COGNITIVE NEUROSCIENCE

An evolutionary conserved neural mechanism for interpersonal coordination in primates

Lucia Maria Sacheli[1,2]*†, Stefano Grasso[3]†, Laura Zapparoli[1,2], Eros Quarta[3], Fabiana Esposito[1], Marika Mariano[1], Margherita Adelaide Musco[1], Carlo Toneatto[1], Gianpaolo Basso[4,5], Stefano Ferraina[3], Alexandra Battaglia-Mayer[3]*‡, Eraldo Paulesu[1,2]‡

Interpersonal coordination is fundamental to social evolution. We investigate its phenomenology and neural underpinnings in humans and macaques, examining the behavioral adaptations required for mutual coordination during motor interactions and the extent to which the underlying brain mechanisms are shared across species. Using a common interpersonal coordination paradigm, we conducted functional magnetic resonance imaging (fMRI) in humans and intracranial local field potential (LFP) recordings in macaques. Despite between-species behavioral discrepancies, both monkeys and humans coordinate through adjustments that proved to be based on proactive adaptation of motor planning and execution. Evidence from fMRI and time-resolved decoding analysis of LFPs converged to show modulation of premotor brain activity associated with interpersonal coordination and its effectiveness. Moreover, a dynamic sequential coding emerged, whereby the action context is represented early during planning, and coordination features near movement onset. Our findings reveal an evolutionarily conserved cortical architecture across primates that supports cooperative motor behavior.

INTRODUCTION

Collaborative motor coordination is a core feature of human sociality that enables the achievement of goals that would otherwise be unattainable by individuals alone. Despite its ubiquity, the cognitive and neural architecture subserving our ability to coordinate with our conspecifics remains to be defined. Even less known is whether interpersonal coordination in humans and nonhuman primates (NHPs) relies on the same basic neurocognitive architecture. Answering these questions would greatly enhance our understanding of how social cognition has evolved across species.

For human interindividual coordination, influential philosophical research (*1*, *2*) suggests that possibly uniquely human psychological structures might be needed [(*3*, *4*); but see (*5*, *6*)]: These have been variously labeled as "shared," "collective," or "we-" intentions (*1*, *2*). On the other hand, more reductionist cognitive (neuro)science views indicate that interpersonal coordination may depend on low-level reactive perceptual and sensorimotor processes [e.g., entrainment, temporal adaptation, and task corepresentation (*7*–*9*)], which do not require any explicit intention to coordinate and may also occur in other species (*10*). However, one may hypothesize the middle-ground position that interpersonal coordination is not, or not only, a matter of higher-order cognitive representations such as shared intention nor is it reducible to automatic behavioral adjustments. One such third position posits that interpersonal coordination may entail the proactive generation of dyadic motor plans encoded within premotor regions (*11*–*13*). The term dyadic motor plan refers to a motor representation adapted to dyadic interactions, which takes into account not only the actions one must execute but also those their partner must perform to achieve the interaction goal. This modality of motor planning is expected to manifest as proactive adjustments in the spatiotemporal features of movement, thereby facilitating effective interpersonal coordination.

Could these representations serve as foundational elements of interpersonal coordination in both humans and NHPs? To investigate this possibility, we used a cross-species and multiscale approach, examining pairs of human participants and two exemplars of *Macaca mulatta* as they engaged in the same experimental task. We analyzed the behavioral adjustments required for a collaborative visuomotor coordination task, contrasting this with the same task performed individually, and studied the neural mechanisms that underlie these coordination processes. Neural underpinnings were studied in humans through functional magnetic resonance imaging (fMRI; Fig. 1A), and in monkeys by analyzing the local field potentials (LFPs) recorded from the dorsal premotor cortex (PMd; F2) of two macaque brains simultaneously (Fig. 1B). Among the different intracortical signals recorded from monkeys, we focused on LFPs because they are the primary neural correlates of the blood oxygen level–dependent (BOLD) fMRI signal (*14*, *15*). In both LFP and fMRI, the neural response was sampled by also including the motor planning phase.

Fifty young adult human participants and two adult male rhesus monkeys [*M. mulatta*; Monkey D (Mk D) and Monkey P (Mk P)] were tested in identical dyadic visuomotor tasks (Fig. 1, A and B), requiring interindividual motor coordination while using a joystick to guide a visual cursor on a computer screen. Humans performed the task with one participant lying in the MRI scanner and their partner playing in an adjacent room (Fig. 1A). Monkeys performed the task sitting side by side with simultaneous LFP recordings from the two brains (Fig. 1B).

Each agent guided a visual object (a yellow circle) displayed on the computer screen from a central target toward a peripheral target (Fig. 1C), either individually (SOLO) or coordinating their actions with the partner in a joint action (JOINT). During SOLO trials, the center of the yellow circle coincided with the location of the cursor,

[1]Department of Psychology and Milan Center for Neuroscience (NeuroMi), University of Milano-Bicocca, Milan 20126, Italy. [2]fMRI Unit, IRCCS Istituto Ortopedico Galeazzi, Milan 20157, Italy. [3]Department of Physiology and Pharmacology, Sapienza University of Rome, Rome 00185, Italy. [4]School of Medicine and Surgery, University of Milano-Bicocca, Milan 20126, Italy. [5]Department of Neuroscience, Fondazione IRCCS San Gerardo dei Tintori, Monza (MB) 20900, Italy.
*Corresponding author. Email: lucia.sacheli@unimib.it (L.M.S.); alexandra.battagliamayer@uniroma1.it (A.B.-M.)
†These authors contributed equally to this work.
‡These authors contributed equally to this work.

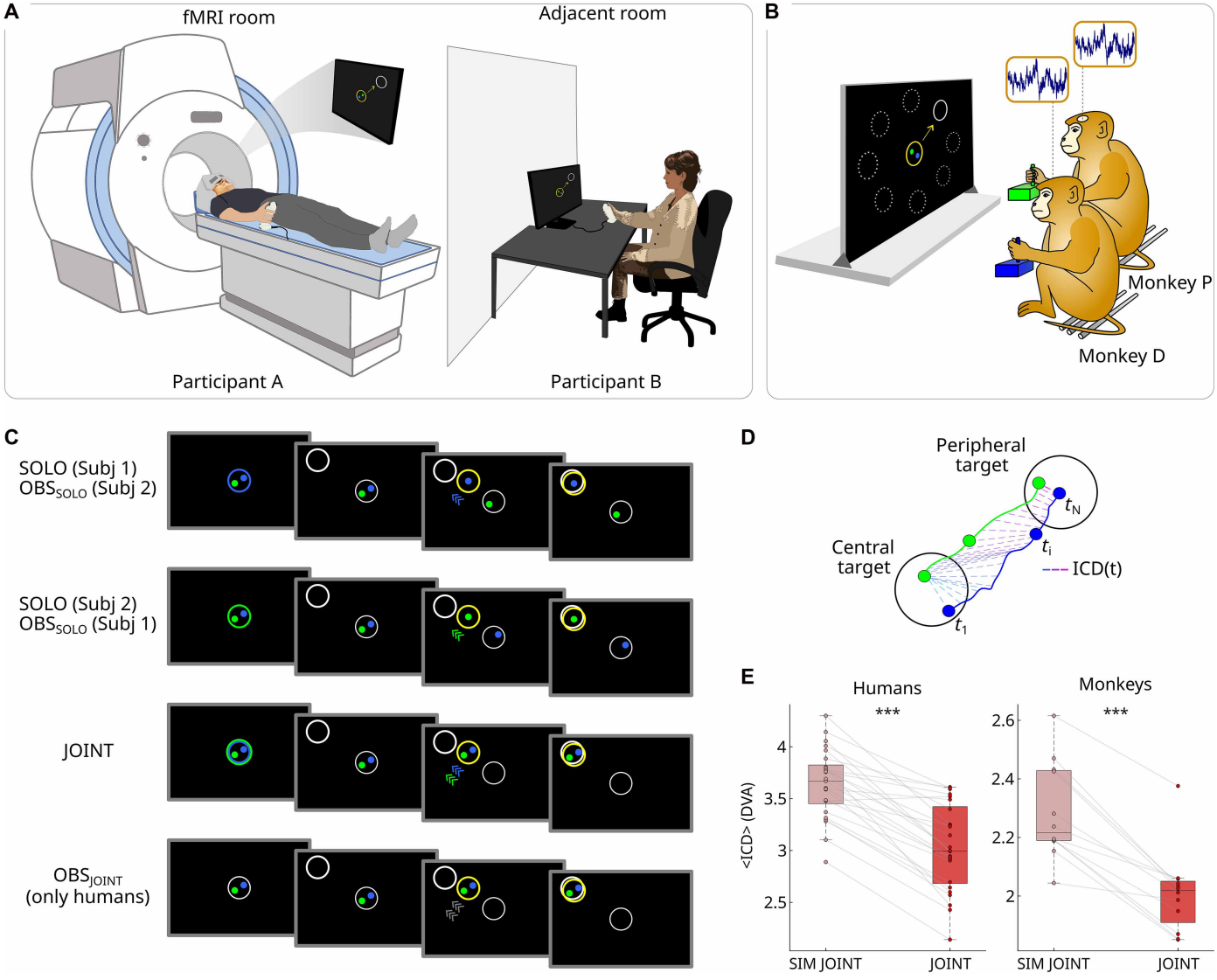

Fig. 1. Experimental setup. The experimental setup adopted with (**A**) human participants and (**B**) monkeys. (**C**) Behavioral task used in both humans and monkeys, with different trial types. Subj, subject (i.e., human participant or monkey). (**D**) Schematic representation of how the ICD was calculated in JOINT trials; green and blue dots represent the positions at different times (*t*) of the two cursors, each controlled by one agent. (**E**) Comparison of <ICD> values recorded during JOINT action execution and the one calculated by randomly pairing individual trajectories recorded within each pair (humans) and session (monkeys) during JOINT trials [SIMULATED (SIM) JOINT; see the Supplementary Materials for methodological details]. ***$P < 0.001$.

and only one participant was responsible for moving the yellow circle to the peripheral target. Meanwhile, the partner observed the cursor movement on the screen and kept their own cursor stationary in the central target (i.e., OBS_{SOLO} condition for the partner). During the JOINT condition, the center of the yellow circle corresponded to the midpoint of the two participants' cursor positions. In these trials, partners had to coordinate their movements to guide the yellow circle toward the peripheral target while maintaining both cursors not exceeding a maximal intercursor distance (ICD) beyond which the trial was aborted (Fig. 1, C and D; see the Materials and Methods section).

In humans only, an additional "observation" condition (OBS_{JOINT}) was introduced as a perceptual baseline, in which both participants passively observed computer-simulated cursor movements based on the trajectories recorded during successful JOINT trials in pilot participants. The trial type (SOLO, JOINT, OBS_{SOLO}, and, in humans, OBS_{JOINT}) was preinstructed by a color change of the central cue presented before the peripheral target appeared (Fig. 1C and tables S1 and S2). In the monkey experiment, each successful trial (SOLO, JOINT, or OBS_{SOLO}) was rewarded with the same amount of juice. In JOINT trials, both monkeys received the same amount of reward upon successful coordination; otherwise, neither animal was rewarded.

To measure interpersonal coordination, we calculated for each trial the mean intratrial ICD (<ICD>; Fig. 1D). Human participants were instructed (and the monkeys were trained) to bring the yellow circle to the peripheral target while keeping their cursors close together. The instantaneous ICD was constantly monitored during JOINT trials: Exceeding a maximum ICD threshold [9 degrees of visual angle (DVA) for humans and 5 DVA for monkeys] led to trial abortion.

Thus, spatiotemporal interpersonal coordination was required throughout the JOINT task and represented the interaction goal for both humans and macaques.

We hypothesized that evidence for dyadic motor planning would emerge if humans and monkeys displayed signs of proactive adaptations in their motor planning and execution. This would manifest as an alteration of their stereotypical motor patterns to enhance coordination, indicated by proactive adjustments of spatiotemporal movement parameters during JOINT trials to minimize their ICD. At the neural level, we tested the role of the PMd—a critical region for motor planning for the type of movement tested (*13*)—in supporting these coordination processes in both species.

RESULTS

Evidence for interpersonal coordination in JOINT action (behavioral analysis #1)

To explore whether the results of the JOINT task provided behavioral evidence for a nonepiphenomenal interpersonal coordination in both humans and NHPs, we compared the <ICD> recorded during JOINT trials with "SIMULATED JOINT" data obtained by randomly pairing, within each dyad (in humans) and session (in monkeys), the trajectories of the two cursors in different JOINT trials in a given direction (e.g., *n*th trial of participant #1 with *m*th trial of participant #2, etc.). To calculate SIMULATED JOINT trials, we paired only successful trials in which the same peripheral target was reached, and the trajectories were aligned on movement time onset (MTon). The results showed that the <ICD> was lower (Table 1; see "Comparison between JOINT and SIM data"), indicating better interpersonal coordination, in the real JOINT than the SIMULATED JOINT data in both species (Fig. 1E). These results suggest that, in JOINT action, individuals were not merely acting simultaneously, guiding "in parallel" their cursors along their stereotypical individual trajectories, but actively modulated their behavior by minimizing the relative distance between their cursors with the aim of achieving effective online interpersonal coordination.

Behavioral adjustments in JOINT action (behavioral analysis #2)

To compare the motor behavior between JOINT and SOLO trials, we analyzed (i) reaction times (RTs), (ii) movement times (MTs), and (iii) the cursor's trajectory deviations (TDs) from the linear pathway linking the starting position (center of the central target) to the center of the peripheral target (Table 1; see "Comparison between the SOLO and JOINT condition").

Human participants overall showed longer RTs and MTs in the JOINT condition compared to the SOLO condition (Fig. 2, Aa and Ab), indicating that both motor preparation and execution were influenced by the need to coordinate with the partner, which required more time. This time increase may be due to the recruitment of additional/more demanding cognitive operations, or it may be a strategy that participants adopt to facilitate the coordination task by making themselves more predictable. Furthermore, these adjustments in the time domain were accompanied by spatial adjustments, whereby participants guided their cursors along more linear trajectories (smaller TD) than those typical of SOLO actions (Fig. 2Ac).

NHPs were also able to adapt their behavior to optimize the dyadic performance, in line with previous studies (*16*, *17*). It is worth noting that, in the present experiment, the two animals in the SOLO condition showed significantly different timings: Mk D was naturally faster than the other animal (Mk P), in both RTs and MTs [RTs: $t(1421.2) = 22.6$, $P < 0.001$, mean difference (±SE) between animals equal to 114 ± 5.07 ms; MTs: $t(1421.4) = 11.5$, $P < 0.001$, mean difference (±SE) between animals equal to 96.3 ± 8.37 ms)]. In JOINT trials, these inter-individual differences had to be necessarily compensated to solve the task cooperatively. The slowest monkey (Mk P) showed pronounced adjustments to its partner, by reducing both RTs and MTs in the JOINT condition (Fig. 2, Ad and Ae). In contrast, the faster monkey (Mk D) did not change significantly its RTs, yet it showed a significant slowdown during task execution (longer MTs). These behavioral changes collectively contributed to dyadic synchronization, with greater adaptation by Mk P to its mate than vice versa. As in humans, both macaques reshaped their individual SOLO trajectories, adopting more linear paths in JOINT trials (Fig. 2Af).

Relationship between behavioral adjustments and better interpersonal coordination in JOINT action (behavioral analysis #3)

We then explored which of the behavioral adjustments shown in the JOINT compared with the SOLO trials were functional to improve joint performance. We first calculated normalized JOINT behavioral indexes (*J,norm*) by estimating normalized deviations from the mean value recorded from the same participant in the SOLO condition ($\langle s \rangle$)

$$J, norm = \frac{\left(j_j - \langle s \rangle\right)}{\left[\frac{j_j + \langle s \rangle}{2}\right]}$$

where j_j is the value of the dependent variable (RT, MT, and TD) measured in each JOINT trial and $\langle s \rangle$ is the individual mean value of the dependent variable calculated in the SOLO condition.

Using linear mixed models, we tested whether *J,norm* values could predict trial-by-trial changes in the <ICD>, our behavioral index of interpersonal coordination. In this model, by-subjects (in humans) and by-session (in monkeys) random intercepts and slopes were included to account for between-subjects (in humans) or between-sessions (in monkeys) variability: We thus tested whether, within each pair of human participants and within each experimental session performed by each monkey, some specific behavioral adjustments trial-by-trial determined better interpersonal coordination.

In humans, changes in both MTs ($MT_{J,norm}$) and trajectory ($TD_{J,norm}$), but not in RTs ($RT_{J,norm}$), determined better interpersonal coordination, as indicated by lower <ICD> [Fig. 2, Ba and Bc, and Table 1; see "Regressions analyses (RTj,norm, MTj,norm, TDj,norm versus ICD)"]. In other words, although humans overall dedicated more time (and possibly cognitive resources) to plan JOINT actions, the pair performance in terms of interpersonal coordination was trial-by-trial predicted by how much they slowed down during movement execution and adapted their cursor trajectory along a straighter line (as compared to the SOLO condition), thus maintaining their cursors closer while reaching the target together. These results held even when considering the relationship between behavioral variables within a unique multiple regression model, as reported in the Supplementary Materials. Moreover, as expected, the relationship emerged between *J,norm* values substantially differed from the one emerged between behavioral variables in SOLO trials, as reported in the Supplementary Materials.

Table 1. Behavioral results. The table reports the comparison between the <ICD> measured in JOINT and SIMULATED JOINT (SIM) data [human model: <ICD> ~ 1+ JOINT/SIM + (1+ JOINT/SIM|Participant); monkey model: <ICD> ~ 1+ JOINT/SIM + (1+ JOINT/SIM|Session)], the comparison between the SOLO and JOINT condition in different dependent variables (DVs) [human model: DV ~ 1+ SOLO/JOINT + (1+ SOLO/JOINT|Participant); monkey model: DV ~ 1+ SOLO/JOINT + (1+ SOLO/JOINT|Session)], and the results of the regression analyses aimed to assess which behavioral adjustments ($DV_{J,norm}$) best predicted interpersonal coordination in the JOINT condition [human model: <ICD> ~ 1+ $DV_{J,norm}$ + (1+ $DV_{j,norm}$|Participant); monkey model: <ICD> ~ 1+ $DV_{J,norm}$ + (1+ $DV_{J,norm}$|Session)]. To calculate $DV_{J,norm}$, we used the following formula: $J,norm = (j_j - \langle s \rangle)/[(j_j - \langle s \rangle)/2]$, where j_j is the value of the dependent variable measured in each JOINT trial and $\langle s \rangle$ is the individual mean value of the dependent variable calculated in the SOLO condition. The asterisks (*) indicate when the full model did not converge, and the random structure of the model was simplified by removing the random slopes (while maintaining the random intercepts). <ICD>, mean within-trial ICD; RT, reaction times; MT, movement times; TD, trajectory deviations. The human results refer to the participants who played the game while being fMRI scanned; these results were fully replicated by the analysis of the behavior of the partner sitting in the MRI control room (described in Supplementary Text in the Supplementary Materials).

Comparison between JOINT and SIM data

		t values	df	*P*	Mean diff (JOINT/SIM)	SE	95% CI
Humans	<ICD>	−9.75	21.17	<0.001	−0.60 DVA	0.06	−0.72−−0.48
Monkeys	<ICD>	−6.54	1397.8*	<0.001	−0.27 DVA	0.04	−0.35−−0.19

Comparison between the SOLO and JOINT condition

		t values	df	*P*	Mean diff (JOINT/SOLO)	SE	95% CI
Humans	RT	6.95	23.90	<0.001	145.02 ms	20.88	104.10–185.94
	MT	9.18	23.94	<0.001	1312.85 ms	142.94	1032.69–1593.01
	TD	−2.23	23.72	0.036	−39.16 px	17.59	−73.63−−4.69
Monkey D	RT	1.67	1415.56*	0.094	4.40 ms	2.63	−0.75–9.56
	MT	2.87	1417.38*	0.004	21.74 ms	7.58	6.88–36.61
	TD	−1.95	1416.52*	0.052	−0.59 DVA	0.30	−1.18–0.00
Monkey P	RT	−11.99	1418.45*	<0.001	−63.10 ms	5.26	−73.40−−52.77
	MT	−2.53	1418.33*	0.011	−22.74 ms	8.98	−40.35−−5.13
	TD	−3.23	1418.03*	0.001	−0.82 DVA	0.25	−1.32−−0.32

Regressions analyses ($RT_{j,norm}$, $MT_{j,norm}$, $TD_{j,norm}$ versus ICD)

		t values	df	*P*	*B* estimate	SE	95% CI
Humans	$RT_{j,norm}$	−0.62	616.85*	0.535	−0.06 DVA	0.10	−0.25–0.13
	$MT_{j,norm}$	−2.36	22.38	0.027	−0.36 DVA	0.15	−0.66−−0.06
	$TD_{j,norm}$	7.04	23.56	<0.001	0.50 DVA	0.07	0.36–0.64
Monkey D	$RT_{j,norm}$	−2.06	12.68	0.061	−0.33 DVA	0.16	−0.64−−0.02
	$MT_{j,norm}$	6.56	708.8*	<0.001	0.38 DVA	0.06	0.26–0.49
	$TD_{j,norm}$	5.42	9.01	<0.001	0.38 DVA	0.07	0.24–0.51
Monkey P	$RT_{j,norm}$	8.35	677.1*	<0.001	0.99 DVA	0.12	0.76–1.22
	$MT_{j,norm}$	7.67	674.3*	<0.001	0.56 DVA	0.07	0.42–0.70
	$TD_{j,norm}$	4.30	500.4*	<0.001	0.20 DVA	0.05	0.11–0.29

As for NHPs, the dyad showed a "leader/follower" pattern in the behavioral adaptation during joint action, with optimization of the interindividual coordination mostly due to Mk P (the "follower"). For clarity, we use the terms leader and follower to denote relative contributions to behavioral adjustments, reflecting the extent to which each monkey modified its performance to align with its partner's behavior.

Mk P shortened its RTs and MTs, compared to its SOLO performance, to adapt to the faster partner (Fig. 2, Ad and Ae, and Table 1; see "Comparison between the SOLO and JOINT condition"), and these changes predicted smaller <ICD> [Fig. 2, Bd and Be, and Table 1; see "Regressions analyses (RTj,norm, MTj,norm, TDj,norm versus ICD)"]. On the other hand, Mk D, identified as the "leader," did not vary its RTs significantly but showed slower MTs compared with the SOLO baseline. However, the attempt to make its MTs more similar to the slower partner did not result in smaller <ICD>, suggesting a less effective adaptation for Mk D. Last, both monkeys optimized their dyadic coordination (i.e., showed smaller <ICD>) by making their movement trajectories straighter during JOINT trials, replicating what was observed in humans [Fig. 2Bf and Table 1; see "Regressions analyses (RTj,norm, MTj,norm, TDj,norm versus ICD)"]. These results held even when considering the relationship between behavioral variables within a unique multiple regression model, as reported in the Supplementary Materials.

Adaptations to the history of the partner's movement patterns (behavioral analysis #4)

Last, we also explored whether reciprocal behavioral adjustments to the partner's preceding behavior might play a role. This, we reasoned, would have been a further demonstration of the proactive nature of

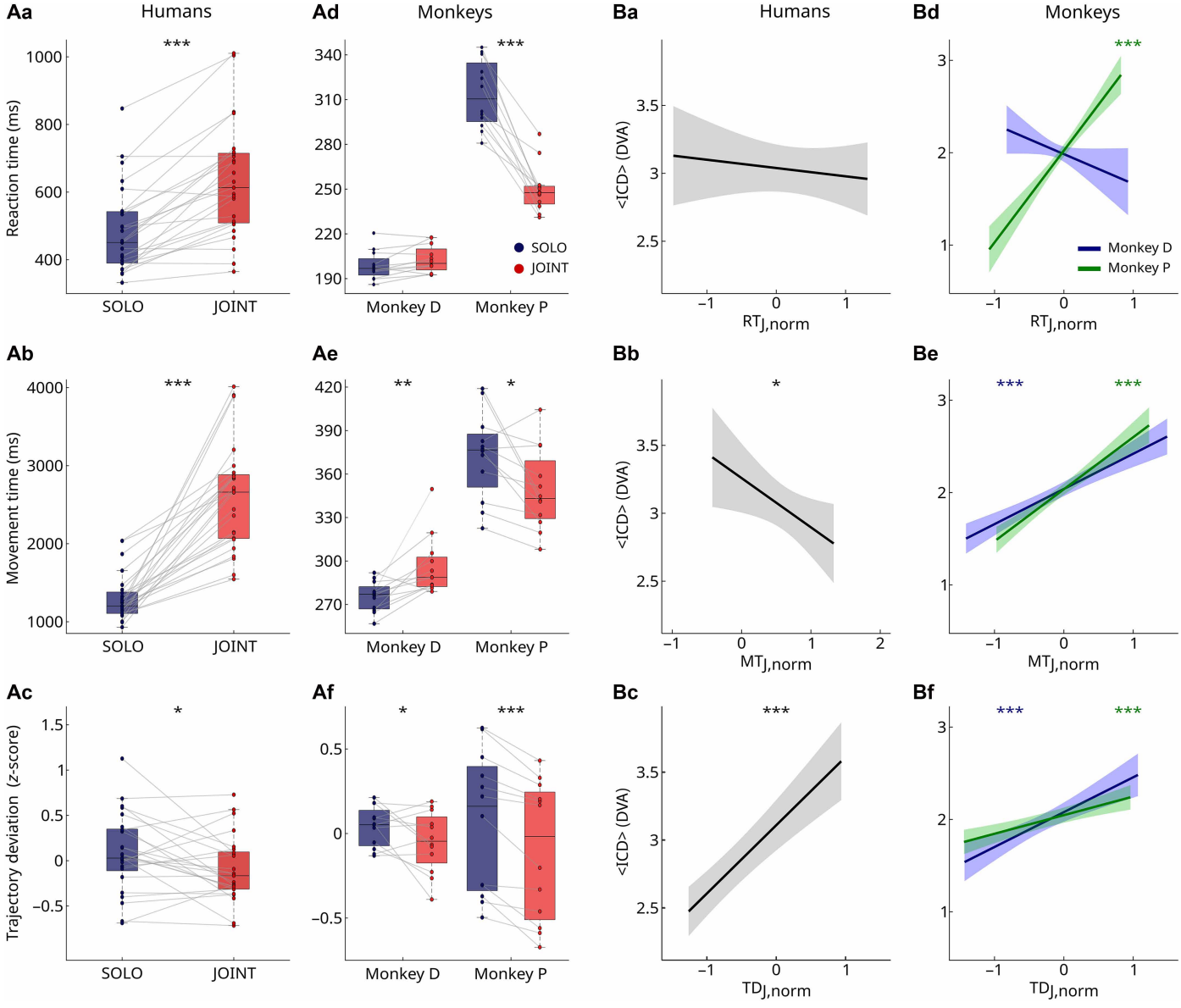

Fig. 2. Behavioral adaptations shown by humans and monkeys during JOINT action. (**A**) Comparisons made on kinematics variables (RTs, MTs, and TDs) between the SOLO and JOINT conditions in humans [(Aa), (Ab), and (Ac)] and monkeys [(Ad), (Ae), and (Af)]. (**B**) Results of the regression analyses showing which of the behavioral adjustments shown in (A) predicted better interpersonal coordination, as indexed by lower <ICD> values, in humans [(Ba), (Bb), and (Bc)] and monkeys [(Bd), (Be), and (Bf)]; here, *J,norm* positive values indicate that JOINT values are higher than SOLO, whereas *J,norm* negative values indicate that JOINT values are smaller than SOLO. This latter analysis was run on JOINT trials only. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

the motor adaptations as it would demonstrate that the participants/monkeys incorporate the history of the partner's motor behavior in its progressive unfolding. To do so, we calculated the absolute value of the difference between each participant/monkey's RT or MT in a given trial and the partner's average RT or MT calculated from all the previous JOINT trials, as follows

$$\Delta RT_{a,<b>} = |RT_a - \langle RT_b \rangle|$$

$$\Delta MT_{a,<b>} = |MT_a - \langle MT_b \rangle|$$

where RT_a (MT_a) refers to the RT (MT) recorded from one agent in each trial, and $\langle RT_b \rangle$ ($\langle MT_b \rangle$) refers to the partner's mean RT (MT) calculated on all previous JOINT trials. The $\Delta RT_{a,<b>}$ and $\Delta MT_{a,<b>}$ delta values thus indicate how much one agent's RT or MT diverges from (or converges on) what the agent might have learned about the partner's behavior. This analysis was restricted to the temporal parameters (RT and MT) because these are the ones showing signs of reciprocal adaptations in monkeys in analysis #2.

Using linear mixed models, we tested whether $\Delta RT_{a,<b>}$ and $\Delta MT_{a,<b>}$ values predict trial-by-trial changes in the <ICD>, the behavioral index of interpersonal coordination. In this model, by-subjects (in humans) and by-session (in monkeys) random intercepts

and slopes were included to account for between-subjects (in humans) or between-sessions (in monkeys) variability: We thus tested whether, within each pair of human participants and within each experimental session performed by the monkeys, the agents' adjustments to their partner's preceding behavior determined better interpersonal coordination.

In humans, whereas the analysis on $\Delta RT_{a,<b>}$ showed no significant effect [$t(26.38) = -0.61$, $P = 0.55$; fixed effect (B) estimate: −0.115 DVA; SE: 0.19 DVA; 95% confidence interval (CI): −0.49 to 0.26 DVA], the results on MTs showed that the higher the between-partners convergence (i.e., the lower the $\Delta MT_{a,<b>}$), the better the coordination, as indicated by lower <ICD> [$t(15.83) = 2.38$, $P = 0.03$; fixed effect (B) estimate: 0.25 DVA; SE: 0.105 DVA; 95% CI: 0.04 to 0.46 DVA].

In NHPs, as expected from previous findings, a higher convergence of RTs (i.e., lower $\Delta RT_{a,<b>}$ values) predicted better interpersonal coordination (i.e., smaller <ICD>) in the "follower" Mk P [$t(675.7) = 6.24$, $P < 0.001$; fixed effect (B) estimate: 2.68 DVA; SE: 0.43 DVA; 95% CI: 1.84 to 3.52 DVA]; in the "leader" Mk D, this effect was not significant [$t(52.8) = 0.56$, $P = 0.575$; fixed effect (B) estimate: 0.42 DVA; SE: 0.75 DVA; 95% CI: −1.04 to 1.88 DVA]. With regard to the MTs, the results showed a significant effect in Mk D [$t(697.6) = -2.25$, $P = 0.025$; fixed effect (B) estimate: −0.56 DVA; SE: 0.25 DVA; 95% CI: −1.05 to −0.07 DVA], which, nevertheless, was in the opposite direction than expected [but coherent with the results emerged from the analysis of *J,norm* values; see Table 1, "Regressions analyses (RTj,norm, MTj,norm, TDj,norm versus ICD)"]: It showed that the less the between-partners convergence (i.e., higher $\Delta MT_{a,<b>}$), the lower the <ICD> (i.e., the better the performance). In Mk P, the results showed no significant effect [$t(693.2) = 1.29$, $P = 0.20$; fixed effect (B) estimate: 0.25 DVA; SE: 0.19 DVA; 95% CI: −0.13 to 0.62 DVA].

Overall, these results indicate that, in human participants, adapting one's behavior to that shown by the partner in previous JOINT trials is beneficial for interpersonal coordination. This is especially true for the temporal features of movement execution, whereas adaptations of RTs did not play any role. Instead, in NHPs, the "follower," slower, and more adaptive monkey (Mk P) adapted its RTs to the ones shown by the partner in previous JOINT trials, and this adjustment was beneficial for interpersonal coordination. Adaptations in MTs, instead, had no similar beneficial effect.

A neural mechanism for interpersonal coordination

The aforementioned behavioral patterns were mirrored by the neurophysiological findings obtained from fMRI in humans and LFP recording in monkeys.

As shown in Fig. 3A, the fMRI data revealed that similar cortical and subcortical neural resources were recruited during SOLO and JOINT trials in humans, when controlling for perceptual stimulation (i.e., comparing SOLO and JOINT task execution with the respective OBS_{SOLO} and OBS_{JOINT} conditions). These regions included primary sensorimotor, dorsal premotor, and the supplementary motor areas bilaterally, as well as the basal ganglia, the thalamus, and the cerebellum (table S4). However, we also found a significant interaction effect between task (execution or observation) and condition (JOINT or SOLO) whereby the execution of the JOINT condition was associated with stronger recruitment of the PMd and posterior parietal cortices bilaterally (Fig. 3, A to C, and table S5). In other words, the activation of dorsal frontoparietal brain regions—rather than other brain resources such as subcortical structures or the cerebellum, or visual-motion areas, associated with low-level entrainment (*18*)—was stronger when the participants had to coordinate with the partner (JOINT) than when playing alone (SOLO) (Fig. 3, A to C).

We then explored the functional role of these stronger activations in guiding interpersonal coordination by combining fMRI data with behavioral measures. At the group level, a regression analysis revealed that the better the pairs' coordination in the JOINT action, as indexed by lower <ICD>, the stronger the interaction effect (i.e., the stronger the brain activity in JOINT than SOLO, controlled for the respective observation tasks) in the PMd bilaterally (Fig. 3D).

Premotor mechanisms guiding joint actions were further examined through the LFPs recorded simultaneously from the area PMd (F2; Fig. 4A) of the two interacting NHPs.

Time-frequency analysis revealed that low-frequency bands below 12 Hz were predominant during both SOLO and JOINT action tasks (Fig. 4B), about 250 ms before and 250 ms after MTon. To better analyze task-related power modulations across recording sites, we studied power variations in different frequency bands. In both animals, frequency-domain analysis revealed power increase in the Theta (4 to 7 Hz) and Alpha (8 to 12 Hz) bands during JOINT behavior compared with SOLO actions (Fig. 4C). We focused on these frequency bands for two main reasons: First, Alpha power (8 to 12 Hz) has been shown to increase with task difficulty in motor tasks requiring reactive control (*19*–*23*). Second, both Theta and Alpha bands encompass key LFP components (P1, N1, and P2; Fig. 4D) known to be modulated by kinematic parameters (e.g., target direction, velocity, and trajectory) as well as by the behavioral context (*24*).

To analyze task-related variations in the time domain, a grand-averaged low-frequency LFP signal (1 to 12 Hz) was computed for Mk P, which showed traces with consistent polarity across sites (see the Supplementary Materials). We found a significant enhancement in signal amplitude during JOINT trials, emerging ~150 ms before movement initiation (MTon) (Fig. 4D; sliding-window t test, window width: 40 ms; step size: 20 ms; df = 3761; $P < 0.001$) and with maximum significance at the time window [−24, 16] ms relative to the MTon [$t(3761) = -6.4$].

We next applied time-resolved decoding with a support vector machine (SVM) to test whether low-frequency premotor LFPs carry information about action context (SOLO versus JOINT). To identify the most informative epoch, we used two distinct data alignments: RT onset (GO alignment) and MTon. Under GO alignment, classification accuracy displayed two pronounced peaks, reaching 80 to 90% across the RT interval beginning ~100 ms after the cue (Fig. 4E, left; permutation test, $P < 0.01$). When realigned to MTon, the overall accuracy was slightly reduced but revealed an additional rise late in execution, coincident with goal attainment (Fig. 4E, right; permutation test, $P < 0.01$). The RT-locked maxima occurred near the canonical low-frequency premotor components P1 and N1, indicating that activity in the range of 4 to 12 Hz (Theta-Alpha) conveys context-specific information about the impending motor plan.

To determine the temporal format of this representation, we performed cross-temporal decoding (*25*–*27*), training and testing the classifier on all pairs of time bins with alignment to either RT or MT. This yields a time-by-time accuracy matrix, in which diagonal elements reflect time-resolved decoding when training and testing coincide (white trace in Fig. 4E), and off-diagonal elements quantify generalization across time. Broad off-diagonal accuracy would indicate a static (time-invariant) code; confinement of accuracy to the diagonal indicates a dynamic, time-specific code; and narrow,

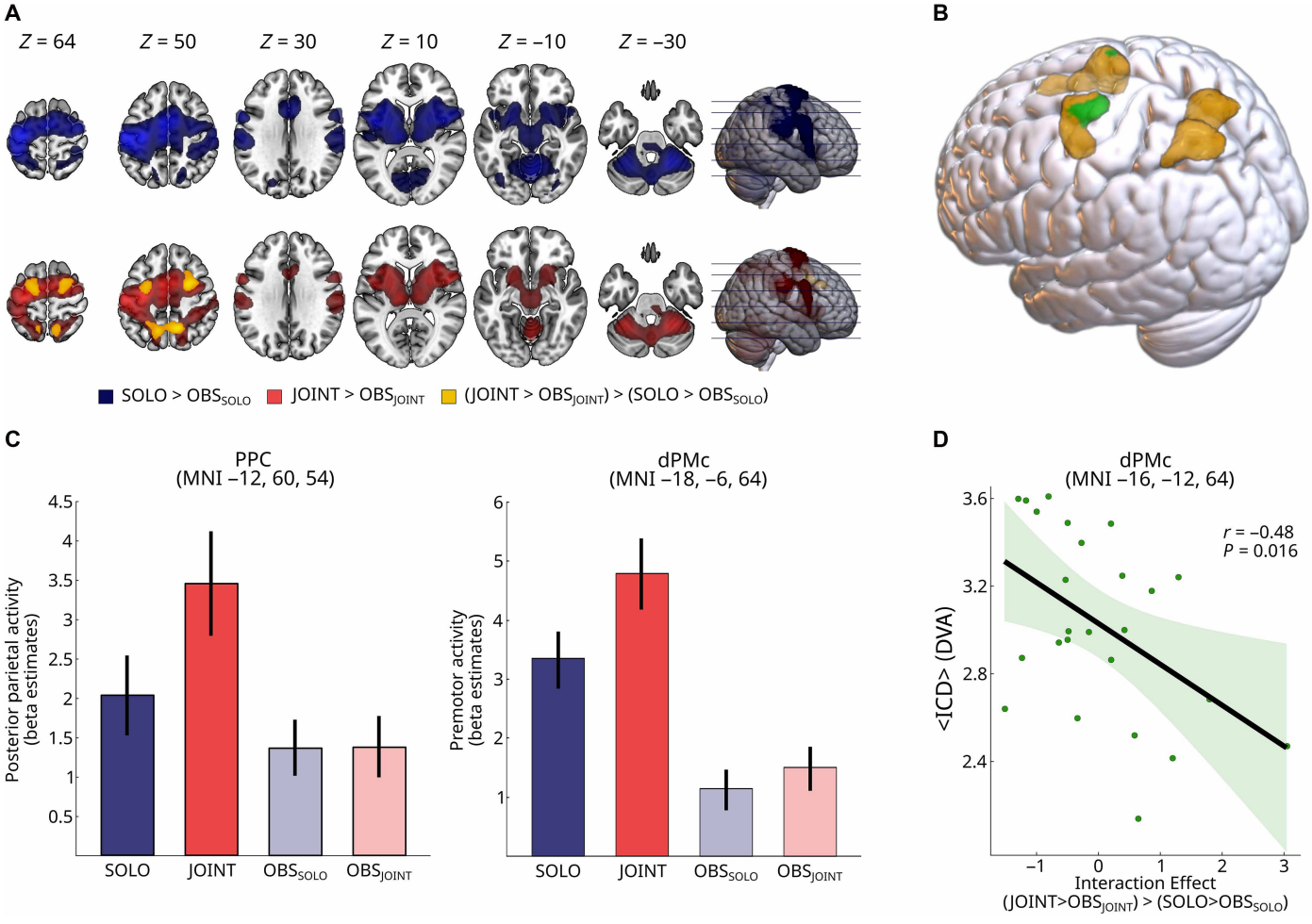

Fig. 3. Human fMRI correlates of action context and interpersonal coordination. (**A**) Brain regions that were more active during the execution than observation of SOLO trials (blue; simple effect of "SOLO > OBS_{SOLO}") and JOINT trials (red; simple effect of "JOINT > OBS_{JOINT}"); the lower map has been superimposed on the one (yellow) reporting the interaction effect "(JOINT > OBS_{JOINT}) > (SOLO > OBS_{SOLO})"; all $P_s < 0.001$ at the voxel level, and $P_{FWERcorr} < 0.05$ at the cluster level. (**B**) The same yellow map is shown on a 3D render together with the two subclusters (green) in which the strength of the interaction effect was significantly associated with the behavioral performance indexed by the mean ICD (<ICD>), as shown by a regression analysis [left premotor cluster ($k = 134$) local maximum: Montreal Neurological Institute coordinates (MNI) −16, −12, 64, z-score = 2.42, and $P = 0.008$; right premotor cluster ($k = 62$) local maximum; MNI 22, −2, 60, z-score = 1.99, and $P = 0.023$]. (**C**) Plots of the BOLD effect (beta values) as measured in the SOLO (blue) and JOINT (red) execution and observation conditions, extracted from the local maxima of the left posterior parietal cortex (PPC) and PMd (dPMc) clusters showing a significant interaction effect. (**D**) Between-participants association between the strength of the interaction effect found significant in the left PMd (dPMc) and the <ICD>.

short-lived diagonal bands indicate transient encoding with limited cross-time generalization.

The resulting matrices revealed a predominantly time-selective (dynamic) code for action context: Decoding emerged ~100 ms after GO and remained confined to a diagonal band throughout the RT period, with little off-diagonal generalization. Thus, premotor low-frequency LFPs encode action context with a time-selective process belonging to the planning phase.

To examine the relationships between the level of partners' coordination (indexed by the <ICD>) during JOINT trials and PMd LFPs, trials were grouped into three datasets based on mean ICD values (large, medium, and small). LFP power increased progressively with higher coordination (lower <ICD>), particularly in Mk P across all bands <30 Hz, and in the 4- to 7-Hz range in Mk D (Fig. 5A). Temporal analysis showed that LFP amplitude predicted the magnitude of the <ICD> ~150 ms before the onset of force application (MTon) (Fig. 5B; sliding-window t test, window width: 40 ms; step size: 20 ms; df = 1236; $P < 0.001$) and with maximum significance at the time window [−4, 36] ms relative to the MTon [$t(1236) = -6.36$].

Similar to the context-decoding analysis, we applied a time-resolved, cross-validated SVM to test whether premotor LFPs also predict the level of interindividual coordination, operationalized as Small versus Large ICD, our behavioral index of dyadic alignment. We found that low-frequency LFPs carried sufficient information to decode not only action context (SOLO versus JOINT) but also coordination level, albeit with distinct temporal profiles (Figs. 4E and 5C). Under MTon alignment, decoding accuracy for coordination quality exhibited clear peaks of ~80 to 90% centered on movement initiation (about 70 ms before and after MTon; permutation test, $P < 0.01$; Fig. 5C, right). By contrast, when aligned to the GO cue, classification performance was substantially

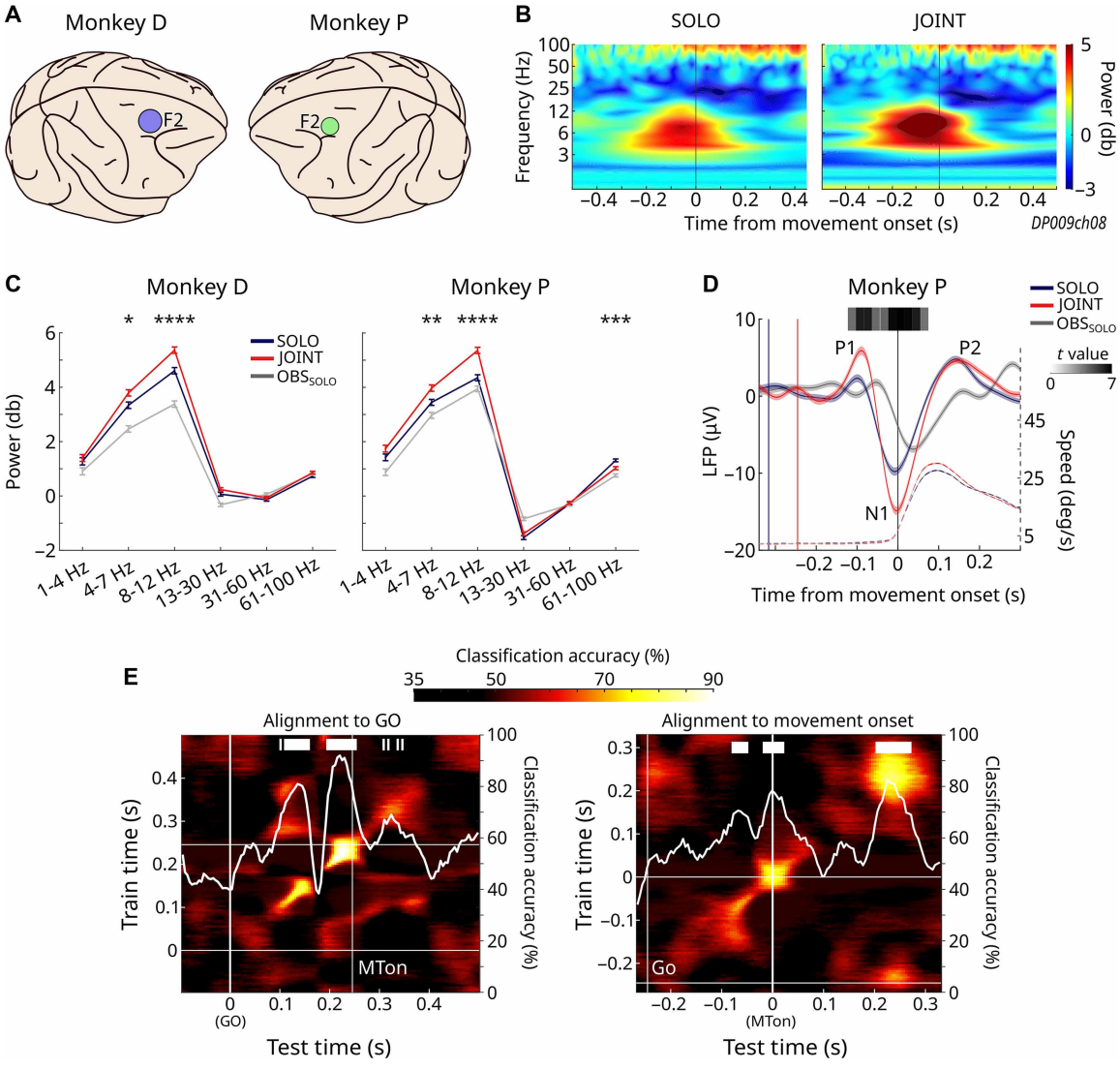

Fig. 4. NHP LFP correlates of action context. (**A**) Recording sites (PMd; F2) in Mk D and Mk P, contralateral to the used (preferred) limb. (**B**) Time-frequency spectrogram LFP from a representative PMd channel during SOLO and JOINT condition. (**C**) Mean (± SEM) band-limited power for Delta (1 to 4 Hz), Theta (4 to 7 Hz), Alpha (8 to 12 Hz), Beta (13 to 30 Hz), low-Gamma (31 to 60 Hz), and high-Gamma (61 to 100 Hz), averaged across trials and sites for SOLO (blue), JOINT (red), and OBS_{SOLO} (gray) conditions. Theta and Alpha power increased during JOINT behavior compared to individual actions in both monkeys. (**D**) Grand-averaged low-frequency LFPs (1 to 12 Hz) in Mk P across trials and sites for SOLO, JOINT, and OBS_{SOLO}, with characteristic components P1, N1, and P2, around movement time onset (MTon) (SOLO and JOINT; 0 ms). OBS_{SOLO} trials are aligned to cursor motion onset. Dashed curves represent average cursor speed profiles for SOLO and JOINT tasks. Vertical lines indicate mean GO signal times (SOLO, blue; JOINT, red). Gray shading highlights time bins with significant differences (*t* test, $P < 0.001$) between SOLO and JOINT trials. deg, degrees. (**E**) Time-resolved and cross-temporal decoding of action context (SOLO versus JOINT) from PMd LFPs. SVM classification with cross-temporal accuracy matrices (heatmaps) assessed the static/dynamic nature of neural coding. Matrices show accuracy as a function of training and testing times, with neural activity aligned to GO (left) or MTon (right). The white trace shows time-resolved accuracy along the diagonal (training = testing). Higher accuracy peaks occurred with RT onset alignment, with evidence of dynamic, time-specific coding of contextual information during RT. An additional accuracy peak emerged with MT alignment, near action completion at goal achievement. Time bins exceeding chance (permutation test, $P < 0.01$) are marked with thick white bars. Thin lines indicate alignment and mean timing of task events. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$.

reduced (Fig. 5C, left), indicating that coordination-related information emerges at the end of the planning phase, as movement begins.

These results suggest that premotor mechanisms supporting interpersonal coordination encode action context and goodness of coordination at different times in sequential order. Specifically, the former is represented early during the RT period and persists up to movement initiation, whereas coordination level emerges later, around MTon. These findings are consistent with prior studies (*25*, *26*) that converge to a common principle of cortical information processing across areas, with task-critical cognitive computations mediated by rapid, time-specific population dynamics. This principle is especially evident in the PMd before movement initiation (*28*), where preparatory neural states set the initial conditions that shape subsequent movement-related activity (*29*).

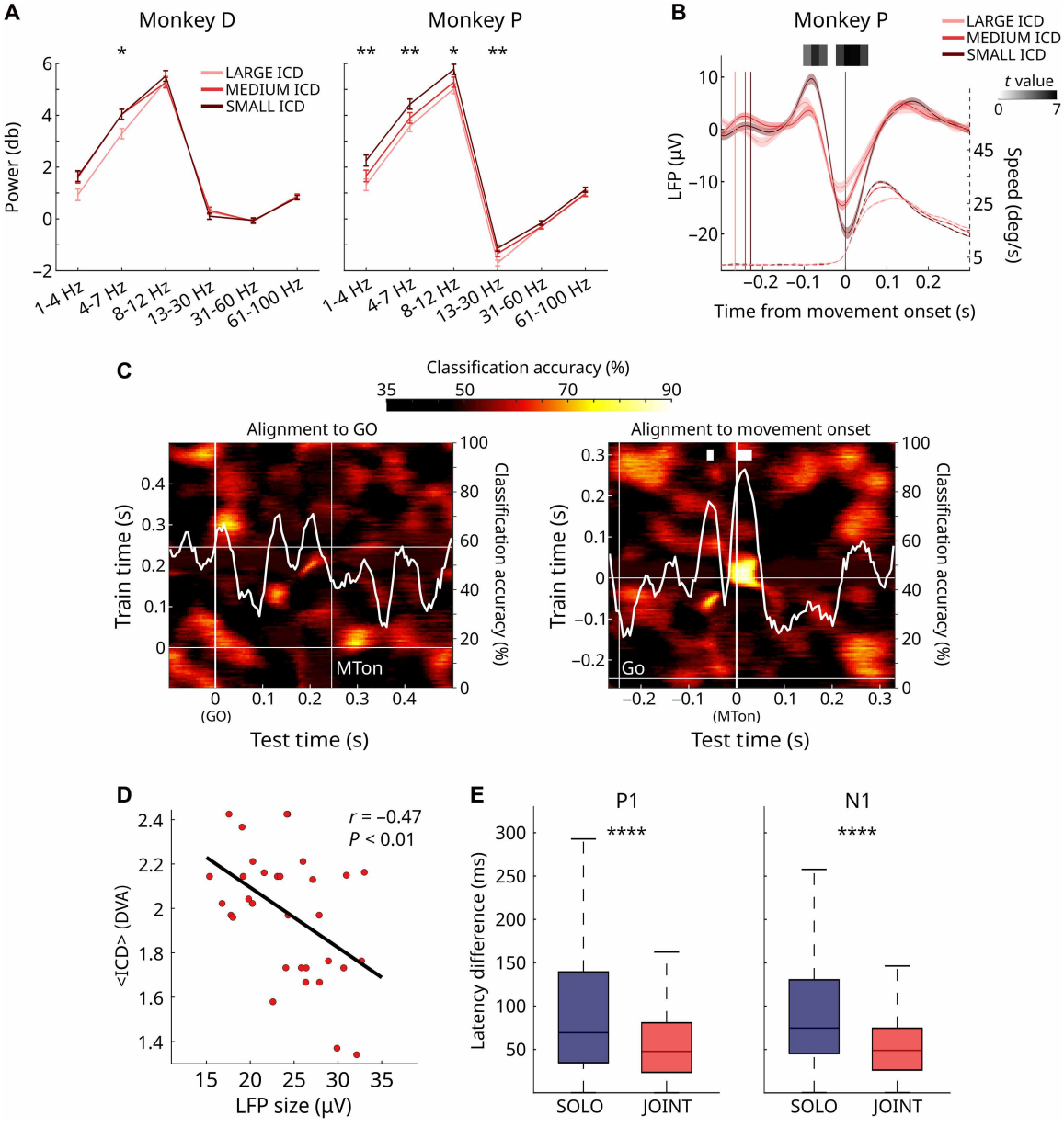

Fig. 5. NHP LFP correlates of interpersonal coordination. (**A**) Mean (± SEM) power in the six frequency bands, computed across trials and recording sites, for three levels of interindividual coordination based on the ICD (<ICD>; large, medium, and small; see Materials and Methods). (**B**) Grand-averaged low-frequency LFPs (1 to 12 Hz) in Mk P, grouped by the same <ICD> categories as in (A). Traces are aligned to movement time onset (MTon, 0 ms), with colored vertical lines indicating the mean GO signal times for each group of trials. (**C**) Decoding of interindividual coordination from premotor LFPs with cross-temporal analysis (see Fig. 4E for details). In contrast to the decoding of action context, higher accuracy peaks emerged when neural data were aligned to MTon, indicating that information about the level of coordination was primarily encoded around movement initiation (conventions and symbols as in Fig. 4E). (**D**) Pearson's correlation between the mean LFP amplitudes during JOINT trials and the <ICD>; here, each data point represents one recording session. LFP amplitudes were computed from the movement direction that elicited the largest response in Mk P (similar results were obtained across all directions, in both monkeys; see the Supplementary Materials). (**E**) Interindividual differences in P1 and N1 latencies were computed in the two task conditions, on a trial-by-trial basis, by coupling signals from all possible combinations of active channel pairs, with each channel recording LFPs from one of the two animals. In the JOINT condition, latency differences were calculated by pairing simultaneously recorded signals from the two brains. In the SOLO condition, signals were paired by randomly selecting recordings from different repetitions, collected by a given channel and in one movement direction. A *t* test revealed a significantly greater synchronization of LFP peaks during JOINT than SOLO trials ($P < 0.0001$). *$P < 0.05$; **$P < 0.01$; ****$P < 0.0001$.

Furthermore, in line with what has been found in humans, we observed a significant correlation between LFP amplitude and the <ICD> magnitude, as shown for Mk P (Fig. 5D).

Last, neural tuning between the two interacting brains was assessed through a trial-by-trial analysis on the timing of P1 and N1 peaks, recorded simultaneously from the two animals. We found that neural responses tended to synchronize during JOINT actions [Fig. 5E; *t* test, SOLO versus JOINT, P1: $t(2699) = 13.62$; N1: $t(2699) = 15.66$; both $P < 0.0001$]. This neural alignment paralleled animals' behavioral coordination in RTs (fig. S4), and it highlights the inextricable link between coagents' temporal alignment and interbrain coupling in JOINT actions.

DISCUSSION

The present study aimed to investigate whether primates' behavior during an interpersonal coordination task better aligns with reactive perceptual and sensorimotor processes, such as entrainment or behavioral mimicry, or it demonstrates proactive adjustments that indicate the contribution of some form of dyadic motor planning. To this end, we examined whether agents adapt their motor behavior based on the fact that they are acting together rather than alone, and we explored the relationship between behavioral adaptations and the optimization of joint performance. Thus, at the behavioral level, we explored (i) whether coagents maximize their interpersonal coordination in JOINT trials, (ii) how they adjust their movement spatiotemporal parameters in the JOINT compared to the SOLO condition, and (iii) how these adjustments relate to the mean intratrial ICD (<ICD>), our index of interpersonal coordination. In addition, we examined the neural correlates of these mechanisms by focusing on the role of the PMd, a crucial brain region involved in motor planning, in facilitating interpersonal coordination.

Overall, our data support the dominance of proactive, rather than purely reactive, motor adjustments during coordination. Although the comparison between the <ICD> calculated in the JOINT and SIMULATED JOINT trials suggests that online corrections may be also involved, all our results point to a prevalent role played by proactive adaptations, as discussed below.

The comparison between the <ICD> measured in real JOINT trials and the one calculated by shuffling the coagents' movement trajectories (within each pair and test session, thus creating a SIMULATED JOINT condition) demonstrated neither epiphenomenal nor stereotypical behavioral adjustments. Agents maximized interpersonal coordination through dynamic minimization of the intratrial <ICD>. If individuals were merely acting simultaneously, guiding "in parallel" their cursors along a stereotypical, more linear trajectory in all JOINT trials, then the <ICD> in SIMULATED JOINT trials would be identical to the one calculated in real JOINT ones. Instead, we found evidence for a reduced <ICD> (i.e., better coordination) in real JOINT trials. This result suggests that coagents actively adjusted their behavior by continuously minimizing the relative distance between their cursors. This was true for both humans and NHPs.

Then, we explored the modulations of spatiotemporal movement parameters that contributed to this trial-by-trial joint optimization. We showed that, although temporal adaptations followed distinct patterns in monkeys and humans (as discussed below), spatial adaptations were, in both species, directed toward remodeling the natural curvilinear trajectories into straighter pathways during JOINT action trials. Such adaptation, which possibly resulted in cursor trajectories that were more predictable for the partner, was nonetheless not stereotypical, but it was adapted trial-by-trial to minimize the <ICD>. This provides strong evidence in favor of a proactive adaptation of movement features to optimize joint performance in both humans and NHPs. JOINT actions were planned and executed in a qualitatively different manner than SOLO actions, prompting individuals to trial-by-trial alter their biomechanically SOLOistic way of performing the task to foster interpersonal coordination.

Last, both intracortical recordings in macaques and BOLD signals in humans indicate that activity in the PMd provides a neural signature of interpersonal coordination. Despite the different temporal resolutions of the two methodologies, the cross-species multiscale approach provides converging evidence for a premotor mechanism engaged before MTon. The precise LFP dynamics recorded in monkeys during the JOINT condition from the homologous PMd regions show an anticipatory modulation ~150 ms before MTon. This supports the interpretation that even the sluggish PMd BOLD signal in humans should reflect proactive motor planning, in line with the well-established association of PMd activity with the preparation phase of motor tasks (*30*, *31*).

Our results also line up with previous evidence on the role of the PMd in guiding joint actions provided by previous fMRI studies in humans (*32*, *33*) and by our single-unit study in monkeys (*13*), which showed the existence of neurons encoding joint performance, also through the integrated representations of self and others' actions.

Notably, in both species, premotor activity during the JOINT condition correlated with the ICD (<ICD>), which can be regarded as an index of the goodness of dyadic performance in the JOINT action. Overall, these findings suggest that the building blocks of interpersonal coordination are shared across humans and macaques and may not necessarily rely on inferential psychological structures [shared/collective intentions (*1*, *2*)] nor on purely reactive visuomotor tracking of the partner's actions. Rather, interpersonal coordination may entail the generation of dyadic motor plans encoded by premotor regions (*11*, *12*) well before movement initiation, as shown by our findings on LFPs. This enables individuals to proactively adapt their behavior to meet interactional demands in real-time. LFPs analysis in monkeys indicated that this proactive coordination process is also reflected in interbrain neural coupling during joint performance.

Furthermore, a decoding analysis confirmed that low-frequency LFP signals carried sufficient information to accurately decode not only contextual information (SOLO versus JOINT) but also the level of interindividual coordination, albeit with distinct temporal patterns. Context-related signals emerged early in the RT period and persisted up to movement initiation, whereas information about the strength of coordination arose later, around the MT onset, marking the critical window for behavioral synchronization in macaques. Testing the stability of this coding through cross-temporal decoding analysis (*26*) revealed that these representations are time specific, following a sequential organization in which an early, dynamic code for action context is followed by a later code for coordination level around MTon. Together, these findings suggest that premotor activity supports a predictive-corrective process, whereby a preparatory dyadic motor plan evolves into execution-phase control essential for achieving effective interindividual coordination. Within this line, a recent dual electroencephalography (EEG) study demonstrated that, under task conditions like those of our study, the dynamic computations characterizing the preparatory phase of joint behavior primarily support the formation of shared collective goals, rather than merely reflecting task complexity (*34*).

The unconstrained whole-brain fMRI analysis performed in human participants revealed that, together with the PMd, the posterior parietal cortex was also more strongly recruited during JOINT compared to SOLO execution. This result is in line with the well-established strong connectivity between these brain regions (*35*, *36*) as well as with their role in the online control of visuomotor behavior (*37*), and it suggests that the whole frontoparietal network is involved during JOINT action planning and execution. The fMRI data analysis also showed that the cerebellum and area MT/V5, regions involved in feedback-based visually guided monitoring and entrainment (*18*), were overall recruited during task execution, yet not to a greater extent during JOINT than SOLO actions. This evidence suggests that feedback-based, visually guided monitoring or entrainment is not sufficient for efficient interpersonal coordination. Instead, JOINT action appears to rely on specific frontoparietal recruitment. Nonetheless, the activity of parietal regions did not show a significant brain-behavioral correlation.

Does the aforementioned evidence provide unequivocal proof of a dyadic motor plan for the form of coordination explored here? Some caution is warranted here. At present, we lack evidence of an "instruction code" that guides behavior on a moment-by-moment basis while simultaneously integrating both the individual's and the coagent's contributions to joint action. The most parsimonious and unequivocal interpretation is that we have evidence of how primates adapt—behaviorally and physiologically—to the demands of cooperation. However, our findings also indicate that such adaptations are largely proactive in nature. Humans and NHPs adapt their behavior in JOINT action since the planning phase (by modulating their RT) and plan their actions in a qualitatively different way (deviating from their stereotypical, more curvilinear trajectory) to trial-by-trial minimize the <ICD>. Viewed in this light, the boundary between the present results and the evidence fulfilling a dyadic motor plan hypothesis (*11*) becomes somewhat narrow.

As there were remarkable points of contact across species, the JOINT behavior of humans and NHP also differed in revealing ways. In all human participants, the RTs became significantly longer in JOINT than SOLO trials. On the contrary, in our NHPs pair, only the more adaptive and responsive monkey (Mk P) changed its RTs in JOINT compared to SOLO trials: It sped up (and not slowed down) its RTs in the JOINT condition, a result that replicates previous findings (*16*, *17*). This behavior, possibly reflecting a "follower" strategy aimed to align with its (faster) partner (Mk D), resulted in better coordination. This difference between humans and monkeys is reminiscent of the developmental pathway shown by children and adults tested in a virtually identical task, as demonstrated in a previous study (*38*). Six- to 8-year-old children (unlike adults but like our monkeys) also resolve to shorter RTs to cope with interpersonal coordination demands. In adult humans, as shown by previous studies based on various social tasks (*39*–*41*), longer RTs may reflect the additional cognitive labor required to plan the action by considering the partner's expected motor behavior. Adult humans (in contrast to monkeys) may also engage in feedback-based monitoring during joint action execution, a function that appears to mature around the age of 7 to 8 (*38*, *42*). Behaviorally, this process is reflected by increased MTs and by the adjustment of one's own MT toward the one observed from the partner in previous JOINT trials, leading to improved dyadic performance. From a neural perspective, it is supported by evidence of broader engagement of the whole frontoparietal network, which is responsible for intentional action control (*37*, *43*).

Our monkeys, instead, showed better interpersonal coordination (i.e., lower <ICD>) when moving faster (i.e., for shorter MTs). The more adaptive Mk P also demonstrated the ability to take the partner's past behavior into account, but only by adapting (speeding up) its RTs accordingly. Regarding the shortening of RTs during JOINT trials, NHPs [like children; see (*38*)] may rely on their implicit learning that initiating their movements faster and along straighter trajectories is more likely to lead to successful (and rewarding) trials when acting together. This suggests that the motor representations that monkeys generate regarding the partner's contribution might be less specified than the ones formed in humans. This interpretation is not based on a single dyad, as comparable behavioral adaptations have been reported in other macaque pairs performing a similar joint action task (*16*, *17*), suggesting that NHPs tend to speed up movement initiation as a coordination smoother (*16*). Overall, these data indicate that monkeys can form a model of their partner's temporal movement features and, in addition, a simpler model of how to make the interaction task easier for themselves when acting with the partner. From both a developmental and evolutionary perspective, adult humans may have refined additional and possibly more adaptive mechanisms that support interaction in increasingly complex social contexts, which nevertheless operate within the same primate premotor building blocks.

Our study also presents some limitations. We list a few here. First, for methodological reasons, we tested the monkeys when they were over-trained on the task at hand. Humans' training was instead limited to a short training session before the fMRI task, but it was sufficient to achieve an overall performance that did not improve throughout the 15-min session performed in the scanner (see Supplementary Text in the Supplementary Materials). Thus, future research might explore what learning processes support the optimization of joint performance and whether monkeys and humans follow similar or divergent patterns in these learning curves. Second, the visuomotor task applied here was kept as simple as possible. Would our findings also hold for more complex forms of motor cooperation? For instance, one might want to explore in the future whether the assignment of specific interacting roles results in similar leader-follower dynamics across species, and the possible engagement of premotor brain activity in these instances. Third, inevitably, monkeys were trained through reward, an aspect that remains to be modeled in both primate species. Last, whereas behavioral adjustments showed remarkable similarity with regard to the spatial features, adaptations in the temporal domain diverged conspicuously. It is worth noting that our animals' MTs were significantly faster than the 3-s threshold imposed for JOINT execution, suggesting an intrinsic emphasis on time efficiency that may extend beyond simply meeting the demands of the task. Future studies might explore whether between-species differences in the temporal features of joint action planning and execution reflect inherent strategic preferences or qualitative differences in how the partner's contribution is represented by each coagent, for instance, by analyzing the agent's reactions to a partner's error that leads to an interaction failure.

MATERIALS AND METHODS

Experimental procedures (humans)

Participants

Fifty healthy volunteers from the University of Milano-Bicocca community were recruited to form 25 gender-matched pairs (13 female pairs and 12 male pairs; age range: 19 to 31; mean age: 22.46 ± 2.45). The sample size was established on the basis of our previous studies (*12*, *44*) showing that a sample size of $N \simeq 20$ provides sufficient power

to detect interaction-specific modulations in the BOLD signal based on a 2 by 2 experimental design and applying a statistical threshold equal to $P_{uncorr} < 0.001$ at the voxel level and $P_{FWERcorr} < 0.05$ at the cluster level, a threshold recommended by Flandin and Friston (*45*). This sample size was also appropriate according to the results of a simulation-based power analysis for mixed models (*46*) run on the behavioral data collected from 22 gender-matched pairs of young healthy human participants (13 female pairs; age range: 20 to 29; mean age: 24.11 ± 1.90) who took part in a pilot behavioral study including the same JOINT execution condition used here. From these data, for each trial, we extracted an index of interpersonal coordination, that is, the mean intratrial <ICD>. This index has been calculated as in the present experiment (see Behavioral data analysis paragraph below). The <ICD> was also trial-by-trial calculated in SIMULATED JOINT data obtained by randomly pairing within each dyad the trajectories of the two cursors in different JOINT trials in each direction (e.g., *n*th trial of participant #1 with *m*th trial of participant #2, etc.; see also Behavioral data analysis paragraph below). By doing so, we obtained <ICD> values that were used to run a simulation-based power analysis for mixed models (*46*) based on the estimate of the effect of condition [JOINT versus SIMULATED (SIM) JOINT] on <ICD> [model: <ICD> ~ 1+ JOINT/SIM + (1+ JOINT/SIM|Participant); intercept estimate: 1.433; fixed-effect estimate: 0.565]. We chose this behavioral effect to run the power analysis because it is the one showing most direct evidence for nonepiphenomenal interpersonal coordination (i.e., a genuine participants' effort to coordinate) in the JOINT action condition. We run the power analysis by hypothesizing a reduction of this effect to the 30% (smallest effect size of interest: 0.1695); the critical *t* value was set at 2. Power estimates were obtained by simulating and then modeling 10,000 independent experiments using the mixedpower package in R (*47*). The power analysis indicated that with *n* equal to the one we selected (25 pairs), the power estimate for the hypothesized effect was 0.80.

Volunteers were eligible if they met the following criteria: (i) right-handedness, as confirmed by the Edinburgh Handedness Inventory (*48*); (ii) no psychiatric nor neurological diseases; and (iii) age between 18 and 32 years. They had normal (or corrected to normal) vision. Students were recruited through a local SONA system (https://milano-bicocca.sona-systems.com) and awarded with formative credits. All participants were naïve as to the purpose of the experiment.

Each pair was formed by participants who did not know each other to exclude the possible confounding effect of previous acquaintance on coordination.

The protocol was approved by the Ethics Committee of the University of Milano-Bicocca (protocol no. 562). Participants gave their written informed consent to take part in the study, in line with the ethical standards of the 1964 Declaration of Helsinki and later amendments. They were debriefed as to the purpose of the study at the end of the experimental procedures. The data collection took place at the RM3T Lab of the Monza San Gerardo Hospital (https://bimr3t.unimib.it). The neuroradiologist screened all the participants who underwent the task in the fMRI scanner to exclude any contraindication to fMRI.

Experimental setup and behavioral task

Per each dyad of participants, only one of them underwent the task in the MRI scanner (Participant A), whereas the other one sat in front of a computer in the adjacent control room (Participant B). See Fig. 1A. During the fMRI scanning, the participants could not see each other but watched the same visual output. Participant A (in the scanner) could see the visual stimuli through a mirror placed over the head coil reflecting an MRI-compatible screen placed behind the scanner (40″; resolution: 3840 by 2160; refresh rate: 120 Hz). The MRI-compatible screen flipped the stimuli right-to-left so that Participant A could see the exact same visual scene as Participant B. Participant B watched the visual stimuli on a full-HD 24″ screen (resolution: 1920×1080; refresh rate: 60 Hz) placed ~50 cm from their eyes.

Both participants wore headphones that dampened the noise of the scanner (Participant A) or reproduced some white noise to facilitate concentration (Participant B).

During the experimental task, each participant controlled a colored visual cursor moving on a black background on the screen by using the thumb of the right hand on an fMRI-compatible Celeritas joystick (Psychology Software Tools; https://pstnet.com/products/celeritas/). Participant A controlled the green cursor and Participant B controlled the blue one.

Stimuli. All stimuli are shown in Fig. 1C. The visuomotor task required participants to move their colored cursors (filled circles having a diameter equal to 0.6 DVA) from the center of the screen to a peripheral target, which was constituted by a white-border circle (3.0 DVA in diameter). The peripheral targets could appear in one of the four possible locations placed at 90 angular intervals and at an eccentricity of 18 DVA in the direction of the four corners of the screen (northwest, northeast, southeast, and southwest). Each trial started with the cursors placed in the middle of the screen, slightly misplaced to avoid overlapping (green cursor: 0.3 DVA on the left; blue cursor: 0.3 DVA on the right). After a variable delay (500 to 1000 ms), a circle (3.0 DVA in diameter) appeared around the cursors ('social cue') and preinstructed participants on the incoming experimental condition (SOLO or JOINT, to be executed or observed; see below). The border of this circle could be green, blue, red, or gray, depending on the incoming experimental condition (see table S1).

At each trial, as soon as one of the cursors started moving, a yellow-border circle (3.0 DVA in diameter) appeared and moved coherently with the cursors: In SOLO conditions, the center of the yellow circle coincided with the location of the moving cursor, whereas, during the JOINT condition, the center of the yellow circle corresponded to the median position between the two partners' cursors. Last, a red and a yellow isosceles triangles (~17 DVA) including an exclamation mark were used as warning stimuli: They were shown to participants when a false start or trial failure occurred (see below).

Experimental conditions. The task included four experimental conditions, indicated by the color of the "social cue" provided at the beginning of each trial, as also summarized in table S1 (see the Supplementary Materials): A green (or blue) cue indicated a SOLO execution (from now on, SOLO) condition for Participant A and a SOLO observation (OBS_{SOLO}) condition for Participant B (and vice versa in the case of a blue cue); a red cue indicated a JOINT execution (from now on, JOINT) condition for both participants; a gray cue indicated a JOINT observation (OBS_{JOINT}) condition for both participants.

During Execution trials, the participants were required to move the yellow circle toward the target shown on a screen, either individually (SOLO condition) or by coordinating their movements with their partner to reach the target together and maintain the cursor as close as possible (JOINT condition). The instantaneous ICD was constantly monitored during JOINT trials (Fig. 1D). If the ICD exceeded the threshold distance of 9 DVA, the trial was aborted. During all observation trials (OBS_{SOLO} and OBS_{JOINT}), the participants

had to rest and observe the screen: If participants moved during an observation trial, a warning stimulus (red triangle) was shown at the end of the trial, and the trial was discarded from the analyses (trial failure). In the OBS_{SOLO} condition, the participant who was not instructed to move observed the partner perform the SOLO task. In the OBS_{JOINT} condition, both partners observed the movements of cursors animated by prerecorded trajectories. These trajectories were randomly extracted from a pool of 224 trajectories of successful trials (see below for definition) executed by four pilot participants (two pairs) playing the JOINT task.

Participants played 32 trials per condition (SOLO, JOINT, OBS_{SOLO}, and OBS_{JOINT}, for a total of 128 trials) in a pseudorandomized order in which we ensured that the same experimental condition was never presented more than three times in a row. Thirty-two "filler" black screens having a randomized duration ranging from 3.5 to 4.5 s were added and used as implicit baselines in fMRI data analyses (see below).

Trial timeline. Figure 1C illustrates the trial timeline, which was identical in all conditions. Each trial started with a black screen on which the cursors were drawn in the center. After a variable delay ranging from 500 to 1000 ms, the "social cue" (green, blue, red, or gray; see Stimuli paragraph above) was shown for a random variable duration (500 to 1000 ms) and then disappeared. At the same time, the peripheral target appeared in one of the four possible locations (see Stimuli paragraph above): This was the GO signal for the participant(s) to start moving. If one participant started moving before the GO signal (false start), a warning stimulus (yellow triangle) was shown, and the trial was aborted and immediately replayed. After the GO signal, as soon as one of the cursors started moving, the yellow circle appeared. As specified above, the center of the yellow circle coincided with the location of the moving cursor in SOLO trials; instead, it corresponded to the median position between the two partners' cursors in JOINT trials. In successful trials, the trial end occurred when the cursor(s) successfully reached the target, that is, when the center of the yellow circle achieved a distance, from the center of the target, lower than 1.5 DVA, to be maintained for at least 100 ms. The trial was aborted (unsuccessful trials) if the target was not reached within 6 s from the GO signal or when the ICD threshold (9 DVA) was exceeded.

At the end of the trial, a filler black screen was shown to ensure that, for each trial, the time interval between the GO signal and the beginning of the subsequent trial was equal to 6 s.

Procedure. Upon arriving in the lab, the participants read through the task instructions and underwent a training phase. During this phase, the participants sat side by side in front of a portable computer while playing the task with standard videogame joysticks to familiarize themselves with it. This training included 8 trials for the SOLO execution condition, 8 OBS_{SOLO} trials, 32 trials for the JOINT execution condition, and 4 OBS_{JOINT} trials, in addition to 4 baseline trials in which only a black screen was shown. Then, Participant A of each pair entered the MRI scan, and a high-resolution T1-weighted MRI scan was acquired. Last, the participants started playing together. Before recording the fMRI data, the participants underwent a second short training (24 trials) aimed at familiarizing the participants with the MRI-compatible joysticks. This second training included 4 SOLO, 4 OBS_{SOLO}, 8 JOINT, 4 OBS_{JOINT} trials, and 4 baselines (black screen).

Software. The task was run, and the behavioral data were recorded using MATLAB R2019b (toolboxes Psychtoolbox and Simulink) (The MathWorks Inc., Natick, MA).

Experimental procedure (monkeys)

Animals and surgery

Two male rhesus monkeys (*M. mulatta*; Mk D, 11.5 kg; Mk P, 11 kg) were used in this study. Animal care, housing, and experimental procedures were performed in conformity with European (Directive 63-2010 EU) and Italian laws (DL. 26/2014) on the use of NHPs in scientific research, approved and authorized by the Italian Ministry of Health. After about 4 months of initial training, animals underwent surgery for headpost implantation. After preanesthesia with ketamine (10 mg/kg, intramuscular), they were anesthetized with a mix of oxygen/isoflurane (1 to 3% to effect). A titanium headpost was implanted on the skull under aseptic conditions. Following surgery, the animals were allowed to fully recover for at least 7 days, under treatment with antibiotic and pain relievers, as from veterinary prescription. Then, after a further week they returned to daily training sessions until they reached a stable performance in all experimental tasks. This occurred about 7 months after the first surgery. At the completion of training, the recording chamber for intracortical multiple-electrode recording (Minimatrix 05, Thomas Recording) was implanted in each animal. For this purpose, each animal was preanesthetized and then anesthetized as above to implant a circular chamber (18 mm in diameter) with four titanium bone screws on the skull, over the frontal lobe of the right hemisphere on Mk D and the left hemisphere on Mk P, to enable recordings from the PMd (F2). The two monkeys underwent all surgical procedures in 2 consecutive days. Recording of neural activity started after 1 week of recovery, under strict veterinary control, and only when both animals were able to perform the tasks as in the immediate presurgery time. Single-unit and LFP recordings were obtained from different layers by using linear arrays of up to five tungsten glass-coated electrodes (0.8 to 2.5 MΩ), with adjacent ones separated by 300 μm. The five-electrode arrays were placed each day in a different site within the recording chambers (everyday similar in the two animals), and the microdrives allowed independent control of electrodes through the dura. Postmortem histological procedures were used to identify the exact location of recording sites.

Experimental setup and behavioral task

Animals were placed together in a darkened, sound-attenuated chamber and seated side by side on two primate chairs (Fig. 1B), in front of a 40-inch monitor (100 Hz; resolution: 800 by 600; 32-bit color depth), located at distance of 150 cm from their eyes. They were trained to control a visual cursor on the screen by applying with their preferred hand a force on an isometric joystick (ATI Industrial Automation, Apex NC), which measured the forces in two dimensions on the horizontal plane with a sampling frequency of 1 kHz. The applied force was proportionally converted into a motion of the cursors on the x and y axes of the vertical plane of the monitor. As for humans, the two cursors, 0.6 DVA in diameter, were presented on a black background and consisted each in a filled circle (blue cursor controlled by Mk D; green cursor controlled by Mk P). During the experiment, the animals used their preferred hand (left Mk D, right Mk P; see Fig. 1B), whereas their other arm was gently restrained. Recording was made from the hemisphere contralateral to the used limb (right for Mk D and left for Mk P; see Fig. 4A). A security distance of 60 cm between animals was always guaranteed to prevent physical contact. The experimental setup was conceived to minimize, during data acquisition, any possible interaction outside the task to avoid potential sources of uncontrolled variables. The orientation of the two chairs prevented direct visual contact.

The behavioral task is also described in (*17*). The behavioral dataset used in this manuscript partially overlaps with the one presented in Lacal and colleagues (*17*); however, the analyzed variables and statistical approach diverge substantially in the two works to address different questions. Monkeys were trained to perform a visuomotor task virtually identical to the one adopted in the fMRI study performed in humans (see above). Animals performed a directional center-out task consisting in moving a visual cursor from a central to a peripheral target by using an isometric joystick. Neural signals were recorded in three of the four experimental conditions described above for human participants, consisting in two executions (SOLO and JOINT) and one observation condition (OBS_{SOLO}), in which the animal observed the outcome of the partner's action during its SOLO trials (see below). Peripheral targets were located in one of the eight possible positions, at 45 angular intervals and at an eccentricity of 8 DVA. The trial type (SOLO, JOINT, or OBS_{SOLO}) was first preinstructed during a dedicated variable time interval, through a specific color cue ("social cue"; see table S2 in the Supplementary Materials), similarly to what was done for the human participants.

The cue about the required action type was followed by a directional cue, consisting in a peripheral target toward which cursors were moved (see the next paragraph for more details). In the SOLO condition, one of the two monkeys was required to move its own cursor from the central to the peripheral target, by applying a dynamic hand force on its own isometric joystick. In the JOINT condition, instead, both monkeys had to coordinate their hand forces in intensity and direction to bring the yellow circle (see above) from the central target to the peripheral location. During the SOLO condition of one animal, the other had to keep its own cursor in the central target while observing the execution of the SOLO trials by its mate. Thus, the SOLO trials of one monkey corresponded to the OBS_{SOLO} trials for the other animal. No specific eye movement requirements were imposed, allowing animals to freely adopt their preferred gaze strategies throughout all task conditions. In the OBS_{SOLO}, although monkeys were free to direct their gaze anywhere on the screen, if anything, the gaze deviation from central fixation was predominantly oriented toward the task-relevant target during the observation of the partner's SOLO trials (fig. S5), suggesting monitoring of the partner's goal even when there was no direct involvement in a coordination task.

Trial timeline. The task began with the presentation of an outlined white circle (2 DVA in diameter) at the center of the screen, and the two cursors close, but outside, said circle. To start the trial, each animal had to move its cursor inside the central target and hold it there for a variable control time [center holding time (CHT); 700 to 1000 ms]. The CHT was followed by a color change of the central circle, with color representing the "social cue" (see table S2), which instructed the monkeys on the future action condition to be performed, i.e., SOLO D (= OBS_{SOLO} for Mk P), SOLO P (= OBS_{SOLO} for Mk D), and JOINT. This preinstruction time lasted 1000 to 1500 ms, and the animals were required to maintain their cursors within the central target. When a peripheral target (white circle; 2 DVA in diameter) appeared (GO signal) in one of the eight possible positions, each animal had to move its cursor toward the peripheral location within an RT of 100 to 800 ms, individually or together, or it had to keep its cursor in the central target, as previously instructed by the color of the "social cue." Once the peripheral target was reached, the animal was required to maintain the cursor in that position for a variable short target holding time (THT; 100 to 200 ms): This prevented the animal from hitting the circle and continuing to move the cursor beyond its edges. At the end of a successful trial, animals were rewarded with a constant amount of liquid reward (0.5 ml). During the SOLO condition of one animal, the partner who observed (OBS_{SOLO}) had to keep holding its cursor inside the central target until the trial ended to gain the same liquid reward, irrespective of the partner's action outcome (success or error). In JOINT action, for the entire duration of their cursors' motion, monkeys were asked to maintain the cursors within a maximum ICD of 5 DVA, which was marked by the yellow circle encompassing the two cursors. As for human participants, exceeding the ICD threshold led to trial abortion and none of the two monkeys obtained the reward. Within each session, trials corresponding to different action conditions (SOLO, OBS_{SOLO}, and JOINT) and different directions were presented in an intermingled manner and pseudorandomized within a session of a minimum of 192 successful trials (three conditions x eight directions x eight replications).

In monkeys, trials were classified as successful when temporal constraints (MT < 3000 ms) and spatial constraints (in JOINT trials only, ICD < 5 DVA) were fulfilled.

Software. The National Institutes of Health (NIH)–funded software package REX was used to control stimuli presentation and to collect behavioral events, force-related data, and eye movements, using custom scripts. Eye data were sampled (at 220 Hz) through an infrared oculometer (Arrington Research) and stored together with joystick force signal and key events, which were sampled at 1 kHz.

Behavioral data analysis (humans and monkeys)

Both human and monkey behavioral data were computed using MATLAB (The MathWorks Inc., Natick, MA) and analyzed with JAMOVI v.2.3.28 (*49*), retrieved from https://jamovi.org. Whenever possible, we used the same behavioral measures and statistical tests for analyses of human and monkey data.

We analyzed the behavioral performance focusing on successful SOLO and JOINT trials only. As reported above, trials were classified as successful when temporal constraints (MT < 6 s for humans and 800 ms for monkeys) and spatial constraints (in JOINT trials only, ICD < 9 DVA in humans and 5 DVA in monkeys) were fulfilled. From these trials, we extracted the following variables:

1) RT, i.e., the time interval between the appearance on the screen of the peripheral target (GO signal) and the instant when the participant(s)'s cursor(s) started moving [movement time onset (MTon)]; for monkeys, the MTon was identified as the instant at which the cursor's velocity exceeded, by 3 SDs and for at least 90 ms, the average velocity signal, measured in the interval spanning from 50 ms before to 50 ms after the presentation of the peripheral target; for humans, the MTon was set as the first frame in which the cursor position changed coordinate on one of the two axis after the presentation of the peripheral target;

2) MT, i.e., the time interval between the instant when the participant(s)'s cursor(s) started moving (MTon; see above) and the instant when the peripheral target was reached (i.e., the yellow circle reached a distance of <1.5 DVA from the center of the peripheral target and this proximity was maintained for 100 ms);

3) TD, an index taken as a measure of the cursor trajectory distortion compared with an "ideal linear trajectory" running from the starting position of the cursor to the center of the peripheral target; to calculate this variable, we divided the cursor(s)' positions sampled during movement execution [i.e., the cursor(s)' trajectory] in 10 bins and calculated the mean position within each bin; then, we measured the Euclidean distance of each of these 10 positions with respect to the line linking the starting position of the cursor to the center of

the peripheral target, and we summed up these 10 values to obtain an overall measure of the distance of these checkpoints from the "ideal linear trajectory";

4) Mean intratrial ICD (<ICD>), i.e., the mean of the instantaneous Euclidean distances (ICDs) between cursors, during JOINT trials, computed throughout the entire duration of the cursor MT, from the instant when the fastest cursor started moving toward the peripheral target to the instant when the target was reached, i.e., the yellow circle reached a distance of <1.5 DVA from the center of the peripheral target and this proximity was maintained for 100 ms; the <ICD> was also trial-by-trial calculated in SIMULATED JOINT data obtained by randomly pairing within each dyad the trajectories of the two cursors in different JOINT trials in each direction (e.g., *n*th trial of participant #1 with *m*th trial of participant #2, etc.).

For human data, RTs, MTs, TDs, and <ICD> were preprocessed to exclude, within each participant and each experimental condition, trials showing outlier values in at least one dependent variable. We considered as outliers the values that fell 3 SDs above or below the individual mean for each experimental condition. According to this criterion, we excluded a total of 32 successful trials (2.2%) of those executed by human participants playing as Participant A (i.e., in fMRI) and 29 successful trials (2.0%) of those executed by human participants playing as Participant B (i.e., outside the MRI scanner). From the monkey dataset, we did not exclude any successful trial.

Behavioral analysis #1: Evidence for interpersonal coordination in JOINT action

We first explored whether the results of the JOINT task provided evidence for genuine interpersonal coordination in both humans and monkeys. To do so, we compared the <ICD> recorded in the JOINT condition with the one measured in SIMULATED JOINT trials, created by randomly pairing the trajectories of the two cursors of the JOINT trials within each dyad (e.g., *n*th trial of participant #1 with *m*th trial of participant #2, etc.). To calculate SIMULATED JOINT trials, we paired only successful trials in which the same target direction was reached, and the trajectories were aligned on MTon. To assess whether the <ICD> was lower (indicating better interpersonal coordination) in the real JOINT than SIMULATED JOINT trials, we ran a linear mixed model on the <ICD> data, considering as fixed effect the Condition (real JOINT versus SIMULATED (SIM) JOINT). Concerning the random effect structure, by-subjects (or by-session, in monkeys) random intercepts and slopes were included to account for between-subject (or between-session) variability. The linear mixed models applied in this analysis were the following (Table 1; see "Comparison between JOINT and SIM data")

$$\text{Human model: } <\text{ICD}> \sim 1 + \text{JOINT/SIM} + \left(1 + \text{JOINT/SIM} \mid \text{Participant}\right)$$

$$\text{Monkey model: } <\text{ICD}> \sim 1 + \text{JOINT/SIM} + (1 + \text{JOINT/SIM} \mid \text{Session})$$

Behavioral analysis #2: Behavioral adjustments in JOINT action

Then, to assess the behavioral adaptations applied by participants in the real JOINT (as compared to the SOLO) trial types, we ran a series of linear mixed models on the dependent variables (DVs) listed above (i.e., RTs, MTs, and TDs), considering as fixed effect the Condition (SOLO versus JOINT). Concerning the random effect structure, by-subjects (or by-session, in monkeys) random intercepts and slopes were included to account for between-subject (or between-session) variability. When convergence issues emerged, the random slope was removed, whereas the random intercept was always included. The linear mixed models applied in this analysis were the following (Table 1; see "Comparison between the SOLO and JOINT condition")

$$\text{Human model: DV} \sim 1 + \text{SOLO/JOINT} + \left(1 + \text{SOLO/JOINT} \mid \text{Participant}\right)$$

$$\text{Monkey model: DV} \sim 1 + \text{SOLO/JOINT} + (1 + \text{SOLO/JOINT} \mid \text{Session})$$

Behavioral analysis #3: Behavioral adjustments leading to better interpersonal coordination in JOINT action

We explored which of the behavioral adaptations that emerged in analysis #2 promoted better interpersonal coordination (in terms of lower <ICD>). To do so, we first calculated normalized JOINT behavioral indexes (*J,norm*) by calculating normalized deviations from the mean value recorded from the same participant in the SOLO condition ($\langle s \rangle$)

$$J, norm = \frac{\left(j_j - \langle s \rangle\right)}{\left[\frac{j_j + \langle s \rangle}{2}\right]}$$

where j_j is the value of the dependent variable measured during the execution of each JOINT trial and $\langle s \rangle$ is the individual mean value of the dependent variable collected during the execution of SOLO trials.

Then, we ran a series of linear mixed models to test whether these *J,norm* values (calculated for RTs, MTs, and TDs), added to separate linear mixed models as independent variables, could predict trial-by-trial changes in the <ICD>. In this model, by-subjects (in humans) and by-session (in monkeys) random intercepts and slopes were included to account for between-subject (in humans) or between-session (in monkeys) variability. When convergence issues emerged, the random slope was removed, whereas the random intercept was always included. The linear mixed models applied in this analysis were the following [Table 1; see "Regressions analyses (RTj,norm, MTj,norm, TDj,norm versus ICD)"]

$$\text{Human model: } <\text{ICD}> \sim 1 + \text{DV}_{\text{J,norm}} + \left(1 + \text{DV}_{\text{J,norm}} \mid \text{Participant}\right)$$

$$\text{Monkey model: } <\text{ICD}> \sim 1 + \text{DV}_{\text{J,norm}} + \left(1 + \text{DV}_{\text{J,norm}} \mid \text{Session}\right)$$

Behavioral analysis #4: Adaptations to the history of partner's movement patterns

We also explored whether reciprocal behavioral adjustments to the partner's behavior might play a role. To do so, we calculated the absolute value of the difference between each participant's RT or MT in a given trial and the partner's average RT or MT calculated from all the previous JOINT trials, as follows

$$\Delta \text{RT}_{\text{a,<b>}} = \left|\text{RT}_{\text{a}} - \langle \text{RT}_{\text{b}} \rangle\right|$$

$$\Delta \text{MT}_{\text{a,<b>}} = \left|\text{MT}_{\text{a}} - \langle \text{MT}_{\text{b}} \rangle\right|$$

where RT_a (MT_a) refers to the RT (MT) recorded from one agent in a given trial, and $\langle RT_b \rangle$ ($\langle MT_b \rangle$)refers to the partner's mean RT (MT) calculated on all previous JOINT trials. The $\Delta RT_{a,<b>}$ and $\Delta MT_{a,<b>}$ delta values thus indicate how much one agent's RT or MT diverges from (or converges on) what the agent might have learned about the partner's behavior. This analysis was restricted to the temporal parameters (RT and MT) because these are the ones showing signs of reciprocal adaptations in monkeys in analysis #2.

Using a series of linear mixed models, we tested whether $\Delta RT_{a,<b>}$ ($\Delta MT_{a,<b>}$) values predict trial-by-trial changes in the <ICD>, our behavioral index of interpersonal coordination. In this model, by-subjects (in humans) and by-session (in monkeys) random intercepts and slopes were included to account for between-subjects (in humans) or between-sessions (in monkeys) variability. When convergence issues emerged, the random slope was removed, whereas the random intercept was always included. The linear mixed models applied in this analysis were the following:

Human models

$$<\mathrm{ICD}> \sim 1 + \Delta RT_{a,<b>} + \left(1 + \Delta RT_{a,<b>} | \mathrm{Participant}\right)$$

$$<\mathrm{ICD}> \sim 1 + \Delta MT_{a,<b>} + \left(1 + \Delta MT_{a,<b>} | \mathrm{Participant}\right)$$

Monkey models

$$<\mathrm{ICD}> \sim 1 + \Delta RT_{a,<b>} + \left(1 + \Delta RT_{a,<b>} | \mathrm{Session}\right)$$

$$<\mathrm{ICD}> \sim 1 + \Delta MT_{a,<b>} + \left(1 + \Delta MT_{a,<b>} | \mathrm{Session}\right)$$

In all linear mixed models, the Satterthwaite method for degrees of freedom and the Wald method for CIs were applied. All tests of significance were based on an α level of 0.05.

For human participants, the behavioral results of the participants playing as Participant A (i.e., in the MRI scanner) are reported in the main text, whereas the behavioral results of the participants playing as Participant B, which fully replicated the ones emerged from the analysis of Participant A data, are reported as Supplementary Text in the Supplementary Materials.

As a control analysis, we also examined the potential influence of learning throughout the monkeys' experimental sessions and across the trials of the single experimental session involving each pair of human participants. As described in the Supplementary Materials, no effect of session (for monkeys) or trial number (for human participants) was found. This indicates that learning effects could not account for our results.

Human neuroimaging data analysis

Data acquisition

Whole-brain functional T2-weighted images were acquired using a Philips Ingenia 3.0T CX scanner (Philips S.p.A., Milan, Italy), equipped with gradient-echo echo-planar imaging (EPI; scan parameters: repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, 35 transversal slices, descending not interleaved acquisition, 4-mm slice thickness, with no interslice gap, flip angle (FA) = 75°, field of view = 240 mm, and matrix size = 80 by 80). Each voxel was 3 mm by 3 mm by 4 mm in dimensions. The total number of scans acquired varied between subjects because, whenever the participants made a false start (i.e., moved before the GO signal), the trial was aborted and immediately restarted, thus making data acquisition slightly longer. On average, 564 (±17) volumes were acquired per each subject. The fMRI session lasted about 20 min. The first five volumes recorded from each functional run were removed to allow for steady-state tissue magnetization. Magnetization-prepared rapid acquisition with gradient echo (MPRAGE) high-resolution T1-weighted structural images were also acquired [sagittal three-dimensional (3D) T1 TFE: acquisition matrix = 256 by 256, voxel size = 1 mm by 1 mm by 1 mm, Repetition Time (TR) = 8.1 ms, echo time (TE) = 3.7 ms, and flip angle (FA) = 8]. This structural scan was acquired before the fMRI session.

Preprocessing

We reconstructed images, and then we converted raw data from the DICOM to the NIfTI format using MRIConvert software (https://softpedia.com/get/Science-CAD/MRIConvert.shtml). We adopted MATLAB R2019b (MathWorks) using Statistical Parametric Mapping software (SPM12, Welcome Department of Imaging Neuroscience) to perform the preprocessing of the data and univariate data analyses. We first realigned and unwarped fMRI scans to reduce the influence of head movements during the scan session and the effect of magnetic field distortions. The unwarped images were coregistered with the 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 of the data. Last, deformation fields used for T1 segmentation were applied to normalize the coregistered functional scans. We then interpolated the data matrix to produce voxels with a dimension of 2 mm by 2 mm by 2 mm. The normalized scans were smoothed using a Gaussian filter of 10 mm by 10 mm by 10 mm to improve the signal-to-noise ratio and permit the application of the family-wise error rate (FWER) correction at the cluster level in the second-level (group-level) univariate analyses, as suggested by Flandin and Friston (*45*).

The data of each participant were screened using Artifact Detection Tools (ART; Whitfield-Gabrieli; http://nitrc.org/projects/artifact_detect) to identify outlier scans in global signal and movement. We marked time-points as outliers if scan-to-scan variations in the global signal overcame 3 SDs from the mean and if the compounded measure of movement parameters exceeded the 1-mm scan-to-scan movement average. The outlier time points were accounted for by adding regressors of no interest in the single-subject analyses. Excluded volumes were 1.96 ± 1.39%. As subject exclusion criterion was >20% of outlier time points in each experimental condition, no participant was excluded due to excessive motion.

First-level (single-subject) analysis

In the first-level analyses, conducted at the individual level, the BOLD signal associated with each experimental condition (SOLO, OBS_{SOLO}, JOINT, and OBS_{JOINT}) was analyzed by convolution with a canonical hemodynamic response function (*50*). Global differences in the fMRI signal were removed from all voxels with grand-mean scaling. The time series was high-pass filtered at 128 s and prewhitened by means of an autoregressive model AR (*1*). The first-level analysis implied a fixed-effect analysis in which condition-specific effects were calculated.

In an event-related design, we used the same onset as the one used to calculate the LFP in monkeys (see below), that is, 150 ms before MTon. To ensure that the same time window was considered in SOLO and JOINT trials, we characterized the BOLD response corresponding to the first part of the movement, and the event duration was set to 0. Separate regressors were also included to consider experimental confounds, including unsuccessful trials and the realignment parameters calculated in the preprocessing step.

We then estimated neurofunctional activations associated with each experimental condition (SOLO, OBS_{SOLO}, JOINT, and OBS_{JOINT}) by calculating the simple effects of each condition, i.e., the voxel-by-voxel difference between brain activity during the task and during the observation of a black screen (implicit baseline). These contrast images entered the group-level analyses. To estimate the relationship between the interaction effect that emerged at the group level (see below) and the behavioral performance (indexed by the <ICD>; see the description of the regression analysis below), we also calculated subject-by-subject the linear contrast measuring the brain activations that show a stronger "Execution > Observation" effect in the JOINT as compared to the SOLO condition, as follows: [(JOINT > OBS_{JOINT}) > (SOLO > OBS_{SOLO})].

Second-level (group) analysis

At the group level, we first run a full-factorial analysis of variance (ANOVA), considering SOLO versus JOINT and Execution versus Observation as within-subject factors. We aimed to explore what neurofunctional resources are needed to execute the JOINT as compared to the SOLO condition while controlling for differences in the visual stimulation (i.e., controlled for by the OBS_{JOINT} and OBS_{SOLO} conditions, respectively). Thus, we calculated (i) the main effect of "Execution > Observation" and (ii) the interaction effects. These results are reported in tables S4 and S5. The simple effects of "Execution > Observation" were also calculated separately for the SOLO and JOINT conditions for visualization purposes (see Fig. 3A). These analyses were performed at the whole-brain level and applied a threshold of $P_{uncorr} < 0.001$ at the voxel level and $P_{FWERcorr} < 0.05$ at the cluster level.

The results showed a significant interaction effect [(JOINT > OBS_{JOINT}) > (SOLO > OBS_{SOLO})], demonstrating that bilateral dorsal premotor and posterior parietal brain regions are more strongly recruited during the JOINT as compared to the SOLO condition (see the Results section and table S5). To explore whether these activations were associated with better interpersonal coordination, we ran a regression analysis on the related activation map by adding the individual mean <ICD> of each pair as predictor and the single-subject interaction effect [(JOINT > OBS_{JOINT}) > (SOLO > OBS_{SOLO})] as the dependent variable. These latter results are reported at $P_{uncorr} < 0.05$ at the voxel level in Fig. 3B.

Monkey electrophysiological recordings and data analysis

Neural recording

We recorded neural activity using two separate five-channel multiple-electrode arrays for extracellular recording (Thomas Recording) simultaneously from the brains of the two monkeys by adjusting the depth of each electrode (quartz-insulated platinum-tungsten fibers; 80 μm in diameter and 0.8- to 2.5-MΩ impedance), so as to isolate the single-unit activity. Electrodes were equidistantly disposed in a linear array with interelectrode distance of 0.3 mm and were guided through the intact dura into the cortical tissue (*51*) through a remote controller. The raw neural signal was recorded at ~24-kHz sampling frequency and amplified, digitized, and optically transmitted to a digital signal processing unit (RA16PA-RX5-2, Tucker-Davis Technologies) where it was stored together with the key behavioral events.

LFPs were recorded in 12 daily sessions, from 27 sites in Mk D and 32 sites in Mk P, with adjacent sites separated by at least 300 μm. To isolate LFP signals, offline preprocessing of neural data was implemented. First, the raw signals of each channel were centered at zero by subtracting their average over time for the entire session, and power line noise was removed by using a notch filter (zero-phase Butterworth, second order, cutoff frequency = 50 Hz). LFPs were obtained by band-pass filtering the signal at 1 to 300 Hz (zero-phase Butterworth, fourth order) and downsampling to ≈1 kHz.

Time-frequency domain. For time-frequency decomposition (Fig. 4B), we adopted the wavelet analysis and instantaneous power definition, reported in (*52*, *53*). In brief, LFP time series of each trial were convolved with a set of complex Morlet wavelets defined as

$$W(f,t) = e^{i2\pi tf} e^{\frac{-t^2}{2\sigma^2}}$$

where t is the time, f is the frequency (which increased from 1 to 100 Hz in 100 logarithmically spaced steps), and σ defines the width of the wavelet in the time domain, set to $7/2\pi f$ (*52*–*54*). For our purposes, the value of 7 provided an adequate trade-off between temporal and frequency resolution. Other values in the range from 4 to 10 provided similar time-frequency maps.

After convolution, instantaneous power of the resulting complex signal $Z(t)$ was defined as follows

$$p(t) = real[Z(t)]^2 + imag[Z(t)]^2$$

Power was normalized with respect to a baseline power ($power_b$) and converted to a decibel (dB) scale as follows

$$p_{norm} = 10 * \log_{10}\left[\frac{p(t)}{\langle power_b \rangle}\right]$$

with baseline defined as the neural activity recorded during the control epoch CHT.

As illustrated in the example in Fig. 4B, low-frequency bands below 12 Hz dominated the period from 250 ms before to 250 ms after MTon, during the execution of both SOLO and JOINT action tasks. To more precisely identify the predominant bands within this time window across recording sites, we measured power variations across different frequency bands.

Frequency domain. We first proceeded with static power spectral analysis. We evaluated frequency-dependent variations across tasks, by subdividing the frequency range of 1 to 100 Hz into six bands: Delta (1 to 4 Hz), Theta (4 to 7 Hz), Alpha (8 to 12 Hz), Beta (13 to 30 Hz), Low Gamma (31 to 60 Hz), and High Gamma (61 to 100 Hz), and computed the relative normalized powers for each channel in the three task conditions (SOLO, JOINT, and OBS_{SOLO}), aligning the LFPs at the MTon. For each trial, the power of each band was extracted in the time window [−250 ms, 250 ms] around the MTon, using the Matlab function "bandpower." For each frequency band, significant differences of the respective powers across conditions (SOLO, JOINT, and OBS_{SOLO}) were assessed through a one-way ANOVA ($P < 0.05$; factor: task condition), followed by post hoc pairwise multiple comparisons with Bonferroni's correction (Fig. 4C).

A similar approach was used to compare the LFPs powers in each frequency band, in relation to different values of <ICD>. To this aim, all trials were first subdivided into three equally populated groups (tertiles) based on the size of their <ICD> (small, medium, and large), computed separately for each session (Fig. 5A). Then, for each band, a one-way ANOVA ($P < 0.05$) was performed (factor: ICD size), followed by post hoc pairwise multiple comparisons with Bonferroni's correction, to detect power modulation in relation to <ICD>.

Time domain. Low-frequency LFPs were obtained for each trial by band-pass filtering the signals (zero-phase Butterworth, fourth order) at 1 to 12 Hz and down sampling them to ≈1 kHz. Signals were then normalized by subtracting the baseline, computed as the mean signal recorded during the control time CHT (see Animal experimental

setup and behavioral tasks). The main scope was to evaluate the overall changes in signal amplitude associated to different tasks conditions, through a grand-average analysis of low-frequency LFPs, by pooling all trials from different channels and different sessions.

To this aim, we first averaged the signals recorded from a single channel within one recording session across different trials (minimum 64 = 8 directions × 8 repetitions at least), after aligning the LFPs to MTon. Before averaging the signals across channels and sessions, we analyzed the LFP waveforms of each channel as a function of recording depth across cortical layers. LFP properties are known to be depth dependent (*55*–*58*); thus, electrode positions were considered to accurately describe LFP peak components near active neuronal populations. The dipolar nature of transmembrane currents underlying neural activity generates negative and positive fields, leading to polarity inversions across cortical layers, with minimal amplitudes at dipole centers. This polarity reversal was confirmed in Mk D data (fig. S2A), where recordings spanned a wider range of cortical depths (about 2 mm). In contrast, in Mk P this variability was less pronounced, with signals collected from more superficial layers only (fig. S2B). Notably, the overall pattern of superficial low-frequency LFPs recorded from Mk D closely resembled the signals observed in Mk P, particularly in the sequence of their negative and positive components.

Therefore, for the purpose of analyzing the overall modulation of LFP amplitudes in the time domain at the population level, we averaged signals that shared similar waveform characteristics, pooling for each monkey different channels and different sessions. Specifically, in Mk P we included all recorded signals, whereas for Mk D, we included only LFPs recorded from superficial layers. This approach was motivated by three main reasons. First, it allowed a direct comparison of LFP waveforms between the two animals. Second, this selection captures the main components of evoked potentials, as previously described in the PMd around MTon during motor tasks, with a positivity (P1) before MTon, a negativity around movement onset (N1), and a slower positive and negative (P2 and N2) component during and after movement (*19*, *21*, *59*). Third, decoding of movement motor goals is best at superficial depths and degrades substantially for signals recorded at cortical sites deeper than 1 mm (*60*). On the grand-averaged LFPs, to assess time-resolved differences in evoked potentials between the SOLO and JOINT conditions, we conducted a sliding-window *t* test (window width: 40 ms; step size: 20 ms; $P < 0.001$; Fig. 4D).

Similar procedures were followed to compare LFPs recorded during JOINT trial and associated with different values of <ICD>. For each session, trials were subdivided into three equally populated groups based on their <ICD>, which was categorized in tertiles (Small, Medium, and Large; Fig. 5B). For the low-frequencies band, because of the directionality of these signals (*20*), we also defined for each channel its preferred direction during JOINT action as the target direction with the largest trial-averaged LFP amplitude. We also evaluated the correlation between LFP amplitude and <ICD> through a Pearson's correlation for each channel at its preferred direction (Fig. 5D) or considering for each channel the eight movement directions separately (fig. S1).

To relate the signal amplitude during JOINT trials to the corresponding <ICD> (Fig. 5D), we quantified the overall amplitude of evoked potentials around MTon using the root mean square (RMS) of the LFP signal (*20*).

Thus, for a given channel *ch* and trial *i,* the RMS was computed as

$$LFPsize_{ch,i} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left[LFP_{ch,i}(t)\right]^2}$$

where *t* denotes time bins within the analysis window spanning from −150 to +250 ms relative to MTon. This interval corresponds to the period in which the three main LFP components (P1, N1, and P2) are typically observed (*21*).

Cross-temporal decoding analysis. To determine whether task-related variables could be decoded from LFP activity, we performed classification analyses using an SVM implemented in the Neural Decoding Toolbox (*61*) together with the library LIBSVM (*62*). Two separate analyses were performed aimed at decoding: (i) action context (SOLO versus JOINT) and (ii) degree of interindividual coordination (Large versus Small ICD, defined as described above). For each channel and trial, LFP signals were averaged into nonoverlapping 5-ms bins. The decoder was run on the low frequency LFPs (1 to 12 Hz) using time windows aligned either to the GO signal (−0.1 to 0.5 s) or to the MT onset (−0.25 to 0.3 s). Trials were assigned labels corresponding either to the task condition (SOLO/JOINT) or to the coordination level (Large/Small ICD). To reduce variability and improve robustness of the decoding results, data were grouped into trial sets, with each set consisting of the average of four consecutive trials of the same type. Then we used a 10- and 4-fold cross-validation for the task condition and the coordination level, respectively. The classifier was trained using data from all but one of the sets, and it was tested on data from the remaining set. This procedure was repeated until each set had served as test data (as many times as the number of splits), and the prediction results were averaged across repetitions. To avoid bias from high-activity channels, *z*-score normalization was applied during cross-validation, ensuring comparable contributions of all channels in both training and test trials. This entire decoding procedure was repeated 50 times, each using different training/test splits, and the results were then averaged. To assess significance, we performed a permutation test in which trial labels were randomly shuffled and the complete decoding procedure was repeated 100 times, generating a null distribution of accuracies. Decoding performance was deemed significant in each bin if it exceeded shuffled values ($P < 0.01$). Last, to evaluate the temporal evolution of information coding, we performed a cross-temporal decoding by training and testing classifiers on all possible combinations of time bins, yielding an accuracy matrix with diagonal values reflecting training and testing on the same time bin and off-diagonal values reflecting generalization across time. Further details on the procedure and the analysis code are available at https://readout.info/.

Temporal coupling of neural signals. Last, we analyzed the timing of the P1 and N1 components, which are typical of low-frequency LFPs. Specifically, we first determined trial-by-trial the N1 latency, as the time of occurrence, relative to the GO cue, of the signal's absolute minimum during the RT and MT. The P1 latency was then identified as the time of the absolute maximum occurring within the preceding 150 ms, a time window selected to ensure the capture of this component (*21*). For direct interanimal comparison and methodological consistency with prior analyses, latency calculations were also restricted to LFPs recorded at comparable cortical depths. Thus, all recording sites were included for Mk P, whereas analyses for Mk D were limited to superficial recording depths.

A trial-by-trial analysis was conducted to assess the temporal coupling of neural signals between the two interacting brains. For each JOINT trial, we measured the latency differences of the P1 and N1 components across all possible pairs of simultaneously recorded channels, with one channel from each monkey. As a control, analogous latency differences were computed using SOLO trials by pairing

signals, again from one channel per monkey, that were randomly selected from the eight repetitions of the same movement direction (Fig. 5E and fig. S3).

We further investigated the relationship between interbrain synchronization and dyadic behavior by analyzing the correlation between P1 and N1 latency differences during JOINT trials and (i) the differences in the monkeys' RTs during the JOINT condition (fig. S4A) and (ii) the ICD (fig. S4B). In this analysis, each data point corresponds to one session in a given movement direction, with signals averaged across channels and repetitions.

Supplementary Materials

This PDF file includes:

Supplementary Text
Tables S1 to S6
Figs. S1 to S5
References

REFERENCES

1. J. R. Searle, "Collective intentions and actions" in *Intentions in Communication*, P. R. Cohen, J. Morgan, M. Pollack, Eds. (MIT Press, 1990), pp. 401–416.
2. M. E. Bratman, Shared Intention. *Ethics* **104**, 97–113 (1993).
3. M. Tomasello, *A Natural History of Human Morality* (Harvard Univ. Press, 2016).
4. E. Herrmann, J. Call, M. V. Hernández-Lloreda, B. Hare, M. Tomasello, Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science* **317**, 1360–1366 (2007).
5. D. Papadopoulos, Shared intentionality in nonhuman great apes: A normative model. *Rev. Philos. Psychol.* **14**, 1125–1145 (2023).
6. A. P. Melis, The evolutionary roots of human collaboration: Coordination and sharing of resources. *Ann. N. Y. Acad. Sci.* **1299**, 68–76 (2013).
7. M. J. Richardson, K. L. Marsh, R. W. Isenhower, J. R. L. Goodman, R. C. Schmidt, Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Hum. Mov. Sci.* **26**, 867–891 (2007).
8. N. Sebanz, G. Knoblich, W. Prinz, Representing others' actions: Just like one's own? *Cognition* **88**, 11–21 (2003).
9. P. E. Keller, G. Novembre, M. J. Hove, Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philos. Trans. R. Soc. London Ser. B Biol. Sci.* **369**, 20130394 (2014).
10. M. K. Schweinfurth, D. B. Baldridge, K. Finnerty, J. Call, G. K. Knoblich, Inter-individual coordination in walking chimpanzees. *Curr. Biol.* **32**, 5138–5143.e3 (2022).
11. L. M. Sacheli, E. Arcangeli, E. Paulesu, Evidence for a dyadic motor plan in joint action. *Sci. Rep.* **8**, 5027 (2018).
12. L. M. Sacheli, C. Verga, E. Arcangeli, G. Banfi, M. Tettamanti, E. Paulesu, How task interactivity shapes action observation. *Cereb. Cortex* **29**, 5302–5314 (2019).
13. S. Ferrari-Toniolo, F. Visco-Comandini, A. Battaglia-Mayer, Two brains in action: Joint-action coding in the primate frontal cortex. *J. Neurosci.* **39**, 3514–3528 (2019).
14. N. K. Logothetis, J. Pauls, M. Augath, T. Trinath, A. Oeltermann, Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
15. N. K. Logothetis, B. A. Wandell, Interpreting the BOLD signal. *Annu. Rev. Physiol.* **66**, 735–769 (2004).
16. F. Visco-Comandini, S. Ferrari-Toniolo, E. Satta, O. Papazachariadis, R. Gupta, L. E. Nalbant, A. Battaglia-Mayer, Do non-human primates cooperate? Evidences of motor coordination during a joint action task in macaque monkeys. *Cortex* **70**, 115–127 (2015).
17. I. Lacal, L. Babicola, R. Caminiti, S. Ferrari-Toniolo, A. Schito, L. E. Nalbant, R. K. Gupta, A. Battaglia-Mayer, Evidence for a we-representation in monkeys when acting together. *Cortex* **149**, 123–136 (2022).
18. S. Uccelli, L. M. Sacheli, C. De Bernardi, F. Devoto, G. Basso, E. Paulesu, Neural mechanisms for visuomotor co-regulation in social synchronization. bioRxiv 2025.05.21.655276 [Preprint] (2025); https://doi.org/10.1101/2025.05.21.655276.
19. O. Donchin, A. Gribova, O. Steinberg, H. Bergman, S. Cardoso de Oliveira, E. Vaadia, Local field potentials related to bimanual movements in the primary and supplementary motor cortices. *Exp. Brain Res.* **140**, 46–55 (2001).
20. J. G. O'Leary, N. G. Hatsopoulos, Early visuomotor representations revealed from evoked local field potentials in motor and premotor cortical areas. *J. Neurophysiol.* **96**, 1492–1506 (2006).
21. B. E. Kilavik, J. Confais, A. Ponce-Alvarez, M. Diesmann, A. Riehle, Evoked potentials in motor cortical local field potentials reflect task timing and behavioral performance. *J. Neurophysiol.* **104**, 2338–2351 (2010).
22. C. Mehring, J. Rickert, E. Vaadia, S. C. De Oliveira, A. Aertsen, S. Rotter, Inference of hand movements from local field potentials in monkey motor cortex. *Nat. Neurosci.* **6**, 1253–1254 (2003).
23. J. Rickert, S. Cardoso De Oliveira, E. Vaadia, A. Aertsen, S. Rotter, G. Mehring, Encoding of movement direction in different frequency ranges of motor cortical local field potentials. *J. Neurosci.* **25**, 8815–8824 (2005).
24. P. Pani, F. Di Bello, E. Brunamonti, V. D'Andrea, O. Papazachariadis, S. Ferraina, Alpha- and beta-band oscillations subserve different processes in reactive control of limb movements. *Front. Behav. Neurosci.* **8**, 383 (2014).
25. E. M. Meyers, D. J. Freedman, G. Kreiman, E. K. Miller, T. Poggio, Dynamic population coding of category information in inferior temporal and prefrontal cortex. *J. Neurophysiol.* **100**, 1407–1419 (2008).
26. D. A. Crowe, B. B. Averbeck, M. V. Chafee, Rapid sequences of population activity patterns dynamically encode task-critical spatial information in parietal cortex. *J. Neurosci.* **30**, 11640–11653 (2010).
27. M. G. Stokes, "Activity-silent" working memory in prefrontal cortex: A dynamic coding framework. *Trends Cogn. Sci.* **19**, 394–405 (2015).
28. M. Giamundo, F. Giarrocco, E. Brunamonti, F. Fabbrini, P. Pani, S. Ferraina, Neuronal activity in the premotor cortex of monkeys reflects both cue salience and motivation for action generation and inhibition. *J. Neurosci.* **41**, 7591–7606 (2021).
29. G. F. Elsayed, A. H. Lara, M. T. Kaufman, M. M. Churchland, J. P. Cunningham, Reorganization between preparatory and movement population responses in motor cortex. *Nat. Commun.* **7**, 13239 (2016).
30. M. Weinrich, S. P. Wise, The premotor cortex of the monkey. *J. Neurosci.* **2**, 1329–1345 (1982).
31. R. J. Bufacchi, A. Battaglia-Mayer, G. D. Iannetti, R. Caminiti, Cortico-spinal modularity in the parieto-frontal system: A new perspective on action control. *Prog. Neurobiol.* **231**, 102537 (2023).
32. L. V. Hadley, G. Novembre, P. E. Keller, M. J. Pickering, Causal role of motor simulation in turn-taking behavior. *J. Neurosci.* **35**, 16516–16520 (2015).
33. P. Cardellicchio, E. Dolfini, A. D'Ausilio, The role of dorsal premotor cortex in joint action stopping. *iScience* **24**, 103330 (2021).
34. M. Fanghella, E. Mussini, A. Zazio, F. Genovese, E. Satta, G. Barchiesi, A. Battaglia-Mayer, M. Bortoletto, C. Sinigaglia, Distinct neural dynamics in joint versus side-by-side actions: Insights from dual EEG. *Soc. Cogn. Affect. Neurosci.* **20**, nsaf075 (2025).
35. A. Battaglia-Mayer, R. Caminiti, Corticocortical systems underlying high-order motor control. *J. Neurosci.* **39**, 4404–4421 (2019).
36. J. R. Sheets, R. G. Briggs, M. Y. Bai, A. Poologaindran, I. M. Young, A. K. Conner, C. M. Baker, C. A. Glenn, M. E. Sughrue, Parcellation-based modeling of the dorsal premotor area. *J. Neurol. Sci.* **415**, 116907 (2020).
37. A. Battaglia-Mayer, A brief history of the encoding of hand position by the cerebral cortex: Implications for motor control and cognition. *Cereb. Cortex* **29**, 716–731 (2019).
38. E. Satta, S. Ferrari-Toniolo, F. Visco-Comandini, R. Caminiti, A. Battaglia-Mayer, Development of motor coordination during joint action in mid-childhood. *Neuropsychologia* **105**, 111–122 (2017).
39. L. M. Sacheli, M. Candidi, E. F. Pavone, E. Tidoni, S. M. Aglioti, And yet they act together: Interpersonal perception modulates visuo-motor interference and mutual adjustments during a joint-grasping task. *PLOS ONE* **7**, e50223 (2012).
40. D. Kourtis, M. Woźniak, N. Sebanz, G. Knoblich, Evidence for we-representations during joint action planning. *Neuropsychologia* **131**, 73–83 (2019).
41. G. Novembre, L. F. Ticini, S. Schütz-Bosbach, P. E. Keller, Distinguishing self and other in joint action. Evidence from a musical paradigm. *Cereb. Cortex* **22**, 2894–2903 (2012).
42. K. Van Braeckel, P. R. Butcher, R. H. Geuze, E. F. Stremmelaar, A. Bouma, Movement adaptations in 7- to 10-year-old typically developing children: Evidence for a transition in feedback-based motor control. *Hum. Mov. Sci.* **26**, 927–942 (2007).
43. P. S. Archambault, S. Ferrari-Toniolo, A. Battaglia-Mayer, Online control of hand trajectory and evolution of motor intention in the parietofrontal system. *J. Neurosci.* **31**, 742–752 (2011).
44. L. M. Sacheli, M. A. Musco, E. Zazzera, G. Banfi, E. Paulesu, How shared goals shape action monitoring. *Cereb. Cortex* **32**, 4934–4951 (2022).
45. G. Flandin, K. J. Friston, Analysis of family-wise error rates in statistical parametric mapping using random field theory. *Hum. Brain Mapp.* **40**, 2052–2054 (2019).
46. L. Kumle, M. L. H. Võ, D. Draschkow, Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behav. Res. Methods* **53**, 2528–2543 (2021).
47. L. Kumle, M. L.-H. Võ, D. Draschkow, Mixedpower: A library for estimating simulation-based power for mixed models in R, version v1.0, Zenodo (2018); https://doi.org/10.5281/zenodo.1341047.
48. R. C. Oldfield, The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).
49. The jamovi project, jamovi (version 2.3) [Computer Software] (2022).

50. K. J. Worsley, K. J. Friston, Analysis of fMRI time-series revisited—Again. *Neuroimage* **2**, 173–181 (1995).
51. R. Eckhorn, O. Thomas, A new method for the insertion of multiple microprobes into neural and muscular tissue, including fiber electrodes, fine wires, needles and microsensors. *J. Neurosci. Methods* **49**, 175–179 (1993).
52. M. X. Cohen, K. R. Ridderinkhof, S. Haupt, C. E. Elger, J. Fell, Medial frontal cortex and response conflict: Evidence from human intracranial EEG and medial frontal cortex lesion. *Brain Res.* **1238**, 127–142 (2008).
53. N. S. Narayanan, J. F. Cavanagh, M. J. Frank, M. Laubach, Common medial frontal mechanisms of adaptive control in humans and rodents. *Nat. Neurosci.* **16**, 1888–1895 (2013).
54. L. Zhao, X. Wang, Frontal cortex activity during the production of diverse social communication calls in marmoset monkeys. *Nat. Commun.* **14**, 6634 (2023).
55. D. Xing, C. I. Yeh, R. M. Shapley, Spatial spread of the local field potential and its laminar variation in visual cortex. *J. Neurosci.* **29**, 11540–11549 (2009).
56. A. Maier, G. K. Adams, C. Aura, D. A. Leopold, Distinct superficial and deep laminar domains of activity in the visual cortex during rest and stimulation. *Front. Syst. Neurosci.* **4**, 31 (2010).
57. Y. Kajikawa, C. E. Schroeder, How local is the local field potential? *Neuron* **72**, 847–858 (2011).
58. J. J. Orczyk, A. Barczak, J. Costa-Faidella, Y. Kajikawa, Cross laminar traveling components of field potentials due to volume conduction of non-traveling neuronal activity in macaque sensory cortices. *J. Neurosci.* **41**, 7578–7590 (2021).
59. S. Waldert, G. Vigneswaran, R. Philipp, R. N. Lemon, A. Kraskov, Modulation of the intracortical LFP during action execution and observation. *J. Neurosci.* **35**, 8451–8461 (2015).
60. D. A. Markowitz, Y. T. Wong, C. M. Gray, B. Pesaran, Optimizing the decoding of movement goals from local field potentials in macaque cortex. *J. Neurosci.* **31**, 18412–18422 (2011).
61. E. M. Meyers, The neural decoding toolbox. *Front. Neuroinform.* **7**, 8 (2013).
62. C. C. Chang, C. J. Lin, LIBSVM: A Library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**, 1–27 (2011).
63. C. Magri, U. Schridde, Y. Murayama, S. Panzeri, N. K. Logothetis, The amplitude and timing of the BOLD signal reflects the relationship between local field potential power at different frequencies. *J. Neurosci.* **32**, 1396–1407 (2012).
64. E. Combrisson, F. Di Rienzo, A. L. Saive, M. Perrone-Bertolotti, J. L. P. Soto, P. Kahane, J. P. Lachaux, A. Guillot, K. Jerbi, Human local field potentials in motor and non-motor brain areas encode upcoming movement direction. *Commun. Biol.* **7**, 506 (2024).
65. P. F. Sayegh, K. M. Hawkins, K. L. Hoffman, L. E. Sergio, Differences in spectral profiles between rostral and caudal premotor cortex when hand-eye actions are decoupled. *J. Neurophysiol.* **110**, 952–963 (2013).

Acknowledgments: We are grateful to the reviewers for insightful comments. We are also thankful to P. R. Lanzoni for supporting fMRI data collection and to all the participants who voluntarily took part in this study. This research was made possible also by the BIMR3T laboratory (https://bimr3t.unimib.it/), supported by the Tecnomed Foundation (https://fondazionetecnomed.it/).
Funding: The study was supported by the following nationally funded grants from the Ministero dell'Università e della Ricerca-NextGenerationEU: PNRR M4.C2.I1.1–NextGenerationEU–Progetti di Ricerca di Interesse Nazionale (PRIN) 2022, project title "From an individual to a shared sense of agency in humans and non-human primates" [grant nos. 2022B9NZC4 ; CUP H53D23004200006 (to E.P.) ; and CUP B53D23014490006 (to A.B.-M.)] and Young Researcher 2024 [CUP B83C24003140006 (to E.Q.)]. **Author contributions:** L.M.S.: Conceptualization, methodology (human), validation (human), investigation (human), formal analysis (human), data curation (human), writing—original draft, and visualization. S.G.: Software (NHP), validation (NHP), formal analysis (NHP), data curation (NHP), writing—original draft, and visualization. L.Z.: Conceptualization, methodology (human), investigation (human), formal analysis (human), and writing—review and editing. E.Q.: Software (NHP), formal analysis (NHP), data curation (NHP), writing—review and editing, and visualization. F.E.: Investigation (human), formal analysis (human), and writing—review and editing. M.M.: Investigation (human) and writing—review and editing. M.A.M.: Investigation (human), formal analysis (human), data curation (human), and writing—review and editing. C.T.: Software (human), formal analysis (human), data curation (human), writing—review and editing, and visualization. G.B.: Methodology, investigation (human), resources (human), data curation (human), and writing—review and editing. S.F.: Investigation (NHP), formal analysis (NHP), resources (NHP), and writing—review and editing. A.B.-M.: Conceptualization, methodology (NHP), investigation (NHP), validation (NHP), formal analysis (NHP), resources (NHP), writing—original draft, supervision, project administration, and funding acquisition. E.P.: Conceptualization, methodology (human), formal analysis (human), resources (human), writing—original draft, supervision, project administration, and funding acquisition. **Competing interests:** The authors declare that they have no competing interests. **Data, code, and materials availability:** This study did not generate new materials. All data and code needed to evaluate and reproduce the results in the paper are available at the following link: https://doi.org/10.5281/zenodo.18743721.

Submitted 30 June 2025
Accepted 13 March 2026
Published 17 April 2026
10.1126/sciadv.aea1927

ScienceAdvances

An evolutionary conserved neural mechanism for interpersonal coordination in primates

Lucia Maria Sacheli, Stefano Grasso, Laura Zapparoli, Eros Quarta, Fabiana Esposito, Marika Mariano, Margherita Adelaide Musco, Carlo Toneatto, Gianpaolo Basso, Stefano Ferraina, Alexandra Battaglia-Mayer, and Eraldo Paulesu

Sci. Adv. **12** (16), eaea1927. DOI: 10.1126/sciadv.aea1927

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.