of the stimuli sequence, like stimulus intensity.

Our results are in clear contrast with the hypothesis by Miyake et al. (2004) that the error correction mechanism is an automatic process not affected by attention. In fact, our results reveal that the error correction mechanism in a paced finger-tapping task is affected by attention when it is oriented to a temporal aspect of the task. It is possible that they did not observe any effect of diverting attention because they used a non-temporal secondary task (word recall). Indeed, previous evidence suggests that two temporal tasks performed at once interfere with each other (Brown, 1997; although possible interferences between consecutive intervals should be controlled for, see Karmarkar & Buonomano, 2007). It is interesting to note that word recall did not interfere with finger tapping in their study but it certainly did with an interval

discrimination task in other experiment (Miller et al., 1978). This might suggest that we have to distinguish between different types of temporal tasks as it seems that they possibly do not share the same cognitive resources. Finally, it is also possible that finger tapping could be performed with a high degree of automaticity, and therefore there would be resources left to perform a secondary task, such as the one used by Miyake et al., without interfering with the main task. This hypothesis could be probed by increasing the difficulty of the secondary task.

We found evidence that tapping in NORMAL attention is done with a certain degree of automaticity, as the mental workload reported by the subjects is lower than in HIGH. In view of these findings, we can consider that finger tapping is not a strongly automatic task, in the sense that Kahneman and Chajczyk define it: a task whose performance does not improve when attention is oriented towards it and does not get worse when attention is diverted from it (Kahneman & Chajczyk, 1983). Instead, we can conclude that finger tapping fits better in the partially automatic category because of both its automaticity characteristics and its susceptibility to attention. According to them, a partially automatic task can be accomplished without attention although its performance improves when attention is oriented towards it.

The resynchronization phase is also affected by attention. We found that focusing the attention on the tempo change increases the resynchronization efficiency. This result is consistent with the study by Repp and Keller (2004), in which he observes a slower correction (smaller PCR) for conditions in which attention is diverted from the tempo change. According to the prior entry hypothesis, which states that an attended stimulus is detected faster than an unattended one (Spence & Parise, 2010), we can suppose that subjects in HIGH perceive the tempo change earlier than subjects in NORMAL, leaving them with more time for the processing. This may be one of the causes explaining the differences observed in resynchronization efficiency between the attention conditions.

4.2 | Neural correlates of attention in SMS

We found differences in the early components (in particular, P2) of the auditory-related ERPs between levels of attention. Such differences have been interpreted in the literature as an effect of attention (Alho et al., 1994a; Eimer & Forster, 2003a; Hansen & Hillyard, 1980; Teder et al., 1993) and are compatible with the negative processing wave, which is a negative displacement of the attended ERP versus the non-attended ERP (HIGH and NORMAL in our work, respectively). The negative processing wave begins in the ascending deflection of N1; it may last a few hundred milliseconds and appears mainly in the front-central region (Näätänen, 1982). In an auditory and dichotic stimuli

paradigm, Teder and coworkers report an amplitude modulation at P2 latencies in the same direction as ours: the amplitude of P2 in the attended ERP is significantly lower than in the unattended ERP (see Teder et al., 1993, Figures 1 and 2, ISI 480 ms). Näätänen (1990) showed a deflection of the P2 component in auditory ERPs to attended stimuli (relative to unattended) with comparable latency, amplitude and sign as ours. Eimer and Forster (2003a) showed that, at P2 latencies, somatosensory ERPs to attended stimuli were less positive than unattended (on the other hand, a visually evoked P2 might show the opposite effect: Luck et al., 2000). All considered, this gives support to our claim that in our task we are indeed controlling the level of attention. It is important to note that the stimuli sequence is the same for both levels of attention (as it only includes the pre-perturbation region), thus any differences in the ERPs may be attributed unequivocally to different psychological states.

4.3 | SMS and auditory feedback

Auditory feedback decreases the mean asynchrony in the pre-perturbation region, in line with Aschersleben's report 1995, although the effect is not significant. We suppose that this difference with Aschersleben's results is due to the type of phones we used. Half of the subjects (no EEG recording) wore headphones which suppressed most of the surrounding noise; the other half (EEG recording) used earphones because the EEG did not leave enough space for headphones. Although a statistical test revealed no significant effect of phones factor on asynchrony (see Section 2), the earphone group was only partially isolated from the surrounding noise and the direct sound of taps was more audible, a situation that we believe contributed to diluting the effect of the auditory feedback when it was added.

We found that the resynchronization efficiency decreases when auditory feedback is added. From a naive point of view, this result is unexpected as the auditory feedback would give the subject access to more or better information. We hypothesize that stimuli and feedback are interfering with each other as both are in the same sensory modality and temporally overlapped. It is possible that auditory feedback prevents the subject from recognizing promptly the new stimuli period, and as a consequence he/she might be unable to perform the corrections efficiently.

4.4 | Asymmetric behaviour in SMS

We observed that positive perturbations have a higher resynchronization efficiency than negative perturbations of the same magnitude. This asymmetry is in accordance with results from different paradigms (e.g. perception vs. production), different perturbation types, and different measures to evaluate the resynchronization process (Bavassi et al., 2013; Jang et al., 2016; Praamstra et al., 2003; Repp, 2001, 2002a, 2002b, 2011; Repp & Keller, 2004). This asymmetry could be a signature of a non-linear correction mechanism (Bavassi et al., 2013; López & Laje, 2019). Jang et al. (2016) posed an interesting hypothesis for this asymmetry: that, when seen from a stimulus-aligned perspective, positive perturbations allow for a little additional time for attention and/or preparation to kick in (because the stimulus is delayed and attention is allegedly already half-way up when the stimulus occurs), while negative perturbations do not. Yet, there is no agreed explanation for positive perturbations having higher resynchronization efficiency.

Our results for ΔMA (Figure 4a) show that a step-change period perturbation has its own effect on mean asynchrony, and it is opposite to the effect of non-perturbative changes to the ISI. The effect on average is greater for negative perturbations; all in agreement with previous results (Bavassi et al., 2013). Our main conclusion from this result is that it reveals that post-perturbation MA is a variable to be controlled as it depends on the three factors: attention, feedback and perturbation sign. Therefore, if the resynchronization process is to be evaluated it is important to define a magnitude that could comprise this effect. This is the reason that encouraged us to define the resynchronization efficiency, a magnitude containing in its definition the change in post-perturbation MA. Finally, the fact that a perturbation to the period has its own effect on the MA speaks against the inclusion of more than a single perturbation per trial, a common practice in some experimental designs (e.g. Jang et al., 2016; Praamstra et al., 2003).

5 | CONCLUSIONS

Our results indicate that the error correction mechanism in paced finger tapping may be affected by higher cognitive functions like attention, both in its stationary and resynchronization phases. In the stationary phase attention improves accuracy, while in the resynchronization phase it increases efficiency. Analysis of the auditory feedback manipulation leads us to find counterintuitive results, namely that resynchronization efficiency is diminished when auditory feedback is added. We also found that the amplitude of the P2 component of the auditory ERPs is diminished in the HIGH attention condition, for the first time in SMS and in line with previous literature on different paradigms. In order to achieve this we developed the TapAdjar algorithm to separate auditory- and response-related ERPs that are naturally overlapped in time in a paced finger-tapping task. Finally, the asymmetry in the resynchronization efficiency in front of positive versus negative perturbations adds a converging line of evidence regarding non-linearity of the underlying

mechanism in charge of SMS even for perturbation sizes traditionally labelled as small (10% of the ISI).

CONFLICT OF INTEREST
The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS
LV and RL conceived the study and designed the experiments. LV conducted the experiments and performed data analysis. All authors interpreted the data and wrote the manuscript.

PEER REVIEW
The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15245.

DATA AVAILABILITY STATEMENT
Data and code are available at the Sensorimotor Dynamics Lab website: http://ldsm.web.unq.edu.ar/attention2021/.

ORCID
Leonardo Versaci https://orcid.org/0000-0002-0554-6104
Rodrigo Laje https://orcid.org/0000-0003-1481-303X

REFERENCES
Alho, K., Teder, W., Lavikainen, J., & Näätänen, R. (1994a). Strongly focused attention and auditory event-related potentials. *Biological Psychology*, *38*, 73–90. https://doi.org/10.1016/0301-0511(94)90050-7

Alho, K., Teder, W., Lavikainen, J., & Näätänen, R. (1994b). Strongly focused attention and auditory event-related potentials. *Biological Psychology*, *38*, 73–90. https://doi.org/10.1016/0301-0511(94)90050-7

Aschersleben, G., & Prinz, W. (1995). Synchronizing actions with events: The role of sensory information. *Perception & Psychophysics*, *57*, 305–317.

Badcock, N. A., Mousikou, P., Mahajan, Y., de Lissa, P., Thie, J., & McArthur, G. (2013). Validation of the Emotiv EPOC(®) EEG gaming system for measuring research quality auditory ERPs. *PeerJ*, *1*, e38.

Badcock, N. A., Preece, K. A., de Wit, B., Glenn, K., Fieder, N., Thie, J., & McArthur, G. (2015). Validation of the Emotiv EPOC EEG system for research quality auditory event-related potentials in children. *PeerJ*, *3*, e907. https://doi.org/10.7717/peerj.907

Bavassi, L., Kamienkowski, J. E., Sigman, M., & Laje, R. (2017). Sensorimotor synchronization: Neurophysiological markers of the asynchrony in a finger-tapping task. *Psychological Research Psychologische Forschung*, *81*, 143–156.

Bavassi, M. L., Tagliazucchi, E., & Laje, R. (2013). Small perturbations in a finger-tapping task reveal inherent nonlinearities of the underlying error correction mechanism. *Human Movement Science*, *32*, 21–47.

Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Perception & Psychophysics*, *59*, 1118–1140.

Caspi, A. (2002). The synchronization error: Attentional and timing aspects. Unpublished doctoral dissertation, TelAviv University, Israel.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for p2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, *115*, 732–744.

Crozier, D., Zhang, Z., Park, S.-W., & Sternad, D. (2019). Gender differences in throwing revisited: Sensorimotor coordination in a virtual ball aiming task. *Frontiers in Human Neuroscience*, *13*, 231.

Dione, M., & Delevoye-Turrell, Y. (2015). Testing the co-existence of two timing strategies for motor control in a unique task: The synchronization spatial-tapping task. *Human Movement Science*, *43*, 45–60.

Drewing, K., Aschersleben, G., & Li, S.-C. (2006). Sensorimotor synchronization across the life span. *International Journal of Behavioral Development*, *30*, 280–287.

Eimer, M., & Forster, B. (2003a). Modulations of early somatosensory ERP components by transient and sustained spatial attention. *Experimental Brain Research*, *151*, 24–31. https://doi.org/10.1007/s00221-003-1437-1

Eimer, M., & Forster, B. (2003b). The spatial distribution of attentional selectivity in touch: Evidence from somatosensory ERP components. *Clinical Neurophysiology*, *114*, 1298–1306. https://doi.org/10.1016/S1388-2457(03)00107-X

Franěk, M., Mates, J., Radil, T., Beck, K., & Pöppel, E. (1991). Finger tapping in musicians and nonmusicians. *International Journal of Psychophysiology*, *11*, 277–279. https://doi.org/10.1016/0167-8760(91)90022-P

Gerloff, C., Toro, C., Uenishi, N., Cohen, L. G., Leocani, L., & Hallett, M. (1997). Steady-state movement-related cortical potentials: A new approach to assessing cortical activity associated with fast repetitive finger movements. *Electroencephalography and Clinical Neurophysiology*, *102*, 106–113.

Hansen, J. C., & Hillyard, S. A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology*, *49*, 277–290.

Iversen, J. R., Patel, A. D., Nicodemus, B., & Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition*, *134*, 232–244.

Jang, J., Jones, M., Milne, E., Wilson, D., & Lee, K. H. (2016). Contingent negative variation (CNV) associated with sensorimotor timing error correction. *NeuroImage*, *127*, 58–66.

Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 497–509.

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: Encoding time in neural network states. *Neuron*, *53*, 427–438.

Kolers, P. A., & Brewster, J. M. (1985). Rhythms and responses. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 150–167. https://doi.org/10.1037/0096-1523.11.2.150

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology*, *63*, 2220–2230.

Krause, V., Pollok, B., & Schnitzler, A. (2010). Perception in action: The impact of sensory information on sensorimotor synchronization in musicians and non-musicians. *Acta Psychologica*, *133*, 28–37.

Law, L. N., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the profile of music perception skills. *PLoS One*, *7*, e52508.

López, S. L., & Laje, R. (2019). Spatiotemporal perturbations in paced finger tapping suggest a common mechanism for the processing of time errors. *Scientific Reports*, *9*, 17814. https://doi.org/10.1038/s41598-019-54133-x

Luck, S. J. (2014). *An introduction to the event-related potential technique*. Chap. 6. MIT Press.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, *4*, 432–440. https://doi.org/10.1016/S1364-6613(00)01545-X

Mates, J. (1994). A model of synchronization of motor acts to a stimulus sequence. II. Stability analysis, error estimation and simulations. *Biological Cybernetics*, *70*, 475–484.

Mates, J., & Aschersleben, G. (2000). Sensorimotor synchronization: The impact of temporally displaced auditory feedback. *Acta Psychologica*, *104*, 29–44.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience*, *36*, 313–336.

Miller, G. W., Hicks, R. E., & Willette, M. (1978). Effects of concurrent verbal rehearsal and temporal set upon judgments of temporal duration. *Acta Psychologica*, *42*, 173–179.

Miyake, Y., Onishi, Y., & Pöppel, E. (2004). Two types of anticipation in synchronization tapping. *Acta Neurobiologiae Experimentalis*, *64*, 415–426.

Mueller, V., Brehmer, Y., Von Oertzen, T., Li, S.-C., & Lindenberger, U. (2008). Electrophysiological correlates of selective attention: A lifespan comparison. *BMC Neuroscience*, *9*, 1–21.

Näätänen, R. (1982). Processing negativity: An evoked-potential reflection of selective attention. *Psychological Bulletin*, *92*, 605–640.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, *13*, 201–233.

Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*, 313–329.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*, 63–71.

Pereira, D. R., Cardoso, S., Ferreira-Santos, F., Fernandes, C., Cunha-Reis, C., Paiva, T. O., Almeida, P. R., Silveira, C., Barbosa, F., & Marques-Teixeira, J. (2014). Effects of inter-stimulus interval (ISI) duration on the N1 and P2 components of the auditory event-related potential. *International Journal of Psychophysiology*, *94*, 311–318.

Pietto, M. L., Gatti, M., Raimondo, F., Lipina, S. J., & Kamienkowski, J. E. (2018). Electrophysiological approaches in the study of cognitive development outside the lab. *PLoS One*, *13*, e0206983. https://doi.org/10.1371/journal.pone.0206983

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25. https://doi.org/10.1080/00335558008248231

Praamstra, P., Turgeon, M., Hesse, C. W., Wing, A. M., & Perryer, L. (2003). Neurophysiological correlates of error correction in sensorimotorsynchronization. *NeuroImage*, *20*, 1283–1297.

Repp, B. H. (2001). Processes underlying adaptation to tempo changes in sensorimotor synchronization. *Human Movement Science*, *20*, 277–312.

Repp, B. H. (2002a). Automaticity and voluntary control of phase correction following event onset shifts in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 410–430. https://doi.org/10.1037/0096-1523.28.2.410

Repp, B. H. (2002b). Phase correction in sensorimotor synchronization: Nonlinearities in voluntary and involuntary responses to perturbations. *Human Movement Science*, *21*, 1–37. https://doi.org/10.1016/S0167-9457(02)00076-3

Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, *35*, 355–370.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*, 969–992.

Repp, B. H. (2010). Sensorimotor synchronization and perception of timing: Effects of music training and task experience. *Human Movement Science*, *29*, 200–213.

Repp, B. H. (2011). Tapping in synchrony with a perturbed metronome: The phase correction response to small and large phase shifts as a function of tempo. *Journal of Motor Behavior*, *43*, 213–227.

Repp, B. H., & Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: Effects of intention, attention, and awareness. *Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *57*, 499–521.

Robinson, K., & Rudge, P. (1977). Abnormalities of the auditory evoked potentials in patients with multiple sclerosis. *Brain*, *100*(Pt 1), 19–40.

Schultz, B. G., & van Vugt, F. T. (2016). Tap arduino: An arduino microcontroller for low-latency auditory feedback in sensorimotor synchronization experiments. *Behavior Research Methods*, *48*, 1591–1607.

Spence, C., & Parise, C. (2010). Prior-entry: a review. *Consciousness and Cognition*, *19*, 364–379.

Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), *Psychology of learning and motivation* (Vol. *55*, pp. 37–76). Academic Press.

Teder, W., Alho, K., Reinikainen, K., & Näätänen, R. (1993). Interstimulus interval and the selective-attention effect on auditory ERPs: "N1 enhancement" versus processing negativity. *Psychophysiology*, *30*, 71–81.

Woldorff, M. G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. *Psychophysiology*, *30*, 98–119.

Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology*, *102*, 3191–3202.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

APPENDIX A

TapAdjar: AN ALGORITHM TO REMOVE DISTORTION DUE TO TEMPORALLY ADJACENT ERPs

Introduction

In paced finger tapping with stimuli periods of a few hundreds milliseconds, the subject probably monitors his/her own performance from the last tap while engaging in motor preparation/execution for the next tap and tracking any change in the stimuli sequence (Bavassi et al., 2013, 2017; López & Laje, 2019). The neural processes involved have durations comparable to the stimuli period, so some degree of overlapping between the response- and stimulus-related ERPs is expected. It would be desirable to have a method to separate both contributions in order to obtain an estimation of the isolated stimuli-related ERPs. By attenuating the distortions caused by the naturally overlapping response-related component, the isolated ERPs could be compared with the extensive literature on attention and auditory stimuli (Näätänen, 1982). To this end, in this appendix we propose an adaptation of an existing algorithm used in other experimental paradigms.

We base our proposal on Adjar, an algorithm designed by Woldorff (1993) for a selective attention paradigm with auditory and dichotic stimuli in conditions with short ISIs of 300 ms and shorter. Considering such values of ISIs, an overlap between the ERPs generated by the present and the subsequent stimuli is expected to occur. This algorithm is meant to attenuate the distortions due to adjacent ERPs, allowing in this way an estimation of the ERP that could be obtained if there was no overlap.

We now describe our adaptation of Adjar to paced finger tapping; we call it *TapAdjar*. It is based on the idealization that each epoch recorded by the EEG primarily consists of two large, overlapping components that cannot be measured directly: one is related to the stimulus (beep) and the other is related to the corresponding response (tap; Figure A1a). The general idea is to estimate the 'pure' response-related ERP (Figure A1b) and to subtract it from the recorded ERP (Figure A1c); the result will be a stimulus-related ERP estimate with attenuated contamination from response-related sensorimotor potentials.

Procedure

Step 1

We consider the pre-perturbation region only. We define the m,n-th epoch as the recorded EEG activity around the n-th stimulus (beep) and its corresponding response (tap) of the m-th trial, for each electrode, subject and condition. The

asynchrony corresponding to the m,n-th epoch is the temporal difference between the onset of the m,n-th response and the m,n-th stimulus, $e_{m,n} = t_{\text{resp}_{m,n}} - t_{\text{stim}_{m,n}}$. We assume that the recorded EEG is the superposition of two main components (Hypothesis 1): (a) a voltage variation related to both the m,n-th stimulus and the effects of attention; we call this component $s_{m,n}(t)$; (b) a voltage variation related to the sensorimotor activity of the m,n-th tap; we call this component $r_{m,n}(t)$ (Figure A1a). These two components are usually overlapped, a condition that turns them individually inaccessible when using a single direct measure. Considering the $e_{m,n}$ asynchrony between the onsets of response and stimulus, the EEG voltage of m,n-th epoch can be written as

$$v_{m,n}(t) = s_{m,n}(t) + r_{m,n}(t - e_{m,n}) \tag{3}$$

Then, the epochs $v_{m,n}(t)$ are aligned to responses and averaged:

$$R(t) = \frac{1}{N}\sum_{m,n} v_{m,n}(t + e_{m,n}) \tag{4}$$

where N is the total number of epochs for each electrode, subject and condition; $R(t)$ is the commonly called averaged resp-locked ERP (Figure A1b). We consider $R(t)$ to be an estimate of $r_{m,n}(t)$, the pure response-related component (Hypothesis 2).

Step 2

We estimate $s_{m,n}(t)$, the component related to the auditory m,n-th stimulus. From Equation (3) we obtain

$$s_{m,n}(t) = v_{m,n}(t) - r_{m,n}(t - e_{m,n}) \tag{5}$$

where $r_{m,n}(t)$ can be replaced by its estimation $R(t)$ (Hypothesis 3) appropriately shifted in time according to the asynchrony of that epoch. The result is $s'_{m,n}(t)$, an estimation of $s_{m,n}(t)$:

$$s'_{m,n}(t) = v_{m,n}(t) - R(t - e_{m,n}) \tag{6}$$

The process is repeated for each of the N epochs. Finally, the resulting $s'_{m,n}(t)$ are aligned to the stimulus and averaged, obtaining the 'corrected' averaged stim-locked ERP, or in other words, the estimate of the 'pure' ERP related to the auditory stimuli (Figure A1c):

$$S(t) = \frac{1}{N}\sum_{m,n} s'_{m,n}(t) \tag{7}$$

Justification of the hypotheses

Our proposal requires us to assume some hypotheses that, with the exception of the first one, are equivalent to those
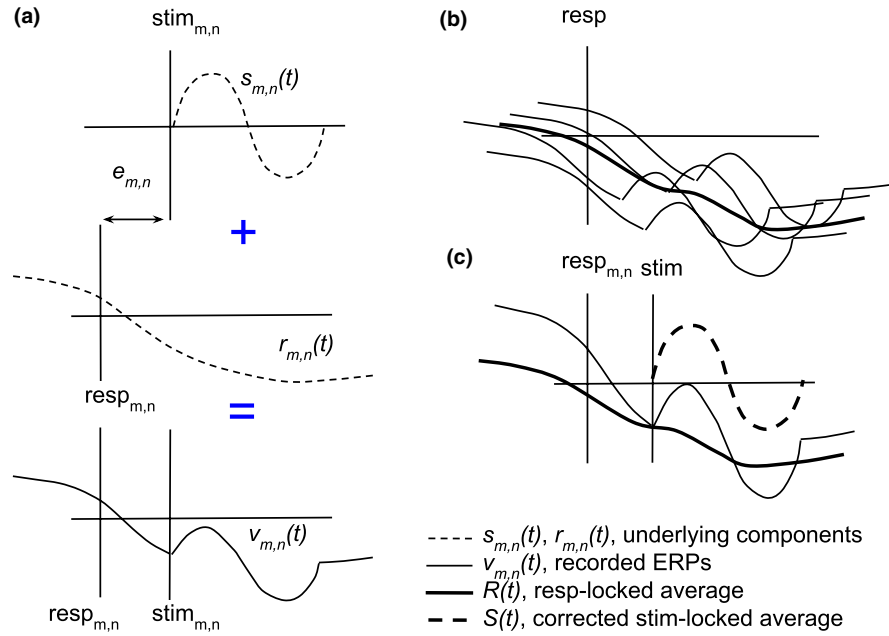
FIGURE A1 (a) Underlying (inaccessible) components in the m, n-th epoch (trial m, beep n). $s_{m,n}(t)$ is the auditory/stimuli-related component; high-frequency content. $r_{m,n}(t)$ is the response-related component; high- and low-frequency content. $v_{m,n}(t) = s_{m,n}(t) + r_{m,n}(t - e_{m,n})$ is the recorded epoch. $e_{m,n}$ is the asynchrony of m, n-th epoch. (B) Single epochs $v_{m,n}(t)$ are aligned to taps and averaged, thus attenuating the high-frequency content. The result $R(t)$ is the average response, that is an estimate of $r_{m,n}(t)$. (c) $R(t)$ is shifted in time according to $e_{m,n}$ and subtracted from $v_{m,n}(t)$ to obtain $s'_{m,n}(t)$ (not shown). This is repeated for all epochs and the results $s'_{m,n}(t)$ are aligned to the stimulus and averaged to obtain $S(t)$ as the estimate of the auditory stimulus-related ERP

assumed by Woldorff. Here, we describe them using the terminology of finger tapping.

Hypothesis 1 *Each epoch is composed of only two large components: one related to the stimulus and the other related to the corresponding response. We also consider that each epoch is the sum of the components pertaining only to that epoch, and we assume that there is a negligible overlap with adjacent epochs (first order of approximation, according to Woldorff terminology Woldorff terminology 1993).*

In a finger-tapping task, there is an overlap between the EEG activity related to stimulus and response (Praamstra et al., 2003). Despite the low-pass filtering effect of epoch averaging due to asynchrony variability, the stim-locked ERPs still show residuals resulting from the sensorimotor activity associated with the responses. This residual activity is clearly observed before the stimulus onset as a low-frequency negative deflection of the potential, and on later latencies it is found overlapped with the auditory ERP components (Praamstra et al., 2003; Robinson & Rudge, 1977). According to the literature, it is reasonable to suppose that this low-frequency deflection is the post-movement potential observed by Gerloff (Gerloff et al., 1997; Figure A2).

In addition to the components related to finger movement, we have to consider the activity generated by other causes.

For example, when the finger hits the sensor a tactile ERP is generated, presumably with frequencies and latencies similar to those of the auditory ERPs (Eimer & Forster, 2003a, 2003b). Besides, in conditions with additional auditory feedback we should take into account a response-related auditory ERP. In summary, the post-movement potential ERP, the tactile ERP and the auditory-feedback ERP can be considered related to the response.

As for the stimulus-related component, we take into account the auditory-stimulus ERP and the effect of attention. We could also consider the appearance of a readiness potential or even a contingent negative variation (CNV) relative to the sensor-switching movement of the finger (it takes place once per trial and only in HIGH conditions, see Section 2), yet our results did not reveal any deflection compatible with those potentials. Finally, the sensor-switching movement occurs outside the region in which we are comparing ERPs (pre-perturbation region), so no potentials related to the sensor-switching movement are considered.

Hypothesis 2 *When the number of averaged epochs is large enough, the resp-locked-ERP presents a negligible contamination from the stimulus-related components.*

As stated above, the stim-locked ERPs present a low-frequency negative deflection clearly visible previous to the onset of the stimulus. This deflection is a consequence
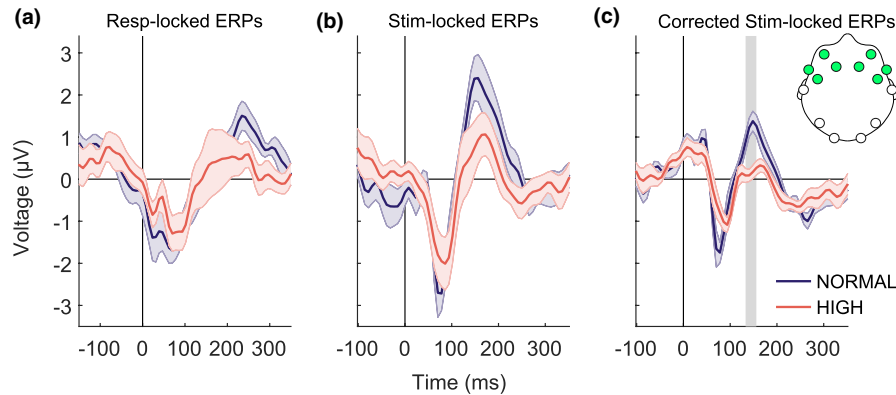
FIGURE A2 (a) Recorded resp-locked ERPs. (b) Recorded stim-locked ERPs. A low-frequency negative deflection is observed near the onset of the stimulus. (c) Corrected stim-locked ERPs (as plotted in Figure 5a). The negative deflection now is absent, and while the baseline is not flat at least both ERPs begin to rise from near zero voltage values. Also a first peak associable with P1 is observed, as it is expected from a canonic auditory ERP. Notes: In (a), the vertical dashed line represents the onset of the tap; in (b and c), represents the onset of the stimulus. The three panels show the general mean ERPs, where feedback conditions were averaged. The TapAdjar correction was applied for each subject, condition and electrode, before the averaging process was performed. Mean ± standard error across subjects

of the components related to the tap (Praamstra et al., 2003). We suppose that the activity related to the tap has both low-frequency components (due to the post-movement potential) and high-frequency components coming from the other contributions. Instead, the activity related to the stimuli has a diminished low-frequency content (Praamstra et al., 2003). This asymmetry in the frequency content between stimuli and taps has an important consequence: when epochs are aligned to the responses and averaged, high frequencies coming from the stimuli are averaged out by the natural variability of asynchronies. This fact makes it possible to isolate or, at least, obtain an estimate of the 'pure' resp-locked ERP. On the contrary, when epochs are aligned to stimuli and averaged, the asynchrony variability again smooths high frequencies in the tap-related activity, but not the low ones, which persist in the stim-locked ERP.

Hypothesis 3 *In terms of EEG activity, we assume that all the responses for a given subject, condition and electrode are equivalent; and that the averaged resp-locked ERP can be considered as an estimation of each individual response-related component; that is, $R(t) \sim r_{m,n}(t)$.*

We will now consider three possible causes of differences in individual response-related activity: overlap of adjacent epochs, EEG noise and asynchrony value; we now show reasons to safely ignore their effect. Regarding the first possible cause, an ISI as the one used in this work (600 ms) is considered large enough to minimize overlap between adjacent epochs (Alho et al., 1994b; Teder et al., 1993; Woldorff, 1993), so we only consider that there exists overlap between stimulus and tap activities belonging to the same

epoch; possible overlaps between epochs are considered as higher order corrections and so we ignore them (Hypothesis 1). The second reason that could cause two individual responses to be different is EEG noise, normally modeled as white noise (Woldorff, 1993); therefore we expect that when several epochs are averaged its random effect is considerably reduced. Finally, the naturally variable asynchrony value might make two individual responses different. That is, the EEG activity associated with a response whose asynchrony is −250 ms might be different to the one associated with a response whose asynchrony is +150 ms. Yet, as it was said in the main text, epochs with extreme values of asynchrony were discarded. We kept epochs with asynchronies between −170 and +80 ms, what reduces the possible variability of ERPs because of this reason.

Results

Figure A2 shows the grand-average ERPs for NORMAL and HIGH conditions (feedback conditions averaged) corresponding to the eight electrodes in the front-central region (AF3, F3, F7, FC5, FC6, F8, F4, AF4). Panel (b) shows the stim-locked ERPs before any corrections were made. It can be observed a low-frequency negative deflection of the potential previous to the onset of the stimulus (dashed vertical line). Praamstra gave evidence that this deflection is caused by the activity related to the tap (Praamstra et al., 2003). In later latencies, the superposition of the activity coming from the tap makes the stim-locked ERPs appear distorted with respect to the auditory canonical ERP shape, with the three well-defined peaks, P1, N1 and P2. Panel (a) shows the resp-locked ERPs. A bootstrapped comparison shows no

significant differences between NORMAL and HIGH ERPs in a window going from -100 to $+300$ ms relative to the onset of the tap. Panel (c) shows the stim-locked ERPs after the TapAdjar correction was applied (as plotted in Figure 5a). The TapAdjar algorithm was applied to each subject, condition and electrode before the average process was performed. We can observe now that the region previous to the onset of the stimulus does not show the negative deflection anymore, and although the ERPs do not show a flat baseline, at least both ERPs appear to rise from near zero voltage values, which is convenient for comparing absolute amplitude values between conditions. Also, we can observe a P1 peak which is not present in the stim-locked ERPs before correction, a fact that increases the similarity of the corrected ERPs with those elicited by an isolated auditory stimulus. Error bands in corrected ERPs are narrower than in uncorrected ERPs; this suggests that the TapAdjar algorithm doesn't have a random effect over the ERPs