Behavioral/Cognitive

Neural Correlates of Visuomotor Co-regulation in Social Synchronization

Stefano Uccelli,[1] Lucia Maria Sacheli,[1,2] Claudia De Bernardi,[1] Francantonio Devoto,[1] Gianpaolo Basso,[3,4] and Eraldo Paulesu[1,2]

[1]Department of Psychology and Milan Center for Neuroscience, University of Milano-Bicocca, Milano 20126, Italy, [2]IRCCS Orthopedic Institute Galeazzi, Milan 20157, Italy, [3]School of Medicine and Surgery, University of Milano-Bicocca, Milano 20126, Italy, and [4]Neuroradiology Unit, Fondazione IRCCS San Gerardo dei Tintori, Monza 20900, Italy

During social activities, people coordinate their movements by exchanging visuomotor information. Interpersonal coordination arises from two complementary processes, involving either voluntary (planned) or spontaneous (emergent) synchronization, whose neurofunctional underpinnings are currently unknown. We investigated the brain correlates of these two synchronization processes during an fMRI finger-tapping task. Thirty participants (17 females, 13 males) were scanned while reproducing a target tempo either synchronizing with (Joint Action, JA) or resisting (Non-interactive, NI) the partner's tapping, whose hand was always visible to the scanned participant. Faster, slower, or equal tempi were used to induce within-dyad tempo discrepancies. Results revealed the emergence of tempo contagion: participants tapped faster in response to their partner's faster taps and vice versa for slower taps. The magnitude of this interpersonal contagion effect was similar across conditions but associated with the activity of different neural structures. Tempo contagion correlated positively with lateral occipitotemporal cortex (LOTc) activations in the JA condition but negatively with cerebellar activations in the NI condition. This suggests that visuomotor information is exploited in opposite ways depending on the task instructions: the LOTc activity contributes to co-regulation to achieve synchronization in JA, whereas the cerebellar activity contributes to preserving individual motor stability in NI, preventing tempo contagion. A negative functional connectivity between the cerebellum and LOTc supports this latter result. These findings have implications for understanding the interplay between planned and emergent synchronization during motor interactions.

Key words: **entrainment; finger-tapping; interpersonal synchronization; joint action; visuomotor control**

Significance Statement

During social activities, people can synchronize with each other intentionally or automatically. Here, the neural correlates of visuomotor social synchronization are investigated for the first time. We scanned brain activity while dyads voluntarily synchronized with the partner or resisted social synchronization. Surprisingly, both tasks activated the same brain areas, but visuomotor information was exploited in opposite ways by different neural structures. Visual brain regions contributed to coordination during social synchronization, while the cerebellum reduced temporal entrainment to the partner, thereby facilitating a focus on one's own tempo. We document that the brain flexibly uses the same visuomotor information differently depending on synchronization goals, balancing between matching others' timing and preserving one's own during social interactions.

Received May 23, 2025; revised March 19, 2026; accepted March 24, 2026.

Author contributions: S.U. and L.M.S. designed research; S.U. and C.D.B. performed research; S.U., L.M.S., C.D.B., and F.D. analyzed data; S.U., F.D., G.B., and E.P. edited the paper; S.U., L.M.S., and E.P. wrote the paper.

We thank the technical staff of the San Gerardo Monza's Hospital for making this study possible, in particular the radiologist technician Paola Lanzoni, and the Tecnomed Foundation (https://fondazionetecnomed.it/) that funded the MRI facility.

This work was supported by a nationally funded grant from the Italian Ministry of University and Research (MUR) [Progetto finanziato dal MUR ai sensi del D.D. n. 1236 del 1.08.2023 nell'ambito della call "Procedura competitiva per lo sviluppo delle attività di ricerca fondamentale", a valere sul fondo italiano per la scienza 2022 – 2023 (BANDO FIS 2), codice progetto FIS-2023-01659, CUP H53C25000610001, PI: LMS].

The authors declare no competing financial interests.

Correspondence should be addressed to Stefano Uccelli at stefano.uccelli@unimib.it or Eraldo Paulesu at eraldo.paulesu@unimib.it.

This paper contains supplemental material available at: https://doi.org/10.1523/JNEUROSCI.1021-25.2026

https://doi.org/10.1523/JNEUROSCI.1021-25.2026

Introduction

Synchronizing movements is fundamental during social interactions. People may maintain their own tempo regardless of their partner, as when a drummer holds the beat without being pulled off course by a guitarist's solo, or they may synchronize with each other, as two musicians playing an ostinato pattern. These two scenarios distinguish emergent and planned synchronization (Knoblich et al., 2011), where people synchronize with others involuntarily ("entrainment") or voluntarily. The interplay between the two processes and their neurofunctional underpinnings is not yet fully understood.

In two fMRI studies (Fairhurst et al., 2013, 2014), participants performed a finger-tapping task synchronizing with a virtual

partner adapting its frequency to the participants' taps, inducing easier or harder coordination. The virtual partner's adaptivity modulated brain activity both in cortical (e.g., the dorsomedial prefrontal cortex) and subcortical (e.g., the hippocampus and the insula) structures. However, these findings have limited generalizability, as virtual partners follow prespecified algorithms and do not replicate the visuomotor dynamics of real-life interaction. Two fMRI studies explored social co-regulation in dyads performing audio-motor tasks. Silfwerbrand et al. (2022) investigated leader–follower coordination dynamics and found the cerebellum modulating both leading and following, suggesting it may be a key brain structure underlying temporal co-regulation during social synchronization. Kohler et al. (2023) observed cortico-cerebellar functional connectivity modulations for familiar or discrepant partner's actions, suggesting that the cerebellum balances self-generated actions relying on feedback predictions.

However, these findings are confined to audio-motor tasks and cannot be directly extended to visuomotor interactions. Visuomotor interactions entail the issue of discerning planned from emergent processes: whether people synchronize voluntarily or involuntarily, sensory information about their partner's movements is the same. Since observing compatible movements elicits entrainment (Oullier et al., 2008; Hove et al., 2010), a crucial question is whether voluntary co-regulation (planned synchronization) might override entrainment (emergent synchronization). We addressed this issue in a dyadic finger-tapping task (Uccelli et al., 2023). Participants observed their partner's hand while synchronizing with each other or resisting synchronization and focusing on one's own tempo. We manipulated tempi within the dyads in a manner to estimate "tempo contagion," a measure of the timing deviation toward the partner's tapping. We observed similar tempo contagion effects when participants either synchronized or resisted, indicating that entrainment survived voluntary corrections. However, tempo contagion did not correlate within participants between the two conditions, suggesting that perhaps the brain underpins emergent and planned synchronization in different ways.

Aims and predictions

We investigated the neurofunctional underpinnings of emergent and planned synchronization in participants who aimed to achieve or resist social synchronization during fMRI. At the behavioral level, we expected a similar tempo contagion between emergent and planned synchronization, in accordance with previous findings (see above). Our main aim was to determine whether similar behavioral effects could be associated with (at least partially) different patterns of brain activity.

We expected activation (1) in high-level visual areas decoding biological movement (e.g., the lateral occipitotemporal cortex, LOTc; Lingnau and Downing, 2015); (2) in the cerebellum, given its contribution to perceptual and motor timing, and more broadly to social cognition (Van Overwalle et al., 2020); and (3) in motor areas typically involved in synchronization (Repp and Su, 2013). Since emergent synchronization is assumed to rely more on automatic processes, it should recruit brain structures dedicated to internal timekeeping (e.g., the cerebellum). In contrast, since planned synchronization is assumed to rely more on voluntary co-regulation, it could recruit structures involved in voluntary motor control (e.g., the supplementary motor area, SMA). As we did not necessarily expect between-task differences in the overall brain activity, we assessed whether the neural activity of (possibly) co-active brain areas correlated differently with tempo contagion. Last, given previous evidence that the cerebellum establishes connectivity with motor (Kohler et al., 2023) and visual (Tzvi et al., 2022) cortical areas, we conducted a functional connectivity analysis, expecting cortico-cerebellar connectivity in planned (but not emergent) synchronization due to voluntary co-regulation.

Materials and Methods

Ethics

The Milano-Bicocca University Ethics Committee approved the experimental protocol (no. 735). Participants signed written informed consent. The study was conducted in accordance with the standards of the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki), of the Italian Board of Psychologists (https://www.psy.it/codice-deontologico-degli-psicologi-italiani), as well as the Ethical Code for Psychological Research of the Italian Psychological Society (https://aipass.org/node/11560).

Sample size determination

We conducted an a priori power analysis to determine the adequate sample size to replicate the tempo contagion effect. In our previous work (Uccelli et al., 2023), we observed an average tempo contagion of ∼2% (SD ∼2.6). We calculated power assuming a similar average under three hypothetical scenarios where the group SD (i.e., the inter-participant variability) was as follows: (1) equal, (2) 3.5 (+30%), or (3) 3.9 (+50%). As we expected a positive tempo contagion average (see below), we estimated power with a right-tailed t test (i.e., not testing for a negative average). Power was calculated with the *pwr* package in R (Champely et al., 2018) with an increasing n. For $n = 30$, power estimates were 0.99, 0.92, and 0.86 for the three scenarios, respectively. Thus, we deemed $n = 30$ as adequate to replicate our tempo contagion estimates in the scanned participants' sample (see next section).

Participants

Sixty volunteers from the Milano-Bicocca community formed 30 dyads (42 females; mean age, 24.5; range, 18–33) not necessarily matched for gender (f-f, 13; f-m, 16; m-m, 1). They had normal hearing and normal or corrected to normal vision, had no history of neurological diseases, had no professional musical expertise, and were naive as to the purpose of the experiment. Students recruited through a local SONA system (https://milano-bicocca.sona-systems.com) received formative credits. Thirty volunteers entered the fMRI scanner as experimental subjects ("IN"; 17 females), whereas the other 30 stood beside the cot as partners ("OUT"). Three IN participants were excluded for the following reasons: one was left-handed and therefore considered a pilot; one revealed excessive movement artifacts; and one revealed a loss of signal in the right prefrontal area due to technical issues. Accordingly, these three participants were excluded. The final sample involved 27 dyads, composed of 54 right-handed (by self-report) participants.

Stimuli

Stimuli were .mp3 files of a cowbell-like click (10 ms; 4,800 Hz) generated in Audacity (https://www.audacityteam.org). For IN participants, three inter-onset intervals (IOIs) were used as target tempi: 400, 500, and 600 ms (150, 120, and 100 bpm). For OUT participants, three IOIs centered around the target tempi of IN participants were used as distractor tempi: one −20 ms faster and one +20 ms slower ("incongruent" tempi) and one medium tempo equal to the target ("congruent" tempi, which served as baseline). We chose 20 ms as it is about double the 2% average tempo contagion observed in our previous work (see the power analysis above). These three triplets (Fig. 1*a*) allowed us to estimate tempo contagion (see below). Participants were unaware that tempi could be either the same or different between partners. Each tempo sounded six times during the synchronization phase.

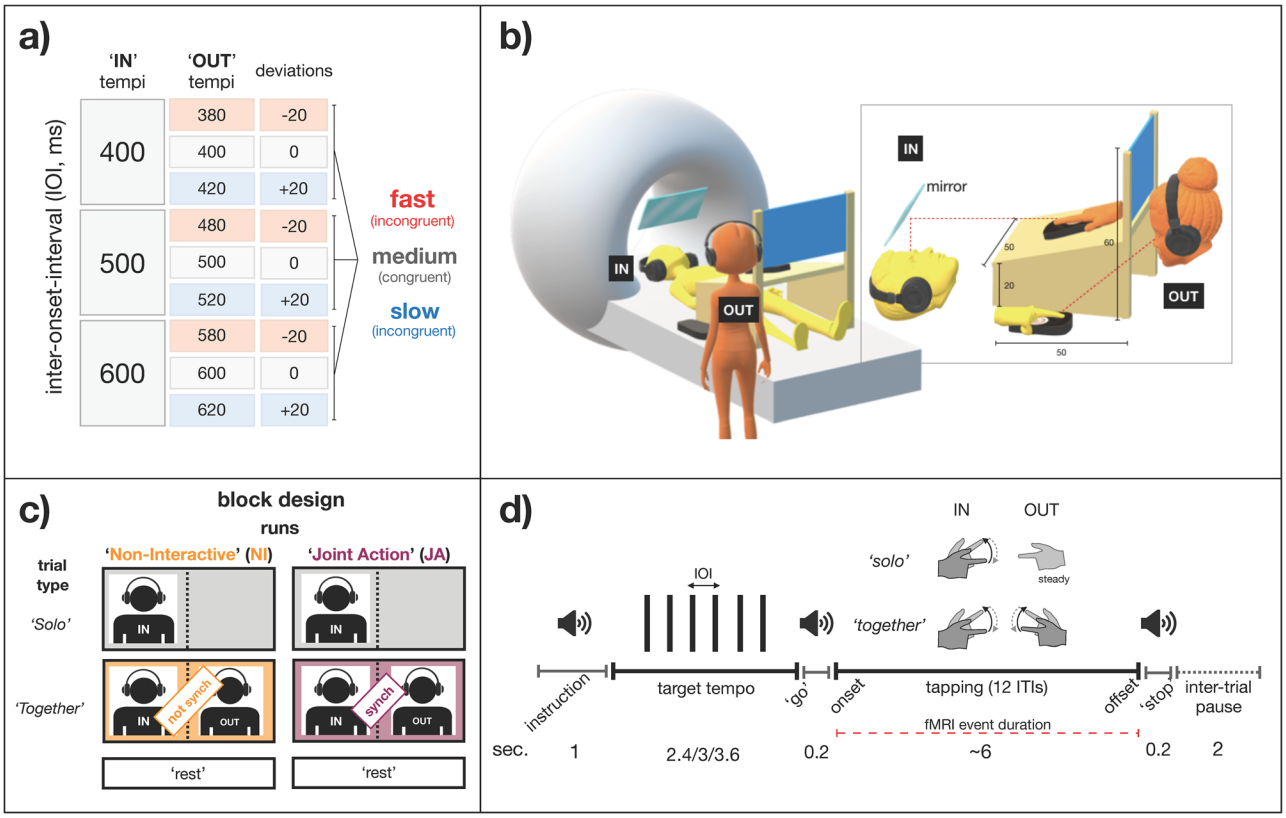

Figure 1. Methods overview. ***a***, Stimuli. "IN" fMRI participants listened to three target tempi (first column). For each, "OUT" fMRI participants listened either to a slower (+20 ms), a faster (−20 ms), or an equal (medium) distractor tempo, yielding three triplets of tempi (middle column). This structure was reduced by re-expressing the data as deviations from the medium–medium pair of tempi (i.e., no effect baseline), isolating the ± 20 ms differences (third column). This re-expression yielded three triples centered around zero, reducing interparticipant variability and improving estimation of tempo contagion. ***b***, Setup. A mirror on the head cage allowed IN participants to observe the OUT participant's hand leaning on a custom-made table above the cot (numbers indicate cm). OUT participants stood beside the MRI cot, observing the IN partner's hand in front of them. Two response boxes collected tapping data. ***c***, Experimental design. The two runs ("Non-Interactive," NI, and "Joint Action," JA) consisted of "Solo" (only IN participants tapping), "Together" (both participants tapping), and rest trials. Runs were identical except for Together trials in which the instruction was not to sync (NI) or sync (JA) with the partner's tapping. ***d***, Trial timeline. Participants listened to the instructions relative to the trial type, and then the appropriate tempo sounded six beats. Once ceased, participants tapped maintaining the tempo according to the instruction until a stop sound signaled the trial's end. This time window corresponded to the fMRI event duration (red dashed line; see Materials and Methods). See text for details.

Setup and apparatus

Figure 1*b* depicts the experimental setup. OUT participants stood beside the MRI cot, seeing the IN participants' right hand. An MRI-compatible table on the cot held the OUT participants' right hand. IN participants were able to see the OUT participants' hand through a mirror placed on the head cage. A panel fixed on the back of the table prevented OUT participants from seeing their hand, as well as focusing the IN participant's vision on the OUT participant's hand. Two MRI-compatible response boxes (Resonance Technology) were used to collect tapping data. Both participants wore soundproofing headphones to listen to tempi and communicate with the control room.

Experimental conditions and procedure

Participants performed an unpaced finger-tapping task while keeping their gaze on the partner's hand. The general instruction was to "synchronize to the target tempo, and once ceased, maintain it at best by tapping your right index finger until the trial ends." Participants completed two runs: a "Non-Interactive" (NI) and a "Joint Action" (JA) run. Each run included the following three types of trials (Fig. 1*c*):

1. "Solo": IN participants listened to the target tempo and replicated it individually while observing the hand of the OUT partner, who was instructed not to take any action and to wait for Together trials. Solo trials were identical in NI and JA runs and served solely as a reference to estimate the neural activity characterizing interpersonal synchronization (i.e., "Together" trials) while controlling for the one necessary for individual synchronization to the target tempo (see MRI data acquisition and analyses below). Solo trials for OUT participants were unnecessary and therefore not included.
2. "Together": Both participants listened to their appropriate tempo and replicated it according to the instructions specific to the run. In the NI run, both participants concurrently tapped while seeing the partner's hand (full vision). The instruction was to "maintain the tempo as best as possible while ignoring the partner's tapping." Importantly, they nevertheless had to keep their eyes open and look at their partner's hand, but trying not to fall in sync with the partner's tapping, although they concurrently tapped having full access to the visual information of the partner's hand movement. In the JA run, both participants (full vision) concurrently tapped while seeing the partner's hand but, in contrast to the NI condition, the instruction was "to maintain the tempo jointly as best as possible while synchronizing with the partner's tapping."
3. "Rest": No tempo sounded. Participants were instructed to relax with their eyes open, waiting for the next trial.

Note that the NI and JA Together trials provided identical visuomotor information and differed only in the type of synchronization instruction. NI and JA runs were counterbalanced across participants.

Dyads were trained by completing 10 practice trials with randomly chosen tempi for each condition on a PC prior to the fMRI experiment. An additional five practice trials were run in the fMRI room before initiating the experiment. Figure 1*d* depicts the trial timeline. At the beginning of each trial, a verbal instruction informed the dyad of the condition ("Solo" or "Together"; "Solo" or "Insieme" in Italian). Next, tempi

sounded six times and dyads synchronized to them ("synchronization" phase) and then the "go" signal sounded. At this point, according to the condition, participants tapped (unpaced) maintaining the tempo until the "stop" signal sounded signaling the trial's end ("tapping" phase). The trial ended once the MATLAB script collected 13 taps (i.e., 12 intervals) from both participants; if the tapping phase exceeded a timeout of 12 s, the trial ended and was considered an error. Rest trials lasted 12 s to match the maximum trial time (i.e., 12 s). Each run consisted of 63 randomized trials, including 18 Solo, 18 Together congruent (i.e., same tempo), 18 Together incongruent (9 fast and 9 slow distractors), and 9 rests. Each run lasted ∼13 min, separated by a brief pause, plus ∼8 min for acquiring the structural (T1) image. Scanning lasted ∼40 min in total.

Behavioral data analysis

We analyzed the intertap interval (ITI), that is, the difference in time between two consecutive taps ($tap_n - tap_{n-1}$), for both IN and OUT participants. From ITIs, we estimated tempo contagion effects (see below) to quantify the magnitude of temporal entrainment. All behavioral analyses were performed in R 4.2.2 (R Core Team, 2024).

Data validation. First, we visually inspected the individual ITI raw distributions as a function of IOIs for detecting outliers. ITIs <150 and >1,000 ms were removed as extreme values. Second, we applied a filter to detect ITIs ±250 ms of the actual IOI in line with our previous work (Uccelli et al., 2023). Following this criterion, 7.3% ITIs were excluded in Solo trials, while 7 and 8.3% ITIs for IN and OUT participants, respectively, were excluded in Together trials. Third, we checked normality assumptions in the ITI distributions to apply a transformation in case of marked asymmetry. The distribution was right-skewed and platykurtic for participants in both roles (IN: skewness = 0.33, kurtosis = 3.9; OUT: skewness = 0.44, kurtosis = 3.86). A *boxcox* procedure returned a λ value of ∼0.5 (i.e., a sq. rt. transformation; Osborne, 2010) for both distributions, indicating that the square root transformation could improve the distributions (Osborne, 2010). The transformation made distributions more symmetrical (IN: −0.07; 4.1; OUT: 0.06; 3.8). Note that we computed individual means on the transformed data, and then we raised them to the power of two before the analysis for the sake of interpretation.

Estimation of tempo contagion. We expected participants to drift ITIs toward each other in incongruent trials. For instance, IN participants should tap faster (or slower) when OUT participants tap faster (or slower), compared with trials in which the partner heard congruent tempi (medium; Fig. 1*a*). Interindividual differences (e.g., in how much the pair speeded up in dyadic tasks, in line with the so-called "rushing" effect; Wolf and Knoblich, 2022) may mask subtle temporal adjustments, making it difficult to estimate tempo contagion. To bypass this issue, we performed a centering variable procedure (Enders and Tofighi, 2007) to exclude interindividual differences in the baseline conditions (i.e., congruent tempi), which are not relevant for analyzing entrainment effects. For a given triplet of tempi (e.g., 480, 500, and 520 ms), we re-expressed individual data as deviations by subtracting the congruent average (i.e., 500) from each ITI datapoint, obtaining positive and negative values relative to zero, signaling slower or faster tap, respectively. For the sake of clarity, given a 500 ms tempo for the IN participant and a 480 ms tempo for the OUT participant, we expected the former to show a negative (faster) deviation and the latter a positive (slower) deviation, compared with when they listened to both the 500 ms baseline. After this re-expression, the three triplets of means consisted of a faster (−20 ms) and a slower (+20 ms) means around a common zero (Fig. 1*a*, third column). This re-expression yielded comparable estimates across participants, conditions, and targets, as it isolates the expected temporal entrainment cleaning data from baseline interindividual differences. This allowed us to estimate the magnitude of tempo contagion effects in each participant and to correlate it, for example, with neurofunctional data. Thus, in the Results section, statements regarding interindividual variability refer to interindividual variability in the magnitude of tempo contagion effects.

Last, we estimated a unique tempo contagion effect by adapting a formula employed in our previous studies (Uccelli et al., 2023; see also Uccelli and Bruno, 2024). We computed the averages of slow and fast ITI deviations, then the difference between such averages, and then expressed the difference as a percentage to the appropriate pair of medium tempi (baseline), as summarized by the following formula:

$$\%\text{tempo contagion} = \frac{(\text{mean(slow)} - \text{mean(fast)})}{\text{medium}} \times 100,$$

where the numerator is the difference between slow (+20 ms) and fast (−20 ms) average means and the denominator is the appropriate medium (baseline) mean. Positive % values indicate tempo contagion effects. The formula was applied separately to the data collected during the NI and the JA runs.

Analysis of tempo contagion. Percent tempo contagion effects entered a 2 × 2 mixed ANOVA, with "Condition" (NI/JA) as the within-subjects factor term and "Participants" (IN/OUT) as the between-subjects factor term. If the two factors did not modulate tempo contagion, the model would return no significant main or interaction effect. Testing was run using the *ezANOVA* function of the R *ez* package (Lawrence and Lawrence, 2016).

MRI data acquisition and analyses

Data acquisition. Whole-brain functional images were acquired using a 3.0 T Ingenia CX scanner (Philips). Gradient-echo T2*-weighted transverse echoplanar images (EPIs) were acquired (scan parameters: TR, 2,000ms; TE, 30 ms; 35 transversal slices; descending not interleaved acquisition; 4 mm slice thickness; with no interslice gap; FA, 75°; FOV, 240 mm; matrix size, 80 × 80; voxel size, 3 × 3 × 4 mm). The first five volumes recorded from each functional run were removed to allow for steady-state tissue magnetization. Each run was planned to acquire a maximum of 460 valid scans. The expected number of scans per run was ≈440, but we extended it to a maximum of 460 to ensure that the run did not end before participants completed all trials. If dyads completed the task beforehand, the run was stopped, and the last scan was discarded before the preprocessing steps to avoid potential distortions due to the forced interruption. On average, 437 scans (± 21.7 scans) were acquired per run. Each run lasted ∼14 min and 30 s.

A high-resolution T1-weighted anatomical scan [TR, 8.1 ms; TE, 3.7 ms; 207 sagittal slices; FOV, 256 mm; matrix size, 256 × 256; FA, 8°; Inversion Time (TI), 1,000 ms] was also acquired before the fMRI session.

Task functional activation analysis

Images were converted from DICOM to NifTI format in mriconvert (https://neuro.debian.net/pkgs/mriconvert.html). Subsequent analyses were performed in MATLAB (R2019b, MathWorks) using Statistical Parametric Mapping (SPM12, Wellcome Department of Imaging Neuroscience).

Preprocessing. First, scans were realigned and unwrapped to account for any movement during the experiment and to reduce magnetic field distortion effects. Second, unwrapped images were coregistered with a T1-weighted structural image of each participant, which was then segmented and stereotactically normalized into the SPM12 template (tmp.nii) to allow for group analyses. Third, deformation fields used for T1 segmentation were applied to coregistered functional scans; here, the data matrix was interpolated to produce voxels 2 × 2 × 2 mm in dimension. Last, the normalized images were smoothed using a Gaussian filter of 10 × 10 × 10 mm to improve the signal-to-noise ratio and allow for valid application of the familywise error rate (FWER) correction at the cluster level in the second-level (i.e., group-level) univariate analyses, as recommended by Flandin and Friston (2019). We used Artifact Detection Tools (ART, Whitfield-Gabrieli, https://www.nitrc.org/projects/artifact_detect) to identify outlier scans for each participant. Timepoints were marked as outliers when scan-to-scan variations in the global signal exceeded 3 SD from the mean and when the compounded measure of movement parameters exceeded 1 mm scan-to-scan

movement. On average, excluded volumes were 2.14%. Only one participant, who showed 135 outliers (31.1%), was excluded from the sample (see above, Participants; the final sample included 27 dyads).

Univariate voxel-by-voxel analyses. We performed a two-step statistical analysis based on the general linear model. The blood oxygen level-dependent (BOLD) signal associated with each experimental condition was analyzed by convolution with a canonical hemodynamic response function (Worsley and Friston, 1995). Global normalization was applied. The time series was high-pass filtered at 128 s and prewhitened by means of an autocorrelation model AR(1).

In a block design, the event onset corresponded to the tapping onset, and the event duration corresponded to the difference between the go signal and the moment at which 13 taps were recorded from both participants (Fig. 1*d*, red dashed line). Separately for the JA and NI runs, these events entered the analyses to estimate the BOLD signal associated with the following conditions of interest: Solo and Together. Rest trials were not modeled, as recommended by Pernet (2014). Regressors of non-interest modeled experimental confounds, including the signal associated with (1) the onset of targets; (2) the onset of the instruction, go, and stop signals; (3) the realignment parameters calculated in the preprocessing; and (4) errors, i.e., trials either exceeding the timeout or having an anomalous duration (<1%).

At the first-level (individual) analysis, we computed fixed condition-specific effects for each participant. We modeled the simple effect of each condition of interest (Solo and Together, separately for the NI and JA runs). Such contrasts allowed us to determine both individual- and dyadic-specific activations and their difference.

At the second-level (group) analysis, we computed a 2×2 full-factorial design including Task (Solo vs Together) and Condition (NI vs JA) as within-subject factors. As a preliminary analysis, we checked for the presence of any significant difference between the Solos in the JA and NI runs (to ensure they did not differ) and explored the simple effect of Solo that served as a reference point. Then, we explored the main effect of Condition (Solo > Together, Together > Solo; see Supplementary Materials) and the interaction effect (NI(Together > Solo) > JA(Together > Solo), JA(Together > Solo) > NI(Together > Solo)). We also conducted a conjunction analysis to describe the brain areas activated during both NI and JA runs (NI(Together > Solo) ∩ JA(Together > Solo)). All contrasts were masked with the main effect of the task (contrast: 1 1 1 1, $p < 0.05_{\text{uncorr}}$) to only include significant task-related activations.

The reported results apply a FWER correction for multiple comparisons at the cluster level ($p < 0.05$). The cluster-wise correction was applied to data having applied a $10 \times 10 \times 10$ Gaussian smoothing and at $p < 0.001_{\text{uncorr}}$ at the voxel level, as recommended by Flandin and Friston (2019).

Linear regression between % tempo contagion and BOLD signal. We planned to extract the effect of interest, i.e., JA or NI task-related activation, NI(Together > Solo) and JA(Together > Solo), to test their association with the % tempo contagion effect. The average *y*-values for each cluster were extracted by using the *MarsBaR* toolbox (Brett et al., 2002). We added the values extracted from each cluster as continuous predictors in a series of linear models fitting % tempo contagion effects and having "Condition" (NI/JA) as an additional factor. We expected task-specific activations extracted from different clusters to modulate tempo contagion effects differently in the JA and NI trials, as possibly shown by a significant interaction effect. To evaluate such potential interactions, we computed 95% CIs around the slopes, judging whether they covered zero.

Task functional connectivity analysis

Functional connectivity analyses were performed in MATLAB using the software CONN 22.a (Whitfield-Gabrieli and Nieto-Castanon, 2012) and SPM12.

Preprocessing and denoising. Scans were preprocessed according to the following steps: realignment, slice timing correction, outlier detection (ART, see above), indirect segmentation and MNI space normalization, and smoothing ($10 \times 10 \times 10$ mm). Scans were denoised using a standard pipeline including the following regressors of potential confounding effects: white matter and CSF timeseries, motion parameters and their first-order derivatives, task activation effects, outlier scans, and linear trends within each run. A high-pass 0.008 Hz frequency filter was applied to the BOLD time series. No low-pass filter was applied. For additional information, see the Supplementary Materials online (Text S1).

First-level analysis. Seed-based functional connectivity maps were estimated, characterizing the spatial pattern of the whole-brain functional connectivity with the seed areas, as defined by the brain clusters that correlated with % tempo contagion estimates in the task activation analyses. Functional connectivity strength was represented by Fisher-transformed bivariate correlation coefficients estimated separately for each seed area and target voxel, modeling the association between their BOLD signal timeseries.

Second-level analysis. Linear regressions were performed to test whether the % tempo contagion was associated with the whole-brain functional connectivity of the selected seeds. In one regression model, we used the NI % tempo contagion to predict the whole-brain functional connectivity strength of the seed areas during the NI condition. In a separate regression model, the JA % tempo contagion was used to predict the functional connectivity strength of the seed areas during the JA condition.

Voxel-level hypotheses were evaluated using multivariate parametric statistics with random effects across subjects and sample covariance estimation across multiple measurements. Results were thresholded at $p < 0.001_{\text{uncorr}}$ at the voxel level and $p < 0.05_{\text{FWE}}$ at the cluster level.

Results

Behavioral results

Solo trials

We checked whether ITIs of IN participants differed between the NI and JA runs. Group averages with standard error of the mean (SEM) were almost equal across the three IOIs (NI vs JA: 392 ± 3 and 391 ± 4 ms; 470 ± 4 and 470 ± 4 ms; 550 ± 5 and 545 ± 5 ms). Although ITIs scaled along the IOIs, on average, participants anticipated the actual tempo, consistent with rushing effects. Visual inspection confirmed that the group averages did not differ between conditions (see Fig. S1 in the Supplementary Materials).

Together trials

Figure 2*a* presents individual and group ITI means as a function of tempi in both runs (left, NI; right, JA). Each point can be compared with zero vertically or horizontally for IN and OUT participants, respectively. Dyads underestimated the 500 and 600 ms tempi (see distances between SEMs from orthogonal crosses), consistent with rushing effects. Given that rushing may potentially mask fine co-adjustments in time, we computed deviations from the baseline average to isolate tempo contagion (see Materials and Methods for details).

Figure 2*b* presents the individual and group ITI deviation means. For instance, for a target and a distractor of 500 and 520 ms, respectively, participants should drift tapping toward the middle (~510). Indeed, most of the fast tempi (red) led to faster ITIs (quadrant III) and most of the slow tempi (blue) to slower ITIs (quadrant I): group averages fall within the two colored regions in which tapping was expected to balance between partners.

Figure 2*c* presents individual and group % tempo contagion effects (see Materials and Methods for the formula description). Group averages were similar across participants and conditions. The mixed ANOVA (see Materials and Methods for details) revealed no main effect of Role (IN/OUT; $F_{(1,52)} = 3.73$, $p =$

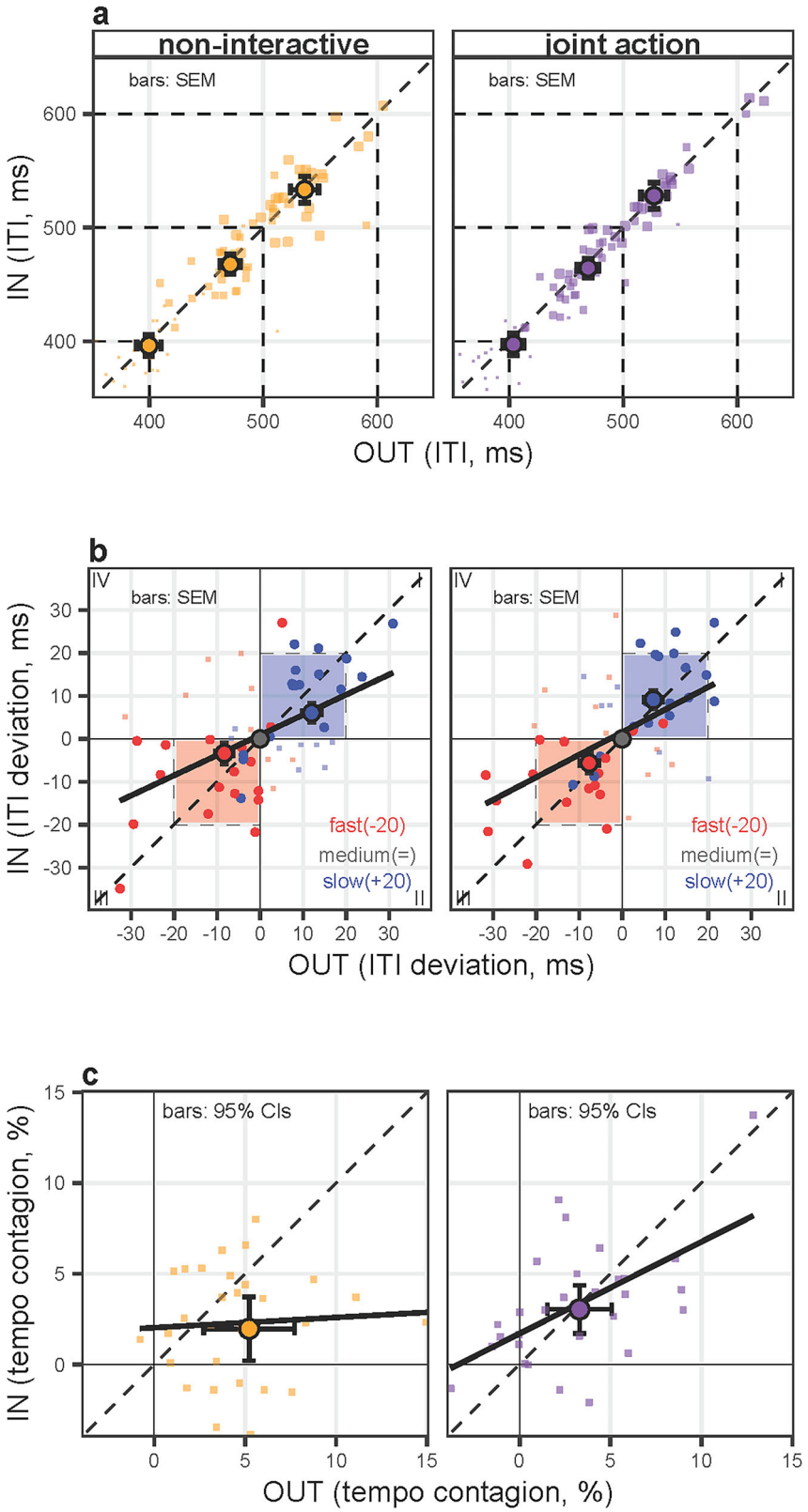

Figure 2. Behavioral results. Individual (small points) and group (large points) data of "IN" fMRI participants (*y*-axes) are plotted against "OUT" participants (*x*-axes) in the two Together "Non-Interactive" and "Joint Action" conditions (columns). ***a***, Intertap interval (ITI). ITIs (in absolute values) scaled on the actual tempi, but dyads rushed at 500 and 600 ms (group averages do not cover the orthogonal crosses). ***b***, Estimation of tempo contagion. ITIs re-expressed as deviations from the medium–medium mean (baseline, gray points). Group averages on the *y*-axis indicate that IN participants tapped faster or slower when the OUT partner tapped for a faster (red points, −20 ms) or slower (blue points, +20 ms) distractor tempo. Conversely, group averages on the *x*-axis do not lie on the black segments (i.e., the actual distractor tempo) but are closer to the baseline average, indicating that OUT participants drifted tapping as well. Colored areas, the ranges in which ITIs are expected if both participants drift toward each other. Thick black line, linear regression obtained by averaging parameters from individual fits. ***c***, Percent tempo contagion. Both IN and OUT participants showed similar % tempo contagion effects across conditions (group averages' CIs do not cross zero lines). See Materials and Methods for the formula description. Thick black lines, linear regression fitting % tempo contagion effects of IN and OUT participants.

0.06, $\eta^2 = 0.036$) and Condition (NI/JA; $F_{(1,52)} = 0.001$, $p = 0.98$, $\eta^2 = 0.000006$), nor the Role × Condition interaction effect ($F_{(1,52)} = 2.66$, $p = 0.11$, $\eta^2 = 0.02$), but a significant effect of the intercept ($F_{(1,52)} = 8.3$, $p < 0.001$, $\eta^2 = 0.45$). Thus, % effects were different from zero (see CIs in Fig. 2*c* not crossing zero lines) but not modulated by experimental manipulations. Moreover, % effects correlated between IN and OUT participants in the JA run [$r_{(25)} = 0.59$, CI (0.27, 0.79); $p = 0.001$] but not in the NI run [$r_{(25)} = 0.21$, CI (−0.18, 0.54); $p = 0.29$]. Last, % effects did not correlate between NI and JA conditions either for IN [$r_{(25)} = -0.0003$, CI (−0.37, 0.37); $p = 0.29$] or OUT [$r_{(25)} = 0.08$, CI (−0.31, 0.45); $p = 0.29$] participants.

Last, as suggested by an anonymous reviewer, we checked dyadic synchronization to describe further NI and JA trials. We performed the same analysis reported in our previous work (Uccelli et al., 2023; Fig. 5). The results are superimposable on our previous estimates, indicating that dyadic synchronization was higher in JA trials than in NI trials. However, for the sake of conciseness, we uploaded the summary of the analysis to the Supplementary Materials (Fig. S2).

fMRI results

Solo network

First, we performed a *t* test to check whether Solo trials showed different activations between NI and JA runs. The two "Solo NI > Solo JA" and "Solo JA > Solo NI" linear contrasts revealed no significant difference. Next, we calculated the simple effect of "Solo" to investigate brain regions specifically active during individual tapping (see also the gray cluster reported in Fig. 3*a*, which represents the simple effect of Solo in the ANOVA, see below). Such brain regions included the bilateral precentral gyri, left postcentral gyrus, left SMA, left insula, left pallidum and putamen, and lobules VI and VIII of the right cerebellum (for details, see Table S1 in the Supplementary Materials).

Together network—"Together > Solo"

Figure 3*a* presents the results of the univariate ANOVA task (Solo vs Together) × Condition (Non-Interactive, "NI," vs Joint Action, "JA"). The results showed a main effect of the task: BOLD signal in the Together tasks increased in the same clusters observed in Solo (see orange and purple bars compared with gray ones in Fig. 3*b*, which plots the simple effects of each task compared with the implicit baseline). These brain regions include the left and right precentral gyri, the left SMA, and the right cerebellum (Fig. 3*bi–iii,vi*). Moreover, both the NI and JA Together tasks activated two large bilateral clusters in the LOTc, which were silent during Solos (Fig. 3*biv,v*; see also Table S2 in the Supplementary Materials for detailed information). The ANOVA also revealed no significant difference between the "NI (Together > Solo)" and "JA (Together > Solo)" linear contrasts (i.e., the interactions; Fig. 3*b*, orange and purple bars). The "Solo > Together" linear contrast did not reveal significant activations. Last, we ran a 2 × 2 × 2 mixed-factors ANOVA including Task (Solo vs Together) and Condition (NI vs JA) as within-subject factors and the run order (first JA vs first NI) as between-subject factors to rule out potential order effects; the results revealed no significant main or interaction effect including the between-subject factor.

Conjunction analysis—"NI(Together > Solo) ∩ JA(Together > Solo)"

Given extensive overlapping activations in the social conditions, we proceeded with a conjunction (inclusive) analysis [NI(Together > Solo) ∩ JA(Together > Solo)] to restrict the regression analysis with % tempo contagion (see next paragraph) to the brain areas co-active in both tasks. The following six clusters were significantly activated in both conditions (Fig. 4*a*): a cluster in the right cerebellum (lobule VI), two clusters in the LOTc bilaterally, a right inferior frontal cluster, a cluster in the right inferior parietal, and a large left fronto-parietal cluster. Table 1 reports the MNI coordinates of such clusters and the corresponding significance ($p < 0.001_{\text{uncorr}}$ at the voxel level and $p < 0.05$ $\text{FWER}_{\text{corr.}}$ at the cluster level).

Linear regressions between conjunction clusters and % tempo contagion

We tested the relation between % tempo contagion effects and the effect of interest (JA/NI > Solo) extracted from each cluster of the conjunction analysis (Fig. 4*a*). Since clusters were rather extended, the local maxima might not be representative of the cluster activity. Thus, we extracted the average BOLD *y*-values by using the *MarsBaR* toolbox. We computed a series of linear models fitting the % tempo contagion effects as a function of the condition and adding the values extracted with *MarsBaR* as a continuous predictor. Since % effects did not differ between conditions (Fig. 2*c*), we were interested in evaluating interaction terms, which would indicate that the neurofunctional activations differently modulated % effects in the two tasks. For the cerebellum cluster, we observed a negative slope in the NI condition ($t_{(50)} = -2.7$, $p = 0.03$) and a flat slope in the JA condition ($t_{(50)} = 0.6$, $p = 0.5$). For the right LOTc, we observed a positive slope in the JA condition ($t_{(50)} = 3.9$, $p = 0.005$) but a flat slope in the NI condition ($t_{(50)} = -1.36$, $p = 0.2$), and the same pattern was observed for the left LOTc (JA condition: $t_{(50)} = 3.7$, $p = 0.02$, NI condition: $t_{(50)} = -1.07$, $p = 0.4$). Figure 4*b* presents these regressions and the corresponding CIs around the slopes (inset plots). For the remaining clusters (i.e., the right parietal cluster and right and left frontal clusters, not plotted in Fig. 4*b*), we observed flat functions relating BOLD to % tempo contagion (see Fig. S3 in the Supplementary Materials).

Task functional connectivity analysis

We sought to determine whether % tempo contagion effects modulated functional connectivity between conjunction clusters and the rest of the brain. Since contagion effects revealed a negative correlation in the right cerebellum in the NI condition and a positive correlation in the bilateral LOTc in the JA condition, we selected these clusters as seeds to restrict the analysis to the behavioral results observed on tempo contagion.

During the NI condition, the functional connectivity between the right cerebellum and a portion of the right LOTc (medial occipital gyrus) was negatively associated with NI % tempo contagion. Similarly, the functional connectivity between the right cerebellum and the right postcentral gyrus was negatively associated with NI % tempo contagion (Fig. 5*a*; MNI coordinates in Table 2). Thus, the higher the NI % tempo contagion, the lower (more negative) the connectivity between the cerebellum and these two clusters (Fig. 5*b*). No significant association emerged between NI % tempo contagion and the functional connectivity patterns of the right or left LOTc. During the JA condition, however, no significant association was observed between JA % tempo contagion and the functional connectivity patterns of the right cerebellum or the right and left LOTc.

During the review process, an anonymous reviewer suggested performing a generalized psychophysiological interaction (gPPI) analysis. gPPI analysis investigates potential task-related changes in functional connectivity but note that our main aim was to

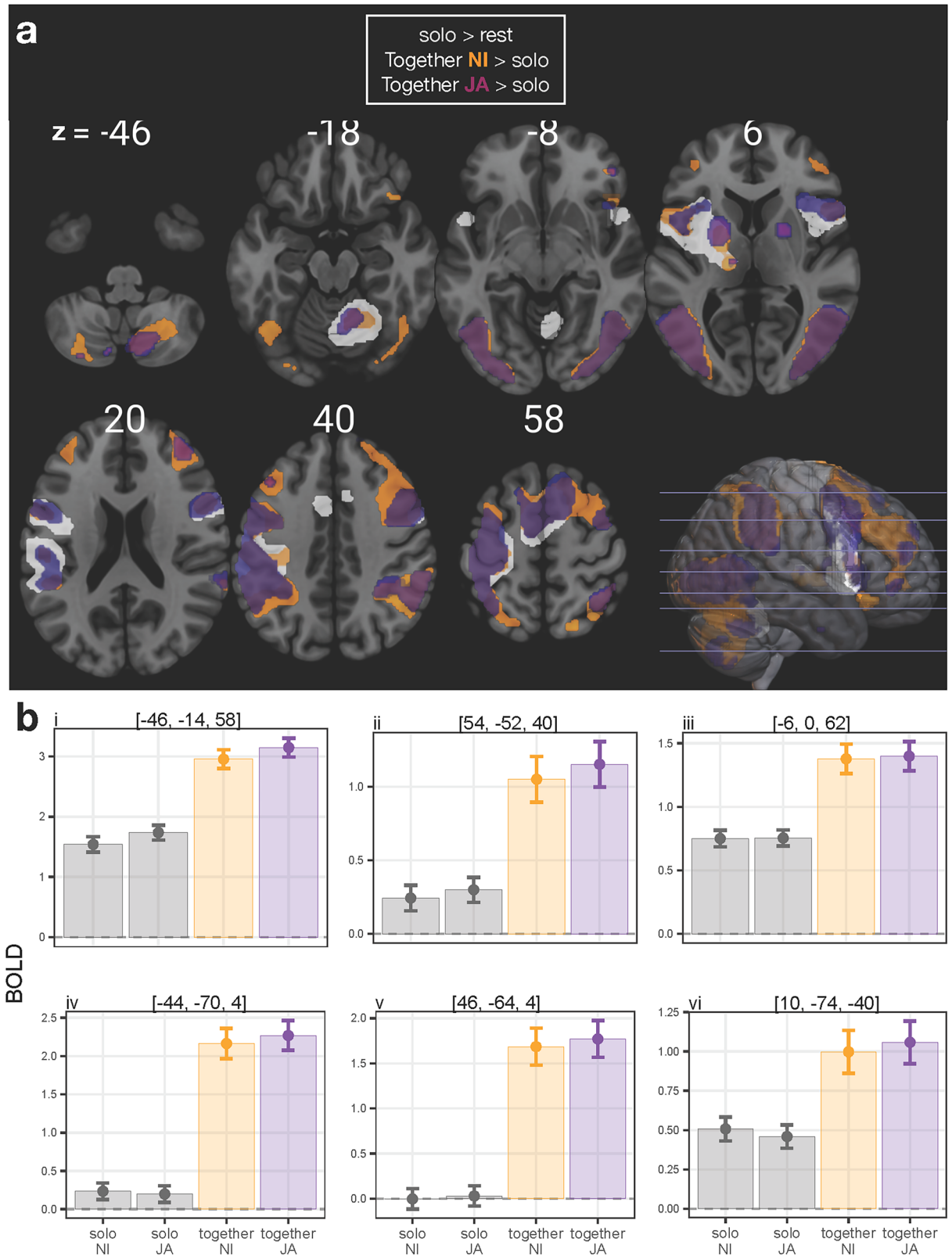

Figure 3. Univariate fMRI results. ***a***, Task (Solo vs Together) × Condition (Non-Interactive, "NI," vs Joint Action, "JA") ANOVA. Gray clusters represent the simple effect of Solo in both runs. Brain regions active during Together trials more than Solo trials are depicted separately for the NI (orange) and JA (purple) runs. Data is plotted by applying a $p_{uncorr} < 0.001$ threshold at the voxel level and a *p*FWER-corr. < 0.05 threshold at the cluster level (see Table S2 in the Supplementary Materials for details on MNI coordinates). ***b***, Histograms plotting the effects of interest in the ANOVA. Local maxima: left (***i***) and right (***ii***) precentral gyrus (PrG), (***iii***) left supplementary motor area (SMA), left (***iv***) and right (***v***) lateral occipitotemporal cortex (LOTc), and right cerebellum (***vi***). Numbers indicate MNI coordinates. Barplots show beta values with 95% CIs. Note that the overlapping of orange and purple CIs indicate no interaction effect (i.e., equally active voxels in NI(Together > Solo) and JA(Together > Solo) linear contrasts).

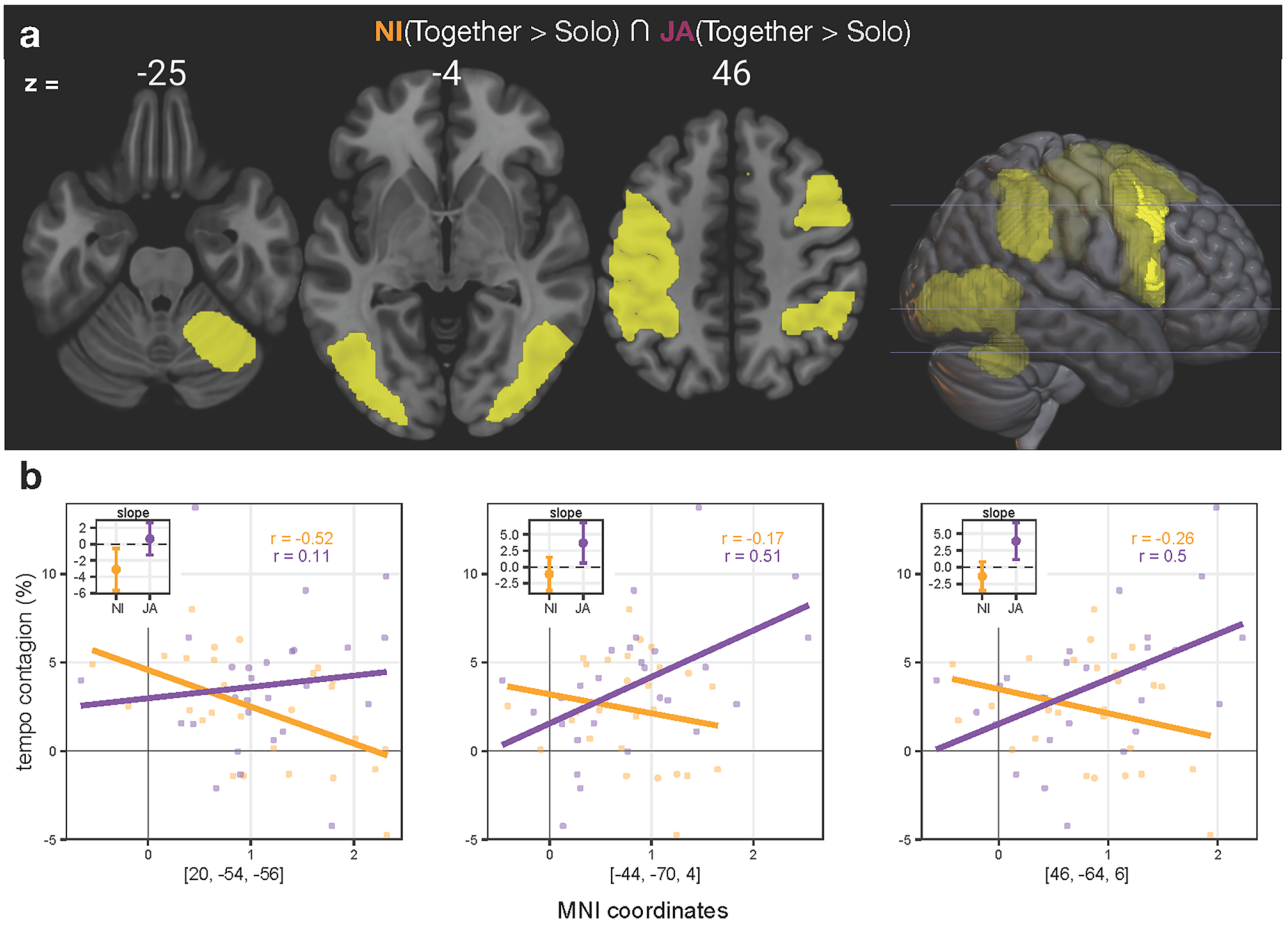

Figure 4. Results of the Conjunction analysis. ***a***, clusters from the NI (Together > Solo) ∩ JA (Together > Solo) conjunction. Data is plotted by applying a $p_{uncorr} < 0.001$ threshold at the voxel level and a *p*FWER-corr. < 0.05 threshold at the cluster level. ***b***, Linear regression between the individual % tempo contagion (*y*-axis) and the average BOLD signal (*x*-axis) of the cluster computed with the *MarsBar* toolbox (left plot, right cerebellum; middle and right plots, the bilateral LOTc). Each point represents a different IN participant. Inset plots, 95% confidence intervals (CIs) around slopes extracted from the linear model (see text for details). CIs not covering zero (dashed line) indicate that the BOLD signal modulated % effects. Orange, Non-Interactive condition (NI); purple, Joint Action condition (JA). Note that the regressions with the right parietal cluster and right and left frontal clusters are not shown since they did not reveal differences in slope, but plots are presented in Figure S3 in the Supplementary Materials.

evaluate which brain structures contributed to transmitting tempo contagion between NI and JA tasks. Nonetheless, we performed the analysis and uploaded it to the Supplementary Materials online (Text S2, Fig. S4) for the interested readers.

Discussion

Real-life social interactions rely on exchanging visuomotor information, which is fundamental for achieving social synchronization. This is the first work to investigate the neurofunctional correlates of emergent and planned synchronization during visuomotor coordination. The results reveal a common neural network for both synchronization types, including the lateral occipitotemporal cortices (LOTc), posterior parietal cortices, the precentral gyri, the supplementary motor areas bilaterally, and the right cerebellum. However, the analyses showed that tempo contagion depended on the recruitment of different neural resources in the two social conditions, suggesting that emergent and planned synchronization depend on how visual and cerebellar brain regions process visuomotor information.

Participants deviated their tapping toward the partner's timing (Fig. 2*b*), showing a similar tempo contagion for the NI and JA conditions (Fig. 2*c*). Thus, entrainment occurred regardless of the intended goal, in line with our previous findings (Uccelli et al., 2023). However, the magnitude of tempo contagion correlated among partners in the JA but not in the NI condition (Fig. 2*c*), and it did not correlate between the NI and JA conditions, suggesting that tempo contagion might stem from different mechanisms. In the NI task, the partner's movement represents a source of interference to which participants are entrained, whereas in the JA task, it is fundamental to sync with the partner's timing. As the results confirmed, depending on the social context and the ensuing instructions, different brain structures process visuomotor information in different ways.

At the neurofunctional level, Solo trials (Fig. 3*a*) activated brain areas consistent with an ALE meta-analysis on finger-tapping tasks (Witt et al., 2008). Together trials recruited the same network as Solo trials, indicating that the brain areas underlying social synchronization are approximately the same as those of individual synchronization, except for the activation of two lateral occipitotemporal cortices (LOTc). LOTc is active during the observation of moving stimuli, including others' actions (Lingnau and Downing, 2015; Zhang et al., 2022), information that was absent during Solo trials (as the OUT participant's hand was stationary). The LOTc is involved in the neural encoding of several socially relevant visual features for dyadic interactions, such as body and hand motion, person orientation, and action representation (Wurm and Caramazza, 2019; McMahon

Table 1. Conjunction analysis "NI (Together > Solo) ∩ JA (Together > Solo)"

	Left hemisphere				Right hemisphere			
Brain area (Brodmann)	*X*	*Y*	*Z*	*Z*-score	*X*	*Y*	*Z*	*Z*-score
Left occipital cluster $k = 1{,}408$, $p_{FWE\text{-}corr.} = 0.001$								
Medial occipital gyrus (37)	−44	−70	4	>8*	–	–	–	–
Inferior occipital gyrus (18)	−28	−94	−8	6.77*	–	–	–	–
Right tempo-occipital cluster $k = 1{,}500$, $p_{FWE\text{-}corr.} = 0.001$								
Medial occipital gyrus (37)	–	–	–	–	46	−64	6	>8*
Medial occipital gyrus (19)	–	–	–	–	38	−88	2	6.74*
Right parietal cluster $k = 1{,}187$, $p_{FWE\text{-}corr.} = 0.002$								
Inferior parietal gyrus (40)	–	–	–	–	56	−52	40	5.12*
	–	–	–	–	54	−52	44	5.03*
	–	–	–	–	42	−48	52	4.58*
	–	–	–	–	46	−46	52	4.57*
	–	–	–	–	32	−50	48	4.56*
Angular gyrus (40)	–	–	–	–	62	−52	28	4.41*
Supramarginal gyrus	–	–	–	–	60	−38	28	4.13
	–	–	–	–	66	−40	26	4.04
Fronto-parietal cluster $k = 8{,}423$, $p_{FWE\text{-}corr.} < 0.001$								
Postcentral gyrus	−46	−14	58	>8*	–	–	–	–
	−38	−20	54	>8*	–	–	–	–
Precentral gyrus (6)	−46	0	52	6.94*	–	–	–	–
Inferior parietal gyrus (40)	−52	−48	50	6.42*	–	–	–	–
	−34	−48	50	5.86*	–	–	–	–
Supplemental motor area (6)	−6	0	62	6.24*	10	6	62	5.67*
	–	–	–	–	12	18	56	5.49*
Supramarginal gyrus (48)	−48	−44	30	5.33*	–	–	–	–
	−52	−24	24	4.62*	–	–	–	–
Superior frontal gyrus (8)	−14	16	58	3.49	–	–	–	–
Superior frontal gyrus (6)	−20	10	66	3.42	–	–	–	–
	−24	8	68	3.33	–	–	–	–
Inferior frontal gyrus, pars opercularis (48)	−48	10	4	4.56*	–	–	–	–
Inferior frontal gyrus, pars opercularis (6)	−58	8	18	4.19	–	–	–	–
	−54	8	12	4.12	–	–	–	–
Right frontal cluster $k = 1{,}843$, $p_{FWE\text{-}corr.} < 0.001$								
Precentral gyrus (6)	–	–	–	–	48	2	50	6.04*
Inferior frontal gyrus, pars opercularis (6)	–	–	–	–	56	12	12	4.48*
Inferior frontal gyrus, pars opercularis (44)	–	–	–	–	52	10	20	4.14
	–	–	–	–	52	10	26	4.11
Right cerebellum cluster $k = 874$, $p_{FWE\text{-}corr.} = 0.009$								
Cerebellum, VI	–	–	–	–	20	−54	−24	5.88*
Cerebellum, VI	–	–	–	–	26	−58	−26	5.80*

X, *Y*, and *Z* are the stereotactic coordinates of the activations in the MNI space. All reported voxels ($p < 0.001_{uncorr}$) are included in clusters surviving the FWER correction at the cluster level. A maximum of 16 coordinates (local maxima) per cluster have been reported, each placed at least 4 mm apart, as reported by default in SPM12.
**Z*-scores statistically significant also after whole brain FWER correction ($p < 0.05$) at the voxel level (FWER, familywise error rate; MNI, Montreal neurological Institute).

and Isik, 2023). The LOTc also includes a region selective to body parts, including the hand, partially overlapping with the MT/V5 cluster and the extrastriate body area (EBA; Downing et al., 2006; Ferri et al., 2013). The EBA was found to be selective for unique interactive information present only when two individuals act simultaneously (Walbrin and Koldewyn, 2019) and to decode the observation of in-sync and out-of-sync facing individuals (Tsantani et al., 2024). Our findings suggest that the LOTc processes temporal information from observing a partner's action to support perceptual-action coupling during social synchronization. In line with this view, EBA shows strong functional connectivity with dorsal stream regions, suggesting a role in supplying information for the planning of goal-directed actions (Zimmermann et al., 2018; Zhang et al., 2021). Yet further research is needed to determine whether LOTc processing operates at a perceptual, an abstract, or both levels of representation (Papeo, 2024).

A portion of the right cerebellum (lobule VI) was active in both NI and JA tasks (Fig. 3*a*). Lobules VI and VIII of the right cerebellum are involved in sequential finger movements (Stefanescu et al., 2013), corresponding broadly to the region encoding the right hand somatotopy according to ALE meta-analyses (van der Zwaag et al., 2013; Boillat et al., 2020). However, the NI and JA tasks did not show a significant difference in the BOLD signal in any of the active clusters (Fig. 3*b*), suggesting that emergent and planned processes rely on the same neural network.

The distinction between emergent and planned synchronization is partially unexplored. Two meta-analyses on fMRI data (Chauvigné et al., 2014; Teghil et al., 2019) investigated the neural bases of synchronization during externally-paced (i.e., to an exogenous rhythmic stimulation primarily derived from the sensory context) and self-paced (i.e., to an endogenous time representation independent from external cues) tapping tasks. These two forms of synchronization resemble, broadly speaking, the distinction between emergent and planned synchronization investigated here, because emergent synchronization relies on automatic entrainment to external stimuli, while planned

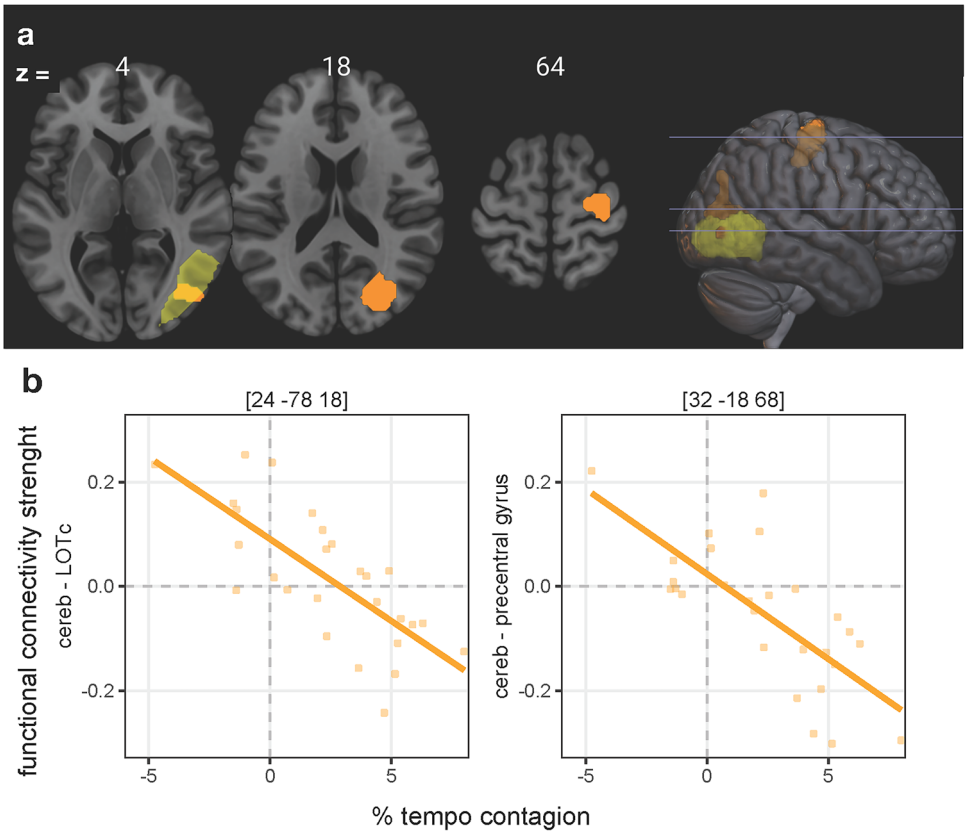

Figure 5. Results of the functional connectivity with the percent tempo contagion in the Non-Interactive condition. ***a***, Orange clusters indicate the brain regions functionally connected with the right cerebellum cluster shown in the activation task (Fig. 4*a*). Note that the cluster in the lateral occipital cortex (LOTc, orange) overlaps with the cluster found in the activation task (yellow). ***b***, Scatterplots of the negative linear relation between the functional connectivity strength and the % tempo contagion effects for both clusters. No significant results were found when correlating % tempo contagion effects from the Joint Action condition. Numbers indicate MNI coordinates.

Table 2. Results of the functional connectivity with the percent tempo contagion in the Non-Interactive condition

	Left hemisphere				Right hemisphere			
Brain area (Brodmann)	*X*	*Y*	*Z*	*Z*-score	*X*	*Y*	*Z*	*Z*-score
Right occipital cluster $k = 904$, $p_{FWE\text{-}corr.} < 0.001$								
Medial occipital gyrus (19)	–	–	–	–	24	−78	18	−5.58
	–	–	–	–	38	−72	6	−4.24
	–	–	–	–	48	−78	4	−3.54
Angular gyrus (39)					44	−76	4	−4.18
Superior occipital gyrus (7)	–	–	–	–	24	−76	38	−3.53
Right frontal cluster $k = 676$, $p_{FWE\text{-}corr.} = 0.001$								
Precentral gyrus (6)	–	–	–	–	32	−18	68	−4.81*
	–	–	–	–	42	−16	56	−3.23
Postcentral gyrus (4)	–	–	–	–	32	−28	48	−4.13

Coordinates in MNI space and *Z*-score of the resulting clusters are reported.
**Z*-scores statistically significant also after whole brain FWER correction ($p < 0.05$) at the voxel level.

synchronization requires voluntary (in a way, "self-paced") processes. In brief, externally paced timing recruited more subcortical structures (e.g., the vermis), whereas internally paced tapping recruited more cortical brain structures (e.g., the SMA). The function of the cerebellum is discussed in relation to explicit (i.e., reports of timing involving a motor performance, such as tempo reproduction tasks) and implicit (e.g., reports not involving a motor performance, such as duration/estimation tasks) timing, showing specific contributions for both (Breska and Ivry, 2016). Although the meta-analyses cited above suggested a partial distinction between external- and self-paced tapping, it is unwise to extend the same to visuomotor contexts for three reasons. First, these meta-analyses focused on audio-motor tasks, which are qualitatively different from visuomotor ones. Second, we asked participants to voluntarily sync or not, each consisting of an active motor coordination strategy different from external-paced stimulation. Third, tempo contagion effects were similar in NI and JA tasks, indicating that entrainment occurred independently of the instruction.

The conjunction analysis (Fig. 4*a*) restricted the regression analyses with tempo contagion to the co-active clusters. As mentioned above, visuomotor information might promote or prevent tempo contagion depending on the social task. The analysis supported this expectation: the higher the BOLD signal in the LOTc, the higher the % tempo contagion, but only in the JA task; in contrast, the higher the BOLD signal in the cerebellum, the lower the % tempo contagion, but only in the NI task (Fig. 4*b*). These findings have a twofold implication. First, planned synchronization is driven by tracking visuomotor information of the partner's movement, probably due to motor mimicry (e.g., the timing of the tap peak amplitude; see Uccelli et al. (2025) and Kroger et al. (2021) for similar considerations). This joint coordination is, however, at the expense of individual performance since participants mutually adjusted timing, yielding tempo contagion. Second, emergent synchronization is triggered by the same visuomotor information, but the higher the cerebellar activity, the more tempo contagion is prevented. The cerebellum is crucial in maintaining stable rhythms (Molinari et al., 2007; Andersen and Dalal, 2021), and its lateral portion specifically integrates visual and motor information critical for processing temporally structured visuomotor input (Tzvi et al., 2022).

Functional connectivity analysis revealed reciprocal communication between the cerebellar and higher-order visual areas, supporting previous evidence that the cerebellum is involved in processing biological motion information (Sokolov et al., 2012) as well as in social cognition tasks (Van Overwalle et al., 2020). In the NI task, the higher the tempo contagion, the lower the connectivity between the right cerebellum and a portion of the right LOTc, or, put otherwise, the stronger the connectivity, the lower the contagion (Fig. 5). This finding indicates that cerebellar activity favors the prevention of tempo contagion during visuomotor synchronization. The analysis also showed a negative connectivity between the cerebellum and a right fronto-parietal cluster (precentral/postcentral gyri), compatible with the idea that entrainment occurs more readily when cortical sensorimotor control is reduced. Both findings are consistent with previous evidence on the role of the cerebellum in visuomotor adaptation (Tzvi et al., 2020; Tzvi et al., 2022) and action observation (Sokolov et al., 2010). In the JA task, it was plausible to expect connectivity between the LOTc and motor or premotor regions as an index of active co-regulation, but no connectivity was associated with tempo contagion. Although this limits our ability to characterize tempo contagion in planned synchronization, it does not undermine the overall interpretation of our findings.

Last, the neural network also included motor, premotor, and sensorimotor clusters known to be active during listening to rhythms (Bengtsson et al., 2009), comprising the inferior frontal gyrus, which is involved in action observation and goal-directed actions (Molnar-Szakacs et al., 2005). Despite their role in the motor control of rhythmic movements, these areas did not correlate with tempo contagion (Fig. S3 in the Supplementary Materials). Our findings suggest that visuomotor temporal adjustments occur automatically, driven by perceptual-action correspondences in cerebellar and visual regions. These adjustments appear to escape the control of higher-order motor areas implicated in action goal representation.

Conclusion

Our findings provide evidence for the neural correlates of emergent and planned social synchronization during visuomotor coordination. Despite the well-established role of motor contagion in social interactions (Blakemore and Frith, 2005), the interplay between emergent and planned processes is equally important for understanding temporal contagion effects. In summary, we suggest that the neural networks responsible for encoding visuomotor information serve a twofold role during social synchronization: the lateral cerebellum prevents entrainment, facilitating focusing on one's own tempo, whereas the LOTc promotes voluntary joint synchronization. We propose that the balance between such antagonistic processing of visuomotor information may represent a requirement for timing regulation during social coordination.

Data Availability

Data and Supplementary Materials are available at https://osf.io/jvwz3/.

References

Andersen LM, Dalal SS (2021) The cerebellar clock: predicting and timing somatosensory touch. Neuroimage 238:118202.
Bengtsson SL, Ullen F, Ehrsson HH, Hashimoto T, Kito T, Naito E, Sadato N (2009) Listening to rhythms activates motor and premotor cortices. Cortex 45:62–71.
Blakemore SJ, Frith C (2005) The role of motor contagion in the prediction of action. Neuropsychologia 43:260–267.
Boillat Y, Bazin PL, van der Zwaag W (2020) Whole-body somatotopic maps in the cerebellum revealed with 7T fMRI. Neuroimage 211:116624.
Breska A, Ivry RB (2016) Taxonomies of timing: where does the cerebellum fit in? Curr Opin Behav Sci 8:282–288.
Brett M, Anton JL, Valabregue R, Poline JB (2002) Region of interest analysis using the MarsBar toolbox for SPM 99. Neuroimage 16:S497.
Champely S, et al. (2018) Package 'pwr'. Available at https://cran.r-project.org/web/packages/pwr/pwr.pdf
Chauvigné LA, Gitau KM, Brown S (2014) The neural basis of audiomotor entrainment: an ALE meta-analysis. Front Hum Neurosci 8:776.
Downing PE, Peelen MV, Wiggett AJ, Tew BD (2006) The role of the extrastriate body area in action perception. Soc Neurosci 1:52–62.
Enders CK, Tofighi D (2007) Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. Psychol Methods 12:121.
Fairhurst MT, Janata P, Keller PE (2013) Being and feeling in sync with an adaptive virtual partner: brain mechanisms underlying dynamic cooperativity. Cereb Cortex 23:2592–2600.
Fairhurst MT, Janata P, Keller PE (2014) Leading the follower: an fMRI investigation of dynamic cooperativity and leader–follower strategies in synchronization with an adaptive virtual partner. Neuroimage 84:688–697.
Ferri S, Kolster H, Jastorff J, Orban GA (2013) The overlap of the EBA and the MT/V5 cluster. Neuroimage 66:412–425.
Flandin G, Friston KJ (2019) Analysis of family-wise error rates in statistical parametric mapping using random field theory. Hum Brain Mapp 40: 2052–2054.
Hove MJ, Spivey MJ, Krumhansl CL (2010) Compatibility of motion facilitates visuomotor synchronization. J Exp Psychol Hum Percept Perform 36:1525.
Knoblich G, Butterfill S, Sebanz N (2011) Psychological research on joint action: theory and data. Psychol Learn Motiv 54:59–101.
Kohler N, Novembre G, Gugnowska K, Keller PE, Villringer A, Sammler D (2023) Cortico-cerebellar audio-motor regions coordinate self and other in musical joint action. Cereb Cortex 33:2804–2822.
Kroger C, Kagerer FA, McAuley JD (2021) Monkey see, monkey tap: mimicry of movement dynamics during coordinated tapping. Exp Brain Res 239: 1465–1477.
Lawrence MA, Lawrence MMA (2016) Package 'ez'. R package version, 4.
Lingnau A, Downing PE (2015) The lateral occipitotemporal cortex in action. Trends Cogn Sci 19:268–277.
McMahon E, Isik L (2023) Seeing social interactions. Trends Cogn Sci 27: 1165–1179.
Molinari M, Leggio MG, Thaut MH (2007) The cerebellum and neural networks for rhythmic sensorimotor synchronization in the human brain. Cerebellum 6:18–23.
Molnar-Szakacs I, Iacoboni M, Koski L, Mazziotta JC (2005) Functional segregation within pars opercularis of the inferior frontal gyrus: evidence

from fMRI studies of imitation and action observation. Cereb Cortex 15: 986–994.
Osborne J (2010) Improving your data transformations: applying the Box-Cox transformation. Pract Assess Res Eval 15:12.
Oullier O, De Guzman GC, Jantzen KJ, Lagarde J, Scott Kelso JA (2008) Social coordination dynamics: measuring human bonding. Soc Neurosci 3:178–192.
Papeo L (2024) What is abstract about seeing social interactions? Trends Cogn Sci 28:390–391.
Pernet CR (2014) Misconceptions in the use of the general linear model applied to functional MRI: a tutorial for junior neuro-imagers. Front Neurosci 8:1.
R Core Team (2024) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/
Repp BH, Su YH (2013) Sensorimotor synchronization: a review of recent research (2006–2012). Psychon Bull Rev 20:403–452.
Silfwerbrand L, Ogata Y, Yoshimura N, Koike Y, Gingnell M (2022) An fMRI-study of leading and following using rhythmic tapping. Soc Neurosci 17:558–567.
Sokolov AA, Gharabaghi A, Tatagiba MS, Pavlova M (2010) Cerebellar engagement in an action observation network. Cereb Cortex 20:486–491.
Sokolov AA, Erb M, Gharabaghi A, Grodd W, Tatagiba MS, Pavlova MA (2012) Biological motion processing: the left cerebellum communicates with the right superior temporal sulcus. Neuroimage 59:2824–2830.
Stefanescu MR, Thürling M, Maderwald S, Wiestler T, Ladd ME, Diedrichsen J, Timmann DA (2013) 7 T fMRI study of cerebellar activation in sequential finger movement tasks. Exp Brain Res 228:243–254.
Teghil A, Boccia M, D'Antonio F, Di Vita A, de Lena C, Guariglia C (2019) Neural substrates of internally-based and externally-cued timing: an activation likelihood estimation (ALE) meta-analysis of fMRI studies. Neurosci Biobehav Rev 96, 197–209.
Tsantani M, Yon D, Cook R (2024) Neural representations of observed interpersonal synchrony/asynchrony in the social perception network. J Neurosci 44:e2009222024.
Tzvi E, Koeth F, Karabanov AN, Siebner HR, Krämer UM (2020) Cerebellar–premotor cortex interactions underlying visuomotor adaptation. Neuroimage 220:117142.
Tzvi E, Loens S, Donchin O (2022) Mini-review: the role of the cerebellum in visuomotor adaptation. Cerebellum 21:306–313.
Uccelli S, Bruno N (2024) The Uznadze effect is temporally dynamic in closed-loop but temporally constant in open-loop grasping. Q J Exp Psychol 77:1238–1249.
Uccelli S, Sacheli LM, Paulesu E (2023) Emergent and planned interpersonal synchronization are both sensitive to 'tempo aftereffect contagion'. Neuropsychologia 181:108492.
Uccelli S, Bacchini B, Paulesu E, Sacheli LM (2025) Effects of emergent and planned interpersonal synchronization on individual spatiotemporal variability. Psychon Bull Rev 32:3214–3227.
van der Zwaag W, Kusters R, Magill A, Gruetter R, Martuzzi R, Blanke O, Marques JP (2013) Digit somatotopy in the human cerebellum: a 7T fMRI study. Neuroimage 67:354–362.
Van Overwalle F, Manto M, Cattaneo Z, Clausi S, Ferrari C, Gabrieli JD, Leggio M (2020) Consensus paper: cerebellum and social cognition. Cerebellum 19:833–868.
Walbrin J, Koldewyn K (2019) Dyadic interaction processing in the posterior temporal cortex. Neuroimage 198:296–302.
Whitfield-Gabrieli S, Nieto-Castanon A (2012) Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. Brain Connect 2:125–141.
Witt ST, Laird AR, Meyerand ME (2008) Functional neuroimaging correlates of finger-tapping task variations: an ALE meta-analysis. Neuroimage 42: 343–356.
Wolf T, Knoblich G (2022) Joint rushing alters internal timekeeping in non-musicians and musicians. Sci Rep 12:1–11.
Worsley KJ, Friston KJ (1995) Analysis of FMRI time-series revisited-again. Neuroimage 2:173–181.
Wurm MF, Caramazza A (2019) Lateral occipitotemporal cortex encodes perceptual components of social actions rather than abstract representations of sociality. Neuroimage 202:116153.
Zhang Y, Lemarchand R, Asyraff A, Hoffman P (2022) Representation of motion concepts in occipitotemporal cortex: fMRI activation, decoding and connectivity analyses. Neuroimage 259:119450.
Zhang Z, Zeidman P, Nelissen N, Filippini N, Diedrichsen J, Bracci S, Rounis E (2021) Neural correlates of hand–object congruency effects during action planning. J Cogn Neurosci 33:1487–1503.
Zimmermann M, Mars RB, De Lange FP, Toni I, Verhagen L (2018) Is the extrastriate body area part of the dorsal visuomotor stream? Brain Struct Funct 223:31–46.