

Coordinated social interactions are supported by integrated neural representations

Silvia Formica , Marcel Brass

Department of Psychology, Berlin School of Mind and Brain, Humboldt Universität zu Berlin, Berlin 10117, Germany

* Corresponding author. Department of Psychology, Berlin School of Mind and Brain, Humboldt Universität zu Berlin, Berlin 10117, Germany.

E-mail: silvia.formica@hu-berlin.de

Abstract

Joint actions are defined as coordinated interactions of two or more agents toward a shared goal, often requiring different and complementary individual contributions. However, how humans can successfully act together without the interfering effects of observing incongruent movements is still largely unknown. It has been proposed that interpersonal predictive processes are at play to allow the formation of a Dyadic Motor Plan, encompassing both agents' shares. Yet, direct empirical support for such an integrated motor plan is still limited. In this study, we aimed at testing the properties of these anticipated representations. We collected electroencephalography data while human participants ($N = 36$; 27 females) drew shapes simultaneously to a virtual partner, in two social contexts: either they had to synchronize and act jointly or they performed the movements alongside, but independently. We adopted a multivariate approach to show that the social context influenced how the upcoming action of the partner is anticipated during the interval preceding the movement. We found evidence that acting jointly induces an encoding of the partner's action that is strongly intertwined with the participant's action, supporting the hypothesis of an integrative motor plan in joint but not in parallel actions.

Keywords: joint action; EEG; multivariate analyses; motor cognition; predictive processing

Introduction

A huge variety of human activities can only be accomplished through coordinated interactions with others. Think of two tango dancers: to perform smoothly and elegantly, they need to coordinate their actions in time and space, constantly adjusting their movements to match their partners. Such concerted activities are often referred to as Joint Actions (Sebanz et al. 2006), situations in which two or more agents interact in order to achieve a shared goal, as, for example, performing a "pirouette." However, the cognitive mechanisms allowing for such coordination are far from being fully understood (Sacheli et al. 2018).

One crucial issue in Joint Actions is that of visuomotor interference (VMI; Kilner et al. 2003, Blakemore and Frith 2005). Observing another agent performing a movement is known to activate the corresponding motor plan in the observer (Rizzolatti et al. 1996), influencing their movement execution. An incongruent observed action produces a disruption of the observer's motor plan and, consequently, poorer motor performance (Brass et al. 2001, Cracco et al. 2015, Forbes and Hamilton 2017, Cracco and Brass 2019). Notably, acting jointly toward a shared goal often requires the coordination of complementary and incongruent movements (Sartori and Betti 2015). It has been proposed that interference is overcome by generating a Dyadic Motor Plan that integrates the anticipated behavior of the other into one's

own motor plan (Sacheli et al. 2018). This perspective is in line with the idea of engaging in a "we-mode" during social interactions (Gallotti and Frith 2013), capturing task representations that go beyond concurrent representations of own and others' individual action contributions by also specifying relations between them that emerge at the level of the group (Kourtis et al. 2019, Marschner et al. 2024). Indirect evidence for such an integration process comes from behavioral data on reduced VMI when acting jointly compared to acting alongside but independently (i.e. in parallel), quantified in faster reaction times and lower distortions in the kinematic profiles of the executed movements (Sacheli et al. 2018, 2019, Clarke et al. 2019, Rocca et al. 2023).

Neurocognitive studies addressing interpersonal coordination mostly focused on how acting jointly modulates motor activation following an observed movement, thus inferring predictive processes during action unfolding (Bolt and Loehr 2021). Only few electroencephalography (EEG) studies investigated the interval preceding the overt movements. Some reported univariate differences in specific event-related potentials between joint and individual action contexts, suggesting an anticipation of the partner's contribution supported by dedicated attentional and motor processes (Kourtis et al. 2013, 2014). Another study highlighted the role of the relational properties between the individual agents' contributions, showing that even in the absence of information

Received: 10 April 2024; Revised: 14 October 2024; Accepted: 29 November 2024

© The Author(s) 2024. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

concerning the individual movements, knowing whether they will be the same or different allows for the formation of cognitive and sensorimotor “we-representations” that benefit the performance of the coordinated action (Kourtis et al. 2019). Altogether, these studies provide initial evidence for differences in neural activity associated with different task demands while preparing for Joint Actions. However, they do not test directly for differences between the same pairs of movements performed in a coordinated (i.e. joint) or independent manner (i.e. parallel).

To fill this gap, we developed a novel Paired Drawing Task, during which we asked participants to draw the instructed shapes either synchronously with a virtual partner (i.e. Joint social context) or in parallel. By controlling visual, motor, and attentional demands of the task, we ensured differences in how the action plans are represented during the preparation interval could only be attributed to their relevance for the upcoming Joint or Parallel action. We set out to address two main questions. First, we tested whether drawing jointly reduced VMI, as reported in previous behavioral studies (Sacheli et al. 2018). Second, in line with the outlined theoretical framework (Sacheli et al. 2018) and with the existing evidence (Kourtis et al. 2013, 2019), we hypothesized that acting jointly induces the formation of dedicated neural representations, proactively anticipating and tying together each agent’s contribution, that are fundamentally distinct from simply encoding two action plans. Therefore, to directly tap into the representational content of the neural activity recorded during the preparation interval, we implemented a multivariate approach to investigate what information was decodable from the spatial and temporal activation patterns. We reasoned that, if during Joint Actions the upcoming movements of the participant and the partner are represented intertwined, it should be difficult to distinguish which specific shape each of the two is about to draw, leading to lower classification accuracies when they are drawing jointly. This pattern of results would be indicative of the emergence of an integrative representation, as proposed by the Dyadic Motor Plan framework.

Materials and methods

Participants

Forty participants were initially recruited, and gave their informed consent prior to the beginning of the experiment in accordance to the approved application to the Ethical Committee of Humboldt University (reference number: 2022-45). Eligibility criteria included age between 18 and 35 years of age, right-handedness, no history of neurological or psychological conditions, no limitations to upper limb mobility, and suitability for EEG procedures. Sample size for this experiment was not computed *a priori*, but chosen to be analogous to that of other studies using similar analytical approaches (Wolff et al. 2020, Muhle-Karbe et al. 2021, Häberle et al. 2023).

Two participants were discarded because of poor EEG data quality (>50% of trials dropped due to excessive noise) and technical issues during data collection. Exclusion criteria based on behavioral performance caused the removal of one participant due to accuracy below 60% in response to catch trials, and one participant because their mean delta time in the Joint task exceeded 3 SDs from the group mean, implying poor compliance to the experiment instructions (see below for the full explanation of the task and measures considered). Following exclusions, the final sample consisted of 36 participants ($M_{\text{age}} = 24.56, \pm 4.86$; 27 females and 9 males), all right-handed and with normal or corrected-to-normal

vision acuity. Participants received a compensation of 30€ for their time.

Materials

Each participant performed the task sitting comfortably at approximately 90 cm from a computer monitor (resolution 3840 × 2160, diagonal display length ~69.47 cm, set on a 30 Hz refresh rate to ensure exact timing). They were asked to use their right hand to draw on a graphic tablet oriented vertically (One by Wacom, model number CTL-472, with an active area of 15.2 × 9.5 cm). Stimuli presentation, collection of responses from both the keyboard and the tablet, as well as the interaction with the EEG recording software, were managed through the Psychopy toolbox, version 2022.1.3 (Peirce 2007). The continuous trace produced by the pen on the graphic tablet was displayed live on the monitor, as a white dot (8 pixels radius) moving on a gray background, with a sampling rate of 30 Hz (i.e. the position of the dot was updated according to the new coordinates of the pen on the tablet every ~33 ms). The entire surface of the graphic tablet was mapped on the right hemisphere of the monitor, meaning that the left edge of the tablet corresponded to the vertical midline of the monitor. Participants started the experiment with extensive practice to familiarize with the experience of drawing on the tablet, and were encouraged to explore the limits of the drawing space to grow accustomed to the correspondence between the surface of the tablet and the monitor. A solid black line (4 pixels) was depicted along the midline, spanning 1200 pixels from the center of the monitor (11.81 degrees of visual angle). At the top and bottom of this line were two black circles of 100-pixel diameter (1 degree of visual angle), which we will refer to as starting and ending points, respectively (Fig. 1).

Participants were informed that they would perform their task together with a partner. They would not meet or interact directly with the partner, but only see their drawings appear on the left side of the monitor, while their own drawing appeared on the right side. Drawings of both the participant and the virtual partner began at the starting point and reached the ending point. Both agents drew either a half circle or a half diamond, depending on the specific instructions for each trial. In reality, the drawings of the partner consisted of a set of pre-recorded, hand-free drawn half circles and half diamonds. This was done to maximize the feeling that these trajectories were the products of another human agent. It is worth pointing out that the mere belief that an observed motion trajectory is human-generated is sufficient to treat it as biological (Gowen et al. 2016, Chandler-Mather et al. 2020). Based on the existing literature, we confidently assumed the kinematics of our prerecorded trajectories would be perceived as the product of human agents. To increase the perceived fluidity of the movement, the trajectories were displayed as thin white lines (5 pixels) originating from the starting point and progressively extending over time to reach the end point, with two new (x, y) coordinates being added to the display on each screen refresh. However, their presentation was carefully controlled in order to match specific timing criteria (see below).

Procedure

Familiarization phase

All participants started the experiment with a familiarization phase to get used to drawing with the tablet, during which they were allowed to draw freely. When they felt comfortable with it, they moved on to the next phase, in which they learnt how to accurately draw the shapes required for the experiment. They were instructed to guide the tip of the pen in correspondence

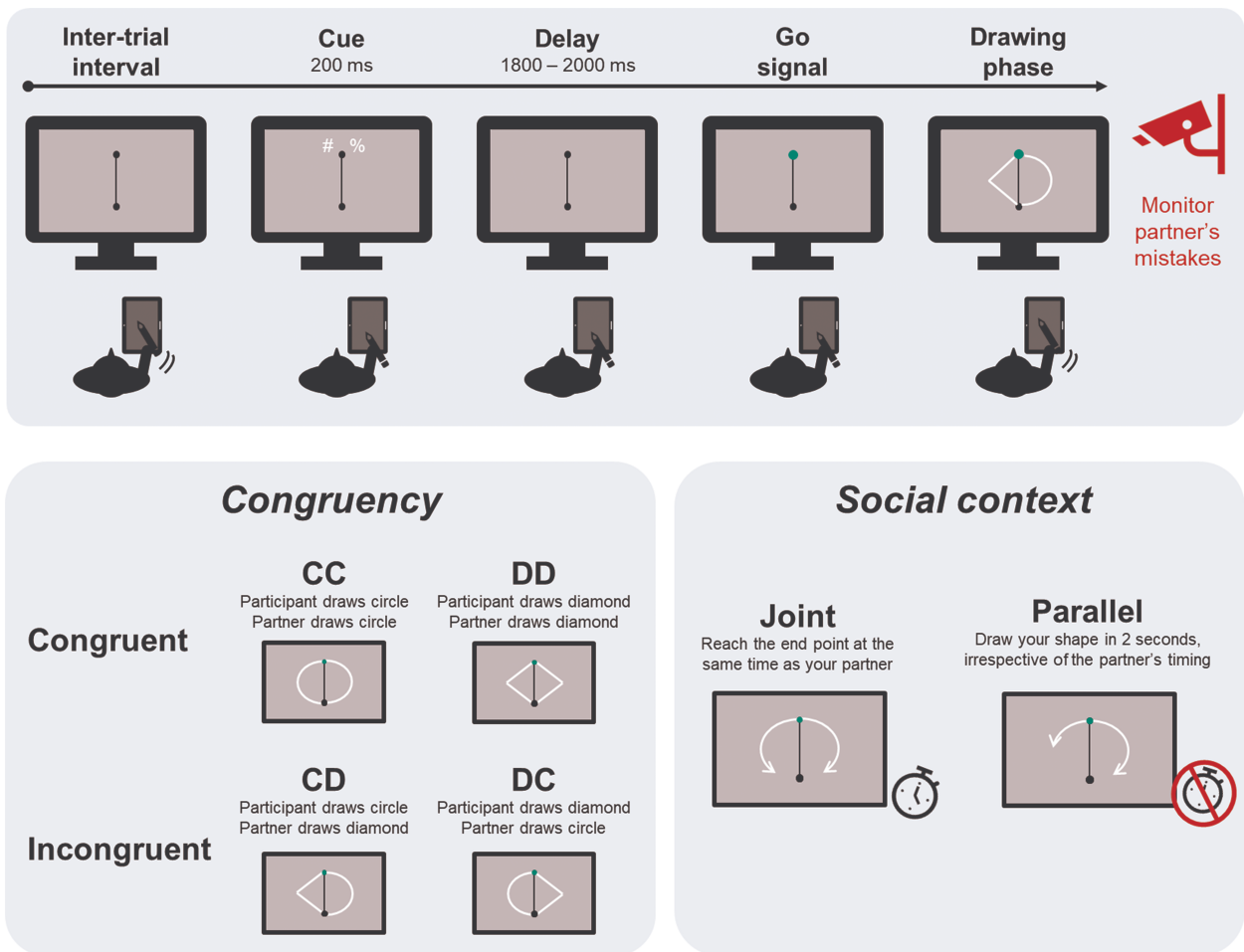


Figure 1. Behavioral paradigm.

Note. Upper panel: Timeline of one example trial. The cue instructs the participant to draw a circle, while the partner will draw a diamond. The shapes are executed after the starting point turns green. Bottom left panel: Congruency manipulation. Depending on the combination of shapes drawn by the participant and the partner, trials could be congruent or incongruent. Bottom right panel: Social context manipulation. Participants carried out a Joint and a Parallel task, requiring to synchronize the drawing time with the partner or to draw consistently in 2 s, respectively.

of the starting point, triggering the appearance of a cue above it, on the right-hand side. The cue indicated the shape to draw: these could be either half a circle or half a diamond, initiating at the starting point and reaching the end point at the bottom of the midline. The cue could be one of four symbols (#, \$, %, &). Each participant was assigned one of 6 possible cues combinations. Namely, for the whole experiment two symbols indicated circles, and two symbols indicated diamonds. At the beginning of each block, one symbol for each shape was instructed to the participant, and the cue pairs kept alternating across blocks.

The cue stayed on display for 200 ms, followed by a cue-target interval with a duration uniformly distributed between 1800 and 2000 ms. Then, the starting point would change its color to green, prompting the participant to start drawing (i.e. go signal). During this first practice phase (i.e. “Shape familiarization”), a gray shaded area appeared on the monitor to guide the participant in their drawing, indicating the area boundaries within which the drawing would be considered accurate (i.e. an area of 100 pixels surrounding the perfect circle and diamond trajectories from the starting to the ending point). To make this learning phase even more effective, the leading tip of their drawn trajectory would be large (15 pixels) if its location was within the indicated shaded area, whereas it would be as small as the size of the remaining

trailing trace if outside of the provided boundaries. Participants completed six of such familiarization trials, without any time constraints.

Once the expected shapes were learnt, participants completed 12 additional trials (6 for each shape) without the aid of the shaded area to guide them (i.e. “Execution practice”). They were instructed to replicate the same shapes they practiced in the previous step, and received analogous online feedback in the form of the pen tip diameter according to the same criteria. Here, they were additionally required to draw their shapes in approximately 2 s, and to initiate the movement as soon as possible when the starting point turned green. Feedback was provided after each trial, evaluating the drawing based on three criteria: (I) the trajectory was considered accurate if at least 60% of its coordinates fell within the ideal trajectory boundaries; (II) the total drawing duration was between 1.6 and 2.4 s; and (III) the starting time (i.e. interval between the starting point turning green and the tip of the pen leaving it) was <800 ms.

Next, participants completed an “Observation practice” phase, during which they were not required to draw, but were introduced to the drawing performed by the virtual partner on the left side of the monitor. Analogously to the Execution practice, the trial started when the tip of the pen was placed within the starting

point, this time triggering the appearance of a symbol above it on the left side. This indicated the movement that the virtual partner would perform after the cue-target interval (1800–2000 ms). After the go signal and a delay spanning from 400 to 600 ms, one of the predrawn trajectories started being displayed, for a total duration ranging from 1.6 to 2.4 s. In 4 trials out of 12, the trajectory being displayed did not match the identity of the presented cue (i.e. “catch trials”). Participants were instructed to monitor the performance of their virtual partner, and to press the “b” key on the keyboard as soon as they noticed a mistake in the shape being drawn. The goal of introducing catch trials was to ensure that attention was allocated to the action of the partner, and that its cued identity was retained during the whole delay interval.

Once these task subcomponents had been practiced individually, they were combined and performed together. Namely, for each trial, two symbols would appear as cues, one indicating a shape for the participant and one for the partner. These could be the same (i.e. congruent trials) or two opposite shapes in the case of incongruent trials. After the go signal, the goal of the participant was to draw their shape, and also to monitor the drawing of the partner to detect potential catch trials.

Social context manipulation: practices and main tasks

The core social manipulation in our experiment consisted of a Joint and a Parallel condition, administered in a block design to all participants, with their order counterbalanced. The two social contexts were designed to be as similar as possible in terms of visual input, amount, and relevance of the information provided to the participant in each trial, motor requirements, and overall movement durations. What distinguished the two conditions was the extent to which the participant had to online adjust their movement to that of the partner. Namely, in the Joint social context, the requirement for participants was to draw accurately, and to synchronize their drawing time to that of the partner, with the goal of reaching the end point simultaneously. On the contrary, in the Parallel context participants had to draw accurate trajectories, trying to keep their timing consistent to approximately 2 s, irrespective of the drawing time of the partner. In other words, the specific kinematic properties of the partner’s movement were behaviorally relevant for the participants only during the Joint task, as they had to adjust their behavior to the speed of the partner’s movement, whereas in the Parallel task only the identity of the shape drawn by the partner had to be monitored to detect catch trials (Fig. 1).

Each of the two conditions was preceded by a task-specific practice phase. Blocks of 12 trials (3 for each combination of shapes between participant and partner, namely both circles, both diamonds, circle and diamond, diamond and circle) were administered. Two trials in each block were catch trials (i.e. the partner drew the opposite shape to that indicated by their cue). The performance in these practice trials was again measured according to three criteria: (I) the trajectory was considered accurate if at least 60% of its coordinates fell within the ideal trajectory boundaries; (II) the starting time of the participant’s drawing was within 800 ms from the go signal; if the starting time exceeded 800 ms, the trial was interrupted and considered an error; (III) the drawing time was in accordance with the task goals. More specifically, in the Joint context, a trial was considered good with respect to timing if the difference (i.e. delta) between the moment the participant and the partner reached the end point was below 200 ms. Small deltas indicate good synchronization to the movement speed of the partner. On the contrary, the Parallel context

required participants to be consistent in their own timing, and trials were considered timely if the drawing was within 200 ms from the instructed 2 s duration. After each sequence of 12 practice trials, the performance in the block was evaluated. If across trials the average trajectory accuracy was above 60%, there were no more than two trials with slow starting time, at least one catch trial was correctly detected, and the task-specific timing requirement was fulfilled, the practice ended and the main task would start. Otherwise, another practice block would start, up to a maximum of eight blocks. On average, 2.6 (± 1.85) blocks were completed for the Joint context, and 2.4 (± 2.02) for the Parallel context. The number of practice blocks needed for the two tasks did not differ significantly across participants ($W = 182.5$, $P = .65$, Common Language Effect Size (CLES) = 0.57).

Following the successful completion of the task-specific practice, 6 blocks of 48 trials were completed for each of the 2 social contexts. Each participant completed 60 trials for each context and shape pair combination, plus 12 additional catch trials, resulting in a total of 576 trials. The trial structure was identical to the practice, with the only difference that feedback was provided only for missed catch trials (as a message stating that “your partner drew the wrong shape!”) and for trials with slow starting time (“You waited too long to start drawing. Try to start drawing as soon as the starting point turns green!”). Trials with a slow start were immediately interrupted and repeated at the end of the block (maximum once). In each block of the main task, eight were catch trials, to ensure in both social contexts participants retained the cue identity for the partner’s shape throughout the whole trial, and allocated comparable attention to the drawing of the partner. Feedback was provided at the end of each block as the average accuracy in terms of trajectory, timing, and detection of the partner’s mistakes.

Crucially, timing parameters were adjusted carefully during the Joint and Parallel contexts, in order to characterize the different task requirements. In the Joint task, the movement duration of the partner was uniformly distributed from 1.8 to 2.2 s. In the Parallel task, the movement duration of the partner was sampled from a slightly larger distribution, spanning from 1.6 to 2.4 s. The reason for this discrepancy was to ensure participants could not rely on the movement duration of the partner to fulfill the timing requirements of the Parallel task (i.e. in some trials the partner would be too fast or too slow with respect to the accepted timing criterium), and therefore ensuring a difference in how the two tasks were approached and performed, while keeping all other factors equal.

Throughout the whole experiment, another timing parameter was updated after each block. The participant’s starting times (i.e. time between the go signal and the pen tip leaving the starting point) were averaged at the end of each block. This value was then used in the subsequent block to adjust the starting time of the partner’s movement, sampled from a uniform distribution centered on the averaged starting time of the participant, and spanning ± 100 ms. While keeping constant for the whole experiment the maximum allowed starting time of 800 ms, this adaptive starting time for the partner’s movement (1) guaranteed that the partner’s performance was credibly attributed to another human agent, and (2) reduced the likelihood of participants waiting for the partner’s drawing to start, in order to detect potential catch trials, before proceeding to execute their own movement. This set-up ensured that drawing and monitoring of the partner’s performance happened simultaneously.

After the preparation and the EEG cap set-up (~40 min), the total duration of the experiment, including the familiarization

phase, the two task-specific practices, and the two main tasks, was of approximately 90 min. Participants could take a self-paced break between each block of the main task, and a longer break was encouraged after the end of the first main task, in order to allow participants to relax before learning and practicing the second task. At the end of the experimental phase, the EEG cap was removed and a short debriefing was provided to the participants.

Experimental design and behavioral analyses

The experiment consisted of the within-subject factor Social context (Joint vs Parallel) and the within factor Combination, with four levels corresponding to the crossing of the shape drawn by the participant and the shape drawn by the partner (CC: both agents draw circles, DD: both agents draw diamonds, CD: the participant draws a circle and the partner a diamond, DC: the participant draws a diamond and the partner a circle; the first letter always indicates the shape of the participant, and the second the shape of the partner). Crucially, the four combinations could be grouped for subsequent analyses in the factor Congruency, with the two levels Congruent (combinations CC and DD) and Incongruent (combinations CD and DC).

With respect to general behavioral performance in the two social contexts, we compared some indicators to evaluate whether participants approached the two tasks differently. First, we compared accuracy in response to catch trials between the two contexts by means of a nonparametric Wilcoxon signed-rank test (due to violated normality assumption, Shapiro-Wilk test $P < .05$). Then, we quantified whether the timing requirements of the two contexts were correctly implemented. For the Joint task, we expected average delta times to not differ significantly from 0, and hence we tested for this comparison with a one-sample t -test. Analogously, in the Parallel task participants were instructed to draw consistently in 2s, thus we compared the average duration of their movement (from the moment the pen left the starting point to the movement it reached it) with this instructed duration. All statistical comparisons were carried out with the Python toolbox Pingouin version 0.5.3 (Vallat 2018) unless otherwise specified.

Trajectory analyses

With respect to overt motor behavior, our main goal was to quantify the VMI elicited by the partner's movement on the drawing of the participant, and to investigate whether this was influenced by the social context manipulation. Trajectories were recorded as the collection of (x, y) coordinates covered by the tip of the participant's pen from the moment it left the starting point (i.e. beginning of the drawing action) to the moment it reached the ending point. Only trajectories recorded during noncatch trials, extending fully from the starting point to the ending point, were used for this analysis.

As one new pen coordinate was sampled at the refresh rate (30 Hz, one sample each ~ 33 ms), trajectories recorded in each trial had different number of data points, depending on the overall drawing duration. Therefore, the first step to make trajectories comparable across trials and conditions was to interpolate and resample them to the same number of data points. For each trajectory (i.e. each trial), we fitted a cubic spline (with the default settings of the function `interpolate.CubicSpline()` from the `scipy` package), and sampled 100 points from it. Next, we visually inspected all the interpolated trajectories and discarded those in which the participant made gross errors (e.g. the pen slipped from their hand resulting in a meaningless scribble). On average, gross errors were identified in 5.45 (± 6.88) trials per participant,

corresponding to 1.13% (± 1.41). Notably, we confirmed that this trimming procedure did not result in discarding more trials in any specific experimental condition, by submitting the count of discarded trials for each participant and condition to a Generalized Linear model with factors Social context and Congruency, using a Poisson distribution and Sum contrast coding (The Jamovi Project, 2023). We found no significant effect (all P s $> .21$), indicating that participants did not commit more gross errors in any specific experimental condition.

The visual inspection implemented was agnostic with respect to the shape indicated by the cue. In other words, we did not visually discard trajectories based on whether the drawn shape matched the instructed one or not. To objectively quantify such swap errors, for each trajectory we identified the rightmost coordinate point, and used it as knot to fit a linear and a quadratic univariate spline (`interpolate.LSQUnivariateSpline` from the `scipy` package, with an order of 1 and 2, respectively). Practically, we fitted two straight lines or two curves between the identified rightmost point, and the first and last point of the trajectory. Linear fits should approximate diamonds better, whereas quadratic fits were expected to better capture circles. Then, for each trial we checked which of the two fits performed better (i.e. had lower residuals) and marked as swap errors trajectories in which the fit inconsistent with the instructed cue had lower residual values. This resulted in an average of 4.22 (± 7.43) trials marked as swap errors per participant, corresponding to 0.92% (± 1.62). Again, we tested whether participants committed more swap errors in any experimental condition by means of a Poisson Generalized Linear Model, with factors Social context and Congruency (The Jamovi Project 2023). We found no significant effect (all P s $> .150$). Since it cannot be disentangled whether swap errors were due to an inaccurate retention of the cue identity, or were the result of VMI, and since they did not differ in number across conditions, we discarded these trials from subsequent analyses. Overall, we retained for analyses an average of 462.08 (± 15.11) trajectories per participant (115.52 ± 4.24 for each Social context \times Congruency condition).

VMI in continuous movements has been defined in previous work as the variability in movement trajectories (Kilner et al. 2003). To quantify VMI in our task, we extracted for each trial an index of distortion of the drawn trajectory from a condition-nonspecific template. For each participant, we first computed a template circle and a template diamond, by averaging all instances of each of the two shapes, across all experimental conditions. Then, we measured the area subtended by the trajectory drawn in each individual trial and the corresponding template shape. This resulted in a single value per trial, indicating how distant the trajectory was from the participant-specific average shape template, with larger values indicating larger deviations from the template (Fig. 4, left panel). Because the resulting distribution of these values severely violated normality, we applied a Box-Cox transformation (Pek et al. 2018, Atkinson et al. 2021), as implemented in the `scipy.stats` package (Virtanen et al. 2020). The optimal lambda was estimated to be 0.336, and the transformed data successfully approximated normality (skewness = -0.037 , kurtosis = -0.164). Therefore, we fitted this Box-Cox transformed distribution to a Linear Mixed Model using the `lme4` package in R (Bates et al. 2014). We included as fixed effects Social context, Congruency, and their interaction; and we included a random intercept for each participant [in `lme4` notation: `Area ~ SocialContext * Congruency + (1 | Subject)`], setting a Sum contrast coding for both predictors ($-1, 1$). Notably, adding random slopes to the model equation resulted in failure of model convergence, and therefore we implemented the simple intercept model. Visual

inspection of the residuals did not reveal a deviation from normality. *P*-values were computed with the Satterthwaite-corrected degrees of freedom (Kuznetsova et al. 2017).

Additionally, based on the Dyadic Motor Plan framework and previous evidence linked to it (Sacheli et al. 2018, 2019, Clarke et al. 2019), we formulated also the more specific and directional hypothesis of larger distortions in incongruent trials of the Parallel task. To test it, we computed the Congruency Effect (i.e. mean distortion in incongruent trials—mean distortion in congruent trials), separately for each of the two tasks. We then compared the Congruency Effects across Tasks with a paired-samples one-tailed *t*-test, assuming larger values for the Parallel task.

To ensure distortions in the drawn trajectories are not caused by confounding factors other than the target experimental manipulations, we performed two additional control analyses. First, we added to the model described in the previous paragraph an additional fixed effect, namely the centered drawing time [in lme4 notation: $Area \sim SocialContext * Congruency * DrawingTime + (1 | Subject)$]. The rationale for this control analysis was to rule out that larger distortions could be due to faster drawing performance, akin to a speed-accuracy trade-off. Second, we tested for the effect of the order in which the two tasks were performed. Namely, participants performed the two tasks in a blocked fashion, with the order of the two tasks counterbalanced across participants. With this control analysis, we aimed at checking that the VMI was not systematically influenced by the order in which the two tasks were administered. To this goal, we fitted the model $Area \sim SocialContext * Congruency * TaskOrder + (1 | Subject)$.

EEG recordings and preprocessing

Electrophysiological data were recorded with a Biosemi ActiveTwo system, with 64 Ag-AgCl electrodes arranged in accordance with the international 10-20 system (Klem et al. 1999). The set-up included a Common Mode Sense–Driven Right Leg (CMS–DRL) electrode pair, two external electrodes attached to the left and right mastoids, and four electrodes (two at the outer canthi of both eyes, one above and one below the left eye) used to monitor horizontal and vertical eye movements. Data were recorded at a sampling rate of 1024 Hz, and we aimed at keeping impedance below 10 k Ω .

All preprocessing steps were carried out using the MNE Python toolbox version 1.3.1 (Gramfort 2013). First, to achieve the maximum signal-to-noise ratio provided by the CMS–DRL set-up, we initially re-reference the data to the electrode POz (spatially located in the cap between CMS and DRL). Then, we implemented a bandpass FIR filter on the continuous data, in the frequency interval 0.1–40 Hz (Hamming window with 0.0194 passband ripple and 53 dB stopband attenuation, with lower and upper transition bandwidths of 0.1 and 10 Hz, respectively). Next, we used the NoisyChannels function, available in the toolbox PyPrep version 0.4.2 (Bigdely-Shamlo et al. 2015, Appelhoff et al. 2022) to detect bad channels in the continuous recordings. We adopted the default parameters of the toolbox and searched for bad channels according to the whole range of methods available, namely: low signal-to-noise ratio, channels with poor correlation with the surrounding ones, channels containing abnormally high amplitudes, excessive high frequency noise, flat or near-flat values, and channels poorly predicted by the others (*ransac* approach). This comprehensive procedure resulted in an average of 3.64 (± 1.83) channels being flagged as bad across participants. Note that for all participants the channel POz was indicated as flat due to it being used as reference.

Our main interest was to test for the preparatory neural representations associated to dyadic movements. Therefore, we focused our analyses in the delay period between the cue and the go signal. The continuous data were epoched time-locked to the onset of the cue (initially from –500 to 3000 ms), linearly detrended, and downsampled to 128 Hz. To remove electrophysiological activity associated with eye movements, an Independent Component Analysis was performed and components were selected to be discarded when they correlated with the activity recorded by the electrooculogram located above the left eye (MNE function `ica.find_bads_eog()`). On average, 1.25 (± 0.6) components were discarded per participant. After re-referencing to the average of all electrodes, we proceeded to discard trials in which the 150 μ V peak-to-peak amplitude threshold was exceeded (on average, 3.28% $\pm 4.83\%$ trials per participant). Finally, epochs were baseline corrected to the average signal in the time window 200–0 ms preceding the cue onset.

As the focus of our hypotheses was the delay period following the cue and preceding the go signal, both regular and catch trials were included in the analyses. As we aimed at including in the analyses only trials with an accurate representation of the other's shape during the cue-target interval, catch trials that were not correctly identified, and regular trials that were indicated to be catch trials were excluded from the analyses. For each individual condition (crossing of the factors Social context and Congruency), we retained an average of 139.74 (± 6.36) trials per participant, corresponding to 97.04% (± 4.42) of available trials.

EEG decoding analyses

Decoding action combinations in congruent and incongruent trials

In line with the Dyadic Motor Plan hypothesis (Sacheli et al. 2018), we expected that, during the preparation interval of the Joint condition, the upcoming participant's and partner's movements would become integrated in a conjoined representation. We reasoned that this would result in low decoding accuracies when trying to discriminate incongruent combinations (CD vs DC), as both would include the same set of two movements. On the contrary, information concerning the two movements is thought to be maintained during Parallel actions in a less intertwined fashion, leading to higher discriminability. Given the complete overlap between the two upcoming actions in the Congruent combination pairs, we expected high discriminability in both Social contexts. Therefore, to test this hypothesis, we performed two-way classifications on the broadband EEG data, with the goal of classifying the two possible movements combinations (i.e. CD vs DC and CC vs DD), separately for the two contexts. The classification procedure was carried out using the Scikit-learn Python toolbox version 1.0.2 (Pedregosa et al. 2011) and MNE Python version 1.3.1 (Gramfort 2013).

To increase signal-to-noise ratio in the input for the classifier, we created “supertrials” by averaging together randomly sampled sets of four trials belonging to the same category (Grootswagers et al. 2017). To minimize the influence of the visual features of the cue on the classifier, for each supertrial we averaged together trials belonging to the same experimental condition, but randomly sampled in equal number from each of the two cue pairs (e.g. one supertrial for the condition CC would be the result of averaging together two CC trials with cue pair 1, and two CC trials with cue pair 2). For each classification, we created 10 supertrials for each condition. The mean activity was then subtracted from each supertrial and each channel separately, to normalize the voltage fluctuation over time (Wolff et al. 2020, Muhle-Karbe et al. 2021).

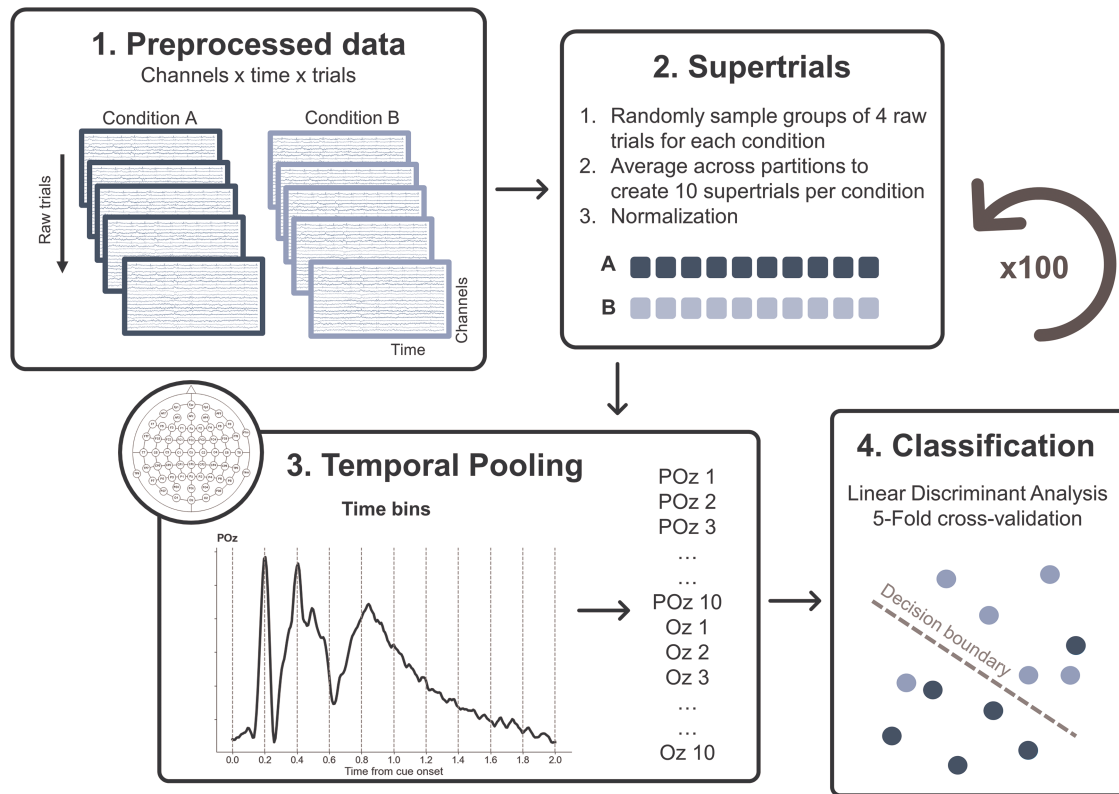


Figure 2. Spatiotemporal decoding.

Note. Epoched data after preprocessing (1) are averaged together in random partitions of the same conditions to create supertrials (2). Time courses for each of the 64 channels are pooled in time, averaging voltages in 10 time-bins (3). The resulting 640 values are used as features to train and test an LDA classifier (4). The whole procedure is repeated 100 times.

Instead of performing decoding separately for each time point, we implemented spatiotemporal decoding, capturing multivariate information encoded not only in spatial patterns but also in their temporal evolution (Wolff et al. 2020, Muhle-Karbe et al. 2021). With this approach, both spatial and temporal features of the neural activity are used for classification within participants, with the advantages of improving decoding accuracy and being insensitive to potential differences in decoding latencies across participants. We were interested in the neural representations arising during the delay period, thus anticipating the overt movement. Therefore, our time window of interest covered the interval between the onset of the cue and the subsequent 2000 ms (i.e. the shortest possible delay interval, accounting for its variability). We averaged the data from these 2 s window in 10 bins of 200 ms. As a result, each single supertrial was treated as a discrete event, with 640 features (i.e. 10 binned-values for each of the 64 electrodes).

These binned supertrials were used as input to train and test Linear Discriminant Analysis (LDA) classifiers, with a least square solver and automatic shrinkage based on the Ledoit-Wolf lemma. We used a stratified five-fold cross-validation approach. This means that the pool of 20 supertrials entering each classification was split 5 times in nonoverlapping training sets of 16 supertrials, and testing sets of 4 supertrials, equally representing the 2 classes, and the classification accuracy was averaged across the 5 folds.

Importantly, to avoid biases due to the random sampling of trials in creating supertrials, for each contrast of interest, we repeated the whole classification pipeline for 100

permutations (i.e. we generated supertrials and performed the 5-fold cross-validated classification 100 times) and we averaged the results (Fig. 2).

Decoding own and partner's movements

Additionally, we tested whether the two instructed movements were decodable separately during the delay period. Therefore, we adopted the same spatiotemporal approach to classify the upcoming movement of the participant, and the upcoming partner's movement, across Congruency levels and separately for the two Social contexts. In other words, we aimed at classifying whether the participant was instructed to draw a circle or a diamond (conditions CC and CD vs conditions DC and DD), and, separately, if the partner's instruction was to draw a circle or a diamond (conditions CC and DC vs conditions CD and DD).

The only difference from the classification pipeline explained above was in the creation of supertrials. To classify the upcoming movement of the participant, we created supertrials averaging together one individual trial from each of the two experimental conditions in which the participant had to draw a specific shape, and one for each physical cue identity. For example, to create one supertrial for the condition "participant was instructed to draw a circle," we averaged together one trial from condition CC (cue pair 1), one trial from condition CC (cue pair 2), one trial from condition CD (cue pair 1), and one trial from condition CD (cue pair 2). This resulted in the same total number of trials (i.e. four) contributing to the formation of one single supertrial, as in the previous contrasts.

Analogous five-fold cross-validated LDA classifications were performed, repeating the random sampling for the supertrials creation for 100 permutations.

Statistical testing of spatiotemporal decoding accuracies

The spatiotemporal classification approach described resulted in one value per participant for each contrast. To reliably test if these empirical distributions were significantly higher than chance, we compared them against null distributions obtained through labels permutations. First, we repeated for each participant the classification procedure outlined above, but this time randomly assigning the supertrials to one of the two classes. This resulted, for each participant, in 100 classification accuracies that we averaged to obtain one value for each participant, constituting the null distributions (Supplementary Fig. S1 and Supplementary Table S1). Then, we used a one-tailed paired-samples t-test to establish, for each contrast, whether the empirical distributions were significantly higher than the null distributions.

Our main hypothesis was that the decoding accuracy of incongruent movement combinations would be lower during Joint task compared to the Parallel task. To directly test for this, we performed a two-tailed paired-samples t-test contrasting decoding accuracies for Incongruent conditions (CD vs DC) between Joint and Parallel tasks. Analogously, we also compared the empirical classification accuracies across the two social contexts when decoding congruent action combinations, own movement, and partner's movement.

Spatial decoding

To have an understanding of which channels contributed most to the decoding obtained with our spatiotemporal approach, we repeated the same set of two-way classifications in a spatial fashion. Instead of treating space (i.e. channels) and time as features

for the LDA classifiers, we performed the decoding on the time-resolved voltage values, separately for each channel. This analysis results in one decoding accuracy for each channel, participant, and contrast of interest, therefore allowing us to display topographies of classification accuracies. For exploratory purposes, we conducted statistical testing to detect clusters of electrodes with decoding accuracies significantly larger than chance. We used a cluster-based permutation approach (Maris and Oostenveld 2007, Sassenhagen and Draschkow 2019), clustering across neighboring channels based on Delauney triangulation.

Results

General behavioral measures

We computed some indicators of behavioral performance to investigate compliance to the experiment instructions (Fig. 3). First, we compared accuracy in catch trials detection between Joint and Parallel tasks. We designed the experiment with the goal of equating the attentional deployment to the partner's drawing across the two conditions. Participants were highly accurate in identifying errors in partner's trajectories during both Joint (Mean = 0.94, SD = 0.05) and Parallel task (Mean = 0.92, SD = 0.06), although slightly better during the Joint task ($W = 162.0$, $P = .035$, $CLES = 0.62$).

Next, we aimed at testing whether the timing requirements imposed by the task were attended correctly by participants. In the Joint task, we expected participants to align their timing with the movement duration of the partner's, resulting in very small intervals between the moment theirs and the partner's trajectories reached the ending point (i.e. delta times). As expected, we found that on average the delta time (Mean = 0.01, SD = 0.051) was not significantly different from 0 ($t_{35} = 1.08$, $P = .28$, $d = 0.18$, $BF_{10} = 0.31$). The total duration of the movement time was not analyzed statistically for the Joint context

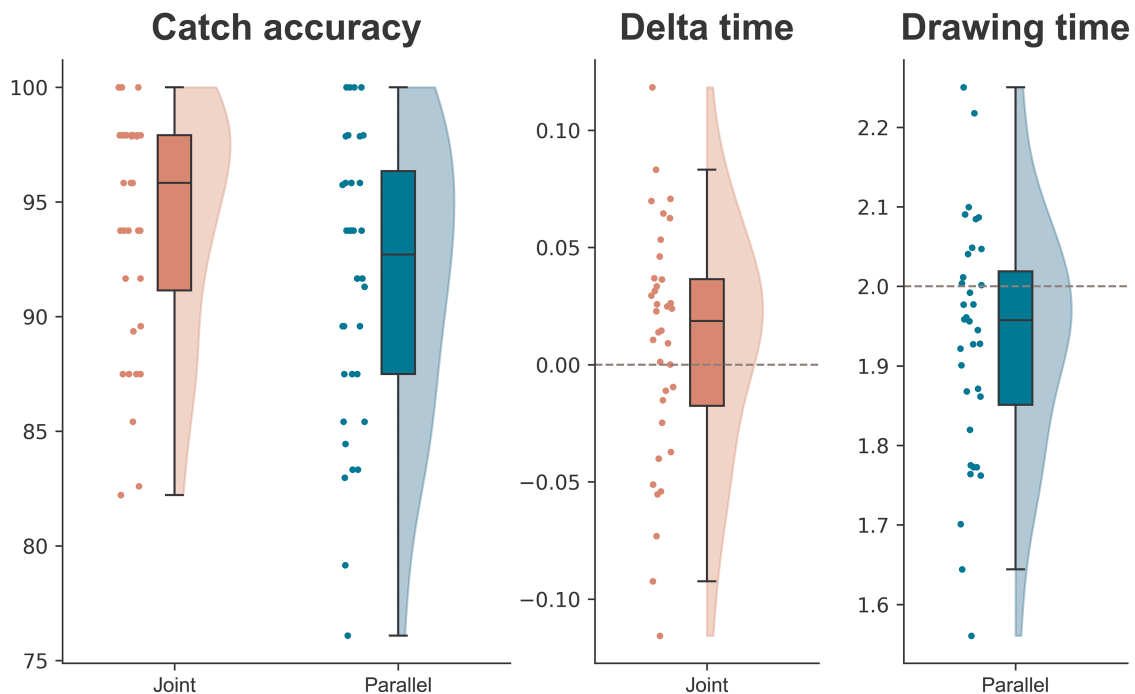


Figure 3. Indicators of behavioral performance.

Note. In each boxplot, the black line inside the box indicates the second quartile (median) of the distribution ($n = 36$). The bounds of the boxes depict the first and third quartiles of the distribution. Whiskers denote the 1.5 interquartile range of the lower and upper quartile. On the left side of each boxplot, dots represent individual scores, whereas on the right side their density is displayed as a half-violin plot.

Visuomotor interference

Distortion in the drawn trajectories

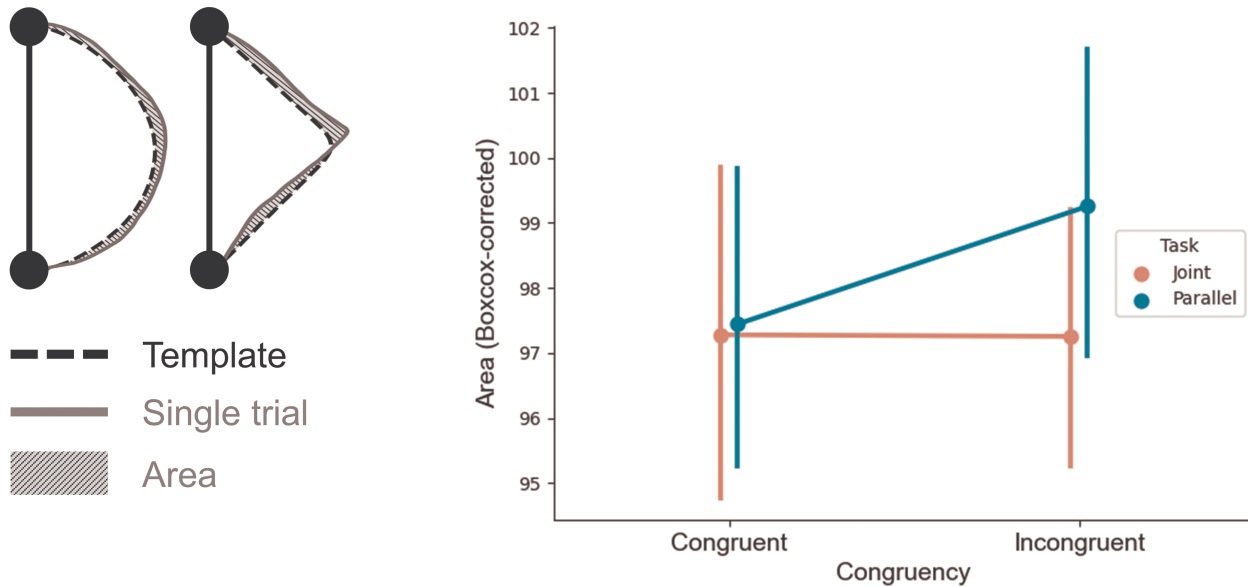


Figure 4. VMI results.

Note. Left panel: VMI is quantified as the area between the average shape templates for the participant, and each of the drawn trajectories. Right panel: Box-Cox corrected area values, for each Social context * Congruency condition. Error bars depict the confidence interval.

(Mean = 2.07, SD = 0.06), as it was biased by the discrepancy between the participant's starting time and the (adaptive) starting time of the partner.

With respect to drawing time in the Parallel task, participants were instructed to complete their trajectory in 2 s. However, they showed a tendency to systematically shorten their drawing time ($t_{35} = -2.71$, $P = .01$, $d = 0.18$), resulting in an average drawing time of 1.93 s (± 0.15).

Overall, these results support the effectiveness of our social manipulation, indicating that participants understood the timing requirements of the two tasks and behaved accordingly, although they were more successful in complying with the requirements of the Joint task.

VMI is larger in the parallel context

Our main hypothesis concerning VMI in the drawn trajectories was that an incongruent movement should result in larger distortions during the Parallel task. We predicted that, in the Joint task, the active anticipation of the partner's drawing, and its integration into a Dyadic Motor Plan, would drastically reduce the distortion in the participant's drawing, in line with other experimental findings (Sacheli et al. 2018, Rocca and Cavallo 2020, Rocca et al. 2023). We quantified VMI by computing the distance of each drawn shape from its corresponding template, computed as an average of all circles and all diamonds drawn by the individual participant.

The results of the fitted Linear Mixed Model revealed a significant effect of Social context ($\beta = -0.53$, 95% CI = [-1.02, -0.05], $t_{16629} = -2.17$, $P = .030$), indicating that trajectories were deviating more from the template in the Parallel compared to Joint task. The effect of Congruency approached but did not cross the significance threshold ($\beta = -0.45$, 95% CI = [-0.94, 0.03], $t_{16629} = -1.84$,

$P = .065$), hinting at larger distortions in incongruent trials. Finally, the interaction effect also approached but did not reach the significance threshold ($\beta = 0.45$, 95% CI = [-0.003, 0.93], $t_{16629} = 1.82$, $P = .069$).

To have a more directional test of our effect of interest, we computed the Congruency Effect for each of the two contexts (see "Methods" section). The one-tailed t-test revealed significantly larger distortions in the Parallel compared to Joint task ($t_{35} = -1.88$, $P = .034$, $d = 0.43$, $BF_{10} = 1.73$). Despite the small effect size and the inconclusive Bayes Factor, these results are in line with the hypothesis that integrating the partner's action into a Dyadic Motor Plan reduces the interfering effect of observing an incongruent movement (Fig. 4).

We fitted two additional models as control analyses. First, we extended our model to include the centered factor Drawing Time and its interactions with the factors Social context and Congruency. The rationale for fitting this model was to account for a potential effect of the speed of drawing on the distortion in the trajectories. The results of this model were analogous to the previous analysis with respect to the main effects of Social context, Congruency and their interaction. Additionally, we found a significant interaction of Congruency and Drawing Time ($\beta = -0.78$, 95% CI = [-1.27, -0.30], $t_{16625} = -3.18$, $P = .001$). Notably, this interaction indicated that slower movements were associated with larger distortions in incongruent trials, discarding the possibility of a speed-accuracy trade-off between movement duration and variability in drawn trajectories (Supplementary Fig. S2). Next, we extended our original model to include the factor Task Order and its interaction with Social context and Congruency. While this model replicated again the main findings concerning the experimental manipulations of interest, the effect of Task Order and all the related interactions were not significant. We took

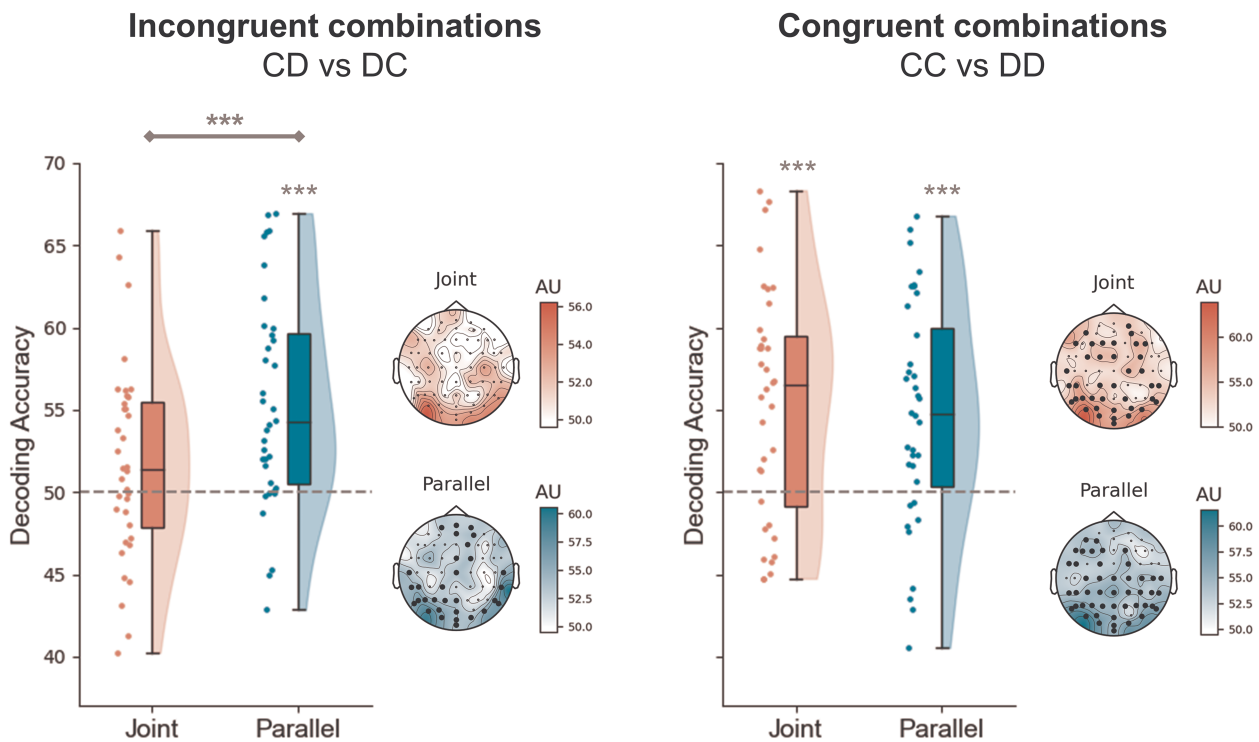


Figure 5. Decoding of incongruent and congruent combinations.

Note. Left panel: Accuracies resulting from the decoding of the conditions CD vs DC in Joint and Parallel tasks. The dashed horizontal gray line indicates the chance level (50%). In each boxplot, the black line inside the box indicates the second quartile (median) of the distribution ($n = 36$). The bounds of the boxes depict the first and third quartiles of the distribution. Whiskers denote the 1.5 interquartile range of the lower and upper quartile. On the left side of each boxplot, dots represent individual scores, whereas on the right side, their density is displayed as a half-violin plot. For visualization purposes, the topographies report the decoding accuracies for each of the 64 channels, separately for each task. Thicker dots indicate channels belonging to spatial clusters with significantly above-chance decoding accuracy. Right panel: Accuracies resulting from the decoding of the conditions CC vs DD in Joint and Parallel tasks. Plotting conventions are the same as for the left panel. ***: $P < .01$.

these results as evidence that the interfering effect of incongruent movements during the Parallel task were not associated with the order in which the participants performed the tasks, nor could be attributed to an increased variability due to faster drawing times.

Social context influences the neural representation of upcoming movements

To test whether the Joint social context induced the formation of a Dyadic Motor Plan, in which the movements of the two agents are tightly integrated, we aimed at classifying conditions pairs, separately for Joint and Parallel tasks (Fig. 5). First, we tested whether the classifier could pick up differences between the two incongruent conditions (i.e. CD vs DC). We reasoned that, if the two actions are represented intertwined in the Joint condition as proposed, it should be more difficult for the classifier to tease these two conditions apart. We found high, above-chance decoding accuracies in the Parallel task ($t_{35} = 4.89$, $P < .01$, $d = 1.09$, $BF_{10} > 100$), while it did not reach significance in the Joint task ($t_{35} = 1.50$, $P = .07$, $d = 0.36$, $BF_{10} = 0.99$). Crucially, we observed a significant difference between Joint and Parallel tasks in the decoding accuracies of incongruent combinations ($t_{35} = 2.74$, $P < .01$, $d = 0.59$, $BF_{10} = 4.37$).

With respect to discriminating between congruent movement combinations (i.e. CC vs DD), the two movement combinations were successfully classified above chance level in both the Joint ($t_{35} = 4.55$, $P < .01$, $d = 1.07$, $BF_{10} > 100$) and Parallel task ($t_{35} = 4.14$, $P < .01$, $d = 0.99$, $BF_{10} > 100$). The observed decoding accuracies did not differ between the two social contexts ($t_{35} = 0.43$, $P = .67$, $d = 0.09$, $BF_{10} = 0.19$).

For qualitative exploration, we implemented the same decoding contrasts also at the spatial level, to identify those electrodes carrying most information. With the exception of the incongruent combinations in the Joint task, the other contrasts yielded electrodes clusters of above-chance accuracy covering mostly occipito-parietal and frontal regions.

One property of the classification approach we adopted is that it is agnostic to which aspect of the data is informing the classifying algorithm. Namely, in our comparisons the classifier might be picking up information encoding the upcoming participant's movement, as well as the partner's movement. Therefore, to better disentangle the individual contributions of these two sources of information, we carried out two additional classification analyses, aimed at discriminating the upcoming participant's movement (Circle: CC and CD vs Diamond: DC and DD), and the upcoming partner's movement (Circle: CC and DC vs Diamond: CD and DD), across Congruency levels (Fig. 6).

The prepared participant's movement was successfully classified above chance in both the Joint ($t_{35} = 4.23$, $P < .01$, $d = 1.10$, $BF_{10} > 100$) and Parallel task ($t_{35} = 4.66$, $P < .01$, $d = 1.15$, $BF_{10} > 100$). We found no evidence indicating that the strength of the participant's movement representation differed between Joint and Parallel contexts ($t_{35} = 0.79$, $P = .431$, $d = 0.20$, $BF_{10} = 0.24$). Spatially, clusters of above-chance electrodes were widespread over the whole scalp.

Interestingly, classifying the partner's movement lead to above-chance accuracies in the Parallel task ($t_{35} = 2.58$, $P < .01$, $d = 0.60$, $BF_{10} = 6.28$), but not in the Joint task ($t_{35} = 0.78$, $P = .22$, $d = 0.18$, $BF_{10} = 0.48$). Decoding accuracies of the partner's

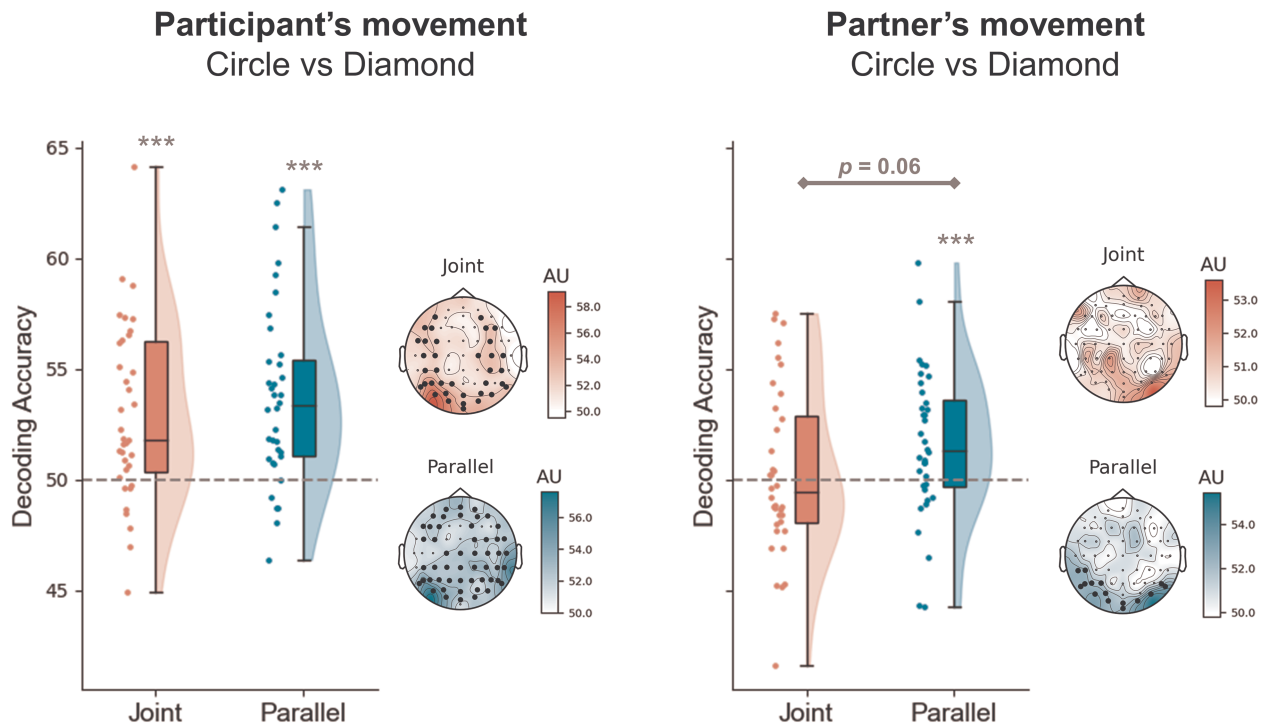


Figure 6. Decoding of participant's and partner's movements.

Note. Left panel: Accuracies resulting from the decoding of the instructed participant's movement (Circle or Diamond), separately for Joint and Parallel tasks. Right panel: Accuracies resulting from the decoding of the instructed partner's movement (Circle or Diamond), separately for Joint and Parallel tasks. See Fig. 5 for plotting conventions.

movement were almost significantly larger in the Parallel compared to the Joint task, although the formal significance threshold was not met ($t_{35} = 1.90$, $P = .06$, $d = 0.38$, $BF_{10} = 0.90$). Consistently, the spatial decoding revealed an occipital cluster of electrodes significantly above chance, only in the Parallel task.

Discussion

According to recent proposals, engaging in successful coordinated interactions entails the proactive anticipation of our partner's actions (Kilner et al. 2007), and their integration with our own motor contribution in a Dyadic Motor Plan toward the shared goal (Sacheli et al. 2018). However, so far the emergence of such integrated representations in preparation for Joint Actions has not been tested directly, but rather was supported by reduced interference effects on overt behavior (Sacheli et al. 2018, Clarke et al. 2019, Rocca et al. 2023). At the neural level, previous studies targeting specifically the preparatory phase preceding Joint Actions reported univariate differences attributed to the interactive task context (Kourtis et al. 2013, 2014, 2019). However, they did not investigate if, and to what extent, the same information is encoded differently when the two agents are asked to perform jointly or independently from each other. To answer this question, our study focused on the preparation phase, and adopted a multivariate approach to compare the patterns of electrophysiological activity between the two social contexts. We reasoned that, if the participant's and the partner's prospective contributions were integrated into a Dyadic Motor Plan, they should be more intertwined and overlapping, resulting in lower discriminability. Confirming this hypothesis, the multivariate approach we implemented revealed that the two incongruent movement combinations (i.e. circle-diamond and diamond-circle) were significantly easier to discriminate when the two agents were acting

in parallel, compared to jointly. This finding provides compelling evidence for the formation of an integrated representation incorporating both upcoming movements. In fact, it highlights how acting jointly induces the encoding of the actions in a format that goes beyond the two individual contributions, but entails also their combination at the dyadic level.

Moreover, and in line with previous studies, we hypothesized that such an integrated representation would lead to reduced VMI, quantified in the distortion of the drawn trajectories. Our results, despite warranting a cautious interpretation due to the inconclusive BF, hint at a pattern that is compatible with our hypothesis and suggest that the active anticipation of observing an incongruent movement induces its diminished interfering effect (Kaiser and Schütz-Bosbach 2018).

We further tested to what extent each of the two individual contributions was decodable from the pattern of neural activity recorded during the preparation interval. The upcoming movement of the participant was successfully decodable in both social contexts, indicating that the motor plan associated with the shape to draw was being strongly prepared (Ariani et al. 2015, 2022). With respect to the action of the partner, in the Joint condition the classifier failed at discriminating whether they were about to draw a circle or a diamond. On the contrary, this same contrast yielded high decoding accuracies during Parallel interactions, indicating that in this condition the upcoming partner's movement was represented in a format that could be more effectively picked up by the classifier.

Crucially, good performance in catch trials detection clearly showed that participants were efficiently maintaining the information pertaining the movements of the partner in both conditions. Therefore, low classification accuracies in the joint social context cannot be due to the action information from the partner being entirely disregarded. Moreover, the two social contexts were

designed to be as matched as possible with respect to cognitive and attentional demands. We consider it highly implausible that participants directed more attention toward the partner's action contribution in the parallel condition. In fact, this task did not require participants to adjust their behavior based on the partner's, unlike the joint condition which relied more heavily on the constant monitoring of the partner's timing. With these considerations in mind, we excluded the possibility of our results being explained by an attentional mechanism, but rather interpreted them as reflective of the different degree of interactivity in the two conditions.

We think these results represent an important step forward in understanding how the available information on the partner's movement is represented as a function of the interactive social context. Previous studies focused predominantly on motor activation in response to an observed movement or during interaction, often in situations in which participants had no prior information concerning the partner's contribution (Bolt and Loehr 2021). As a notable example, Sacheli et al. (2019) revealed a stronger hemodynamic response in the left ventral premotor cortex, tracking the identity of the unfolding partner's movement, while agents took turns in playing a melody (Sacheli et al. 2019). Conversely, we investigated the anticipatory phase of Joint Actions, and provided complementary evidence showing that the upcoming movement of an interactive partner could not be successfully decoded in the interval preceding its unfolding. This finding is consistent with the idea that, when cued in advance, the other's contribution is anticipated and integrated to one's own action and in reference to it, a representational format that does not allow for its classification in isolation but is functional to minimize its subsequent interfering effect. Therefore, our results extend on previous research in suggesting that coordinated interactions rely on complex representational dynamics, potentially undergoing reorganization from the anticipation to the implementation phase (Elsayed et al. 2016, Ariani et al. 2022), and shaped by social context, information availability, and its functional relevance.

One important aspect of our results that needs to be emphasized is that they are agnostic with respect to the nature of the encoded information, being it motoric, visual, abstract and semantic, or a combination thereof. We interpreted them in the light of the proposed existing framework, but the current study was not designed to disentangle between these alternatives. The spatial decoding analysis suggests that occipital electrodes are those carrying more information concerning the partner's upcoming action in the Parallel condition. This topography hints at a larger recruitment of areas associated with visual processing to maintain the partner's contribution throughout the delay period (Christophel et al. 2017). One possible explanation for this might be that in the Parallel condition participants are mostly anticipating the expected visual outcome of the partner's action, rather than its kinematic properties. On the other hand, there is growing evidence showing that the maintenance in working memory (WM) of information concerning motion patterns, biological movements, and hand postures elicits activation of somatosensory cortices, as indexed by suppression of rolandic mu rhythm (Christophel and Haynes 2014, Gao et al. 2015, Galvez-Pol et al. 2018). Moreover, the role of motor activity to smoothly coordinate with an interaction partner has been widely emphasized by previous studies on Joint Action (Bolt and Loehr 2021). In our data, we did not observe larger decoding accuracies in the joint condition over sensorimotor cortices, nor in any other scalp location. Nevertheless, we speculatively raise the intriguing hypothesis

that the social context affects the representational format of the information concerning the partner's movement. Research in the field of WM is strongly demonstrating how task demands influence the encoding and maintenance of items, supporting the idea that information is coded in the most suitable way to effectively drive behavior (Myers et al. 2017, Henderson et al. 2022, van Ede and Nobre 2023, Formica et al. 2024). In keeping with the premises of the Dyadic Motor Plan, we propose that in an interactive context the optimal representational format for the partner's action is motoric in nature, already during the anticipation phase. This would be functional to allow its efficient assimilation with the participant's contribution. In fact, perhaps unsurprisingly, the results on the decoding of the participant's own upcoming movement are compatible with a more motoric representational format, showing widespread scalp areas of high decoding accuracies in both social contexts, including also anterior electrodes over sensorimotor cortices. On the contrary, acting alongside would not require the motoric encoding of the other's action, as this does not need to be integrated with the motor plan of the participant, and can be maintained for future recognition as more visual or abstract semantic content. Future studies should be designed with the explicit goal of teasing apart the characteristics of the encoded anticipatory information, to verify if Joint Actions truly rely on the motoric anticipation of the partner's contribution.

In conclusion, our current results support the Dyadic Motor Plan framework and extend on it, demonstrating that the neural representations of the upcoming joint configuration and of the partner's movement are affected by the social context manipulation. These findings suggest that engaging in Joint Action induces the reformatting of the encoded information into integrated motoric representations, to optimize the subsequent coordinated behavior. Disentangling the specific formats of the encoded actions and investigating how they evolve between preparation and implementation thus remains an open and fascinating avenues for future research.

Author contributions

Silvia Formica (Conceptualization, Investigation, Analyses, Writing), Marcel Brass (Conceptualization, Supervision, Writing).

Supplementary data

Supplementary data is available at SCAN online.

Conflict of interest

None declared.

Funding

S.F. was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under Germany's Excellence Strategy-EXC 2002/1, Science of Intelligence (Project Ref.: 390523135) and the Einstein Foundation Berlin. M.B. is supported by an Einstein Strategic Professorship (Einstein Foundation Berlin, EPP-2018-483).

Data availability

Raw data and analyses scripts will be made publicly available upon publication through the Open Science Framework

References

- Appelhoff S, Hurst AJ, Lawrence A et al. PyPREP: A Python implementation of the preprocessing pipeline (PREP) for EEG data. Zenodo. 2022. <https://doi.org/10.5281/zenodo.6363576>
- Ariani G, Pruszyński JA, Diedrichsen J. Motor planning brings human primary somatosensory cortex into action-specific preparatory states. *ELife* 2022;**11**:e69517. <https://doi.org/10.7554/eLife.69517>
- Ariani G, Wurm MF, Lingnau A. Decoding internally and externally driven movement plans. *J Neurosci* 2015;**35**:14160–71. <https://doi.org/10.1523/JNEUROSCI.0596-15.2015>
- Atkinson AC, Riani M, Corbellini A. The box–cox transformation: review and extensions. *Stat Sci* 2021;**36**:239–55. <https://doi.org/10.1214/20-STS778>
- Bates D, Mächler M, Bolker BM et al. Fitting linear mixed-effects models using lme4. *J Stat Softw* 2014;**67**:1–48. <https://doi.org/10.1126/science.1176170>
- Bigdely-Shamlo N, Mullen T, Kothe C et al. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Front Neuroinform* 2015;**9**:1–20. <https://doi.org/10.3389/fninf.2015.00016>
- Blakemore S-J, Frith C. The role of motor contagion in the prediction of action. *Neuropsychologia* 2005;**43**:260–67. <https://doi.org/10.1016/j.neuropsychologia.2004.11.012>
- Bolt NK, Loehr JD. The motor-related brain activity that supports joint action: a review. *Acta Psychologica* 2021;**212**:103218. <https://doi.org/10.1016/j.actpsy.2020.103218>
- Brass M, Bekkering H, Prinz W. Movement observation affects movement execution in a simple response task. *Acta Psychologica* 2001;**106**:3–22. [https://doi.org/10.1016/S0001-6918\(00\)00024-X](https://doi.org/10.1016/S0001-6918(00)00024-X)
- Chandler-Mather N, Welsh T, Sparks S et al. Biological motion and animacy belief induce similar effects on involuntary shifts of attention. *Atten Percept Psychophys* 2020;**82**:1099–111. <https://doi.org/10.3758/s13414-019-01843-z>
- Christophel TB, Haynes J-D. Decoding complex flow-field patterns in visual working memory. *NeuroImage* 2014;**91**:43–51. <https://doi.org/10.1016/j.neuroimage.2014.01.025>
- Christophel TB, Klink PC, Spitzer B et al. The distributed nature of working memory. *TICS* 2017;**21**:111–24. <https://doi.org/10.1016/j.tics.2016.12.007>
- Clarke S, McEllin L, Francová A et al. Joint action goals reduce visuomotor interference effects from a partner's incongruent actions. *Sci Rep* 2019;**9**:15414. <https://doi.org/10.1038/s41598-019-52124-6>
- Cracco E, Brass M. Reaction time indices of automatic imitation measure imitative response tendencies. *Conscious Cogn* 2019;**68**:115–18. <https://doi.org/10.1016/j.concog.2019.01.001>
- Cracco E, De Coster L, Andres M et al. Motor simulation beyond the dyad: automatic imitation of multiple actors. *J Exp Psychol* 2015;**41**:1488–501. <https://doi.org/10.1037/a0039737>
- Elsayed GF, Lara AH, Kaufman MT et al. Reorganization between preparatory and movement population responses in motor cortex. *Nat Commun* 2016;**7**:13239. <https://doi.org/10.1038/ncomms13239>
- Forbes PAG, Hamilton AFDC. Moving higher and higher: imitators' movements are sensitive to observed trajectories regardless of action rationality. *Exp Brain Res* 2017;**235**:2741–53. <https://doi.org/10.1007/s00221-017-5006-4>
- Formica S, Palenciano AF, Vermeylen L et al. Internal attention modulates the functional state of novel stimulus-response associations in working memory. *Cognition* 2024;**245**:105739. <https://doi.org/10.1016/j.cognition.2024.105739>
- Gallotti M, Frith CD. Social cognition in the we-mode. *TICS* 2013;**17**:160–65. <https://doi.org/10.1016/j.tics.2013.02.002>
- Galvez-Pol A, Calvo-Merino B, Capilla A et al. Persistent recruitment of somatosensory cortex during active maintenance of hand images in working memory. *NeuroImage* 2018;**174**:153–63. <https://doi.org/10.1016/j.neuroimage.2018.03.024>
- Gao Z, Bentin S, Shen M. Rehearsing biological motion in working memory: an EEG study. *J Cognitive Neurosci* 2015;**27**:198–209. https://doi.org/10.1162/jocn_a_00687
- Gowen E, Bolton E, Poliakoff E. Believe it or not: moving non-biological stimuli believed to have human origin can be represented as human movement. *Cognition* 2016;**146**:431–8. <https://doi.org/10.1016/j.cognition.2015.10.010>
- Gramfort A. MEG and EEG data analysis with MNE-Python. *Front Neurosci* 2013;**7**:267. <https://doi.org/10.3389/fnins.2013.00267>
- Grootswagers T, Wardle SG, Carlson TA. Decoding dynamic brain patterns from evoked responses: a tutorial on multivariate pattern analysis applied to time series neuroimaging data. *J Cognitive Neurosci* 2017;**29**:677–97. https://doi.org/10.1162/jocn_a_01068
- Häberle G, Çelikkol AP, Cichy RM. The Influence of the Bullseye Versus Standard Fixation Cross on Eye Movements and Classifying Natural Images from EEG. *bioRxiv* 2023. <https://doi.org/10.1101/2023.03.21.532944>
- Henderson MM, Rademaker RL, Serences JT. Flexible utilization of spatial- and motor-based codes for the storage of visuo-spatial information. *ELife* 2022;**11**:e75688. <https://doi.org/10.7554/eLife.75688>
- The Jamovi Project. *Jamovi* (Version 2.3). 2023. <https://www.jamovi.org> (31 July 2023, date last accessed).
- Kaiser J, Schütz-Bosbach S. Sensory attenuation of self-produced signals does not rely on self-specific motor predictions. *Eur J Neurosci* 2018;**47**:1303–10. <https://doi.org/10.1111/ejn.13931>
- Kilner JM, Friston KJ, Frith CD. Predictive coding: an account of the mirror neuron system. *Cogn Process* 2007;**8**:159–66. <https://doi.org/10.1007/S10339-007-0170-2>
- Kilner JM, Paulignan Y, Blakemore SJ. An interference effect of observed biological movement on action. *Curr Biol* 2003;**13**:522–25. [https://doi.org/10.1016/S0960-9822\(03\)00165-9](https://doi.org/10.1016/S0960-9822(03)00165-9)
- Klem GH, Lüders HO, Jasper HH et al. The ten-twenty electrode system of the international federation. The international federation of clinical neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl* 1999;**52**:3–6.
- Kourtis D, Knoblich G, Woźniak M et al. Attention allocation and task representation during joint action planning. *J Cognitive Neurosci* 2014;**26**:2275–86. https://doi.org/10.1162/jocn_a_00634
- Kourtis D, Sebanz N, Knoblich G. Predictive representation of other people's actions in joint action planning: an EEG study. *Soc Neurosci* 2013;**8**:31–42. <https://doi.org/10.1080/17470919.2012.694823>
- Kourtis D, Woźniak M, Sebanz N et al. Evidence for we-representations during joint action planning. *Neuropsychologia* 2019;**131**:73–83. <https://doi.org/10.1016/j.neuropsychologia.2019.05.029>
- Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Softw* 2017;**82**:1–26. <https://doi.org/10.18637/jss.v082.i13>
- Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 2007;**164**:177–90. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Marschner M, Dignath D, Knoblich G. Me or we? Action-outcome learning in synchronous joint action. *Cognition* 2024;**247**:105785. <https://doi.org/10.1016/j.cognition.2024.105785>

- Muhle-Karbe PS, Myers NE, Stokes MG. A hierarchy of functional states in working memory. *J Neurosci* 2021;**41**:4461–75. <https://doi.org/10.1523/jneurosci.3104-20.2021>
- Myers NE, Stokes MG, Nobre AC. Prioritizing information during working memory: beyond sustained internal attention. *TICS* 2017;**21**:449–61. <https://doi.org/10.1016/j.tics.2017.03.010>
- Pedregosa F, Varoquaux G, Gramfort A et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011;**12**:2825–30.
- Peirce JW. PsychoPy-Psychophysics software in Python. *J Neurosci Methods* 2007;**162**:8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Pek J, Wong O, Wong ACM. How to address non-normality: a taxonomy of approaches, reviewed, and illustrated. *Front Psychol* 2018;**9**:2104. <https://doi.org/10.3389/fpsyg.2018.02104>
- Rizzolatti G, Fadiga L, Gallese V et al. Premotor cortex and the recognition of motor actions. *Cognit Brain Res* 1996;**3**:131–41. [https://doi.org/10.1016/0926-6410\(95\)00038-0](https://doi.org/10.1016/0926-6410(95)00038-0)
- Rocca M, Cavallo A. Wired actions: anticipatory kinematic interference during a dyadic sequential motor interaction task. *J Exp Psychol Gen* 2020;**150**:1387–97. <https://doi.org/10.1037/xge0001003>
- Rocca M, Sacheli LM, Romeo L et al. Visuo-motor interference is modulated by task interactivity: a kinematic study. *Psychonomic Bull Rev* 2023;**30**:1–14.
- Sacheli LM, Arcangeli E, Paulesu E. Evidence for a dyadic motor plan in joint action. *Sci Rep* 2018;**8**:5027. <https://doi.org/10.1038/s41598-018-23275-9>
- Sacheli LM, Verga C, Arcangeli E et al. How task interactivity shapes action observation. *Cereb Cortex* 2019;**29**:5302–14. <https://doi.org/10.1093/CERCOR/BHZ205>
- Sartori L, Betti S. Complementary actions. *Front Psychol* 2015;**6**:557. <https://doi.org/10.3389/fpsyg.2015.00557>
- Sassenhagen J, Draschkow D. Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology* 2019;**2018**:1–8. <https://doi.org/10.1111/psyp.13335>
- Sebanz N, Bekkering H, Knoblich G. Joint action: bodies and minds moving together. *TICS* 2006;**10**:70–76. <https://doi.org/10.1016/J.TICS.2005.12.009>
- Vallat R. Pingouin: statistics in Python. *J Open Source Software* 2018;**3**:1026. <https://doi.org/10.21105/joss.01026>
- van Ede F, Nobre AC. Turning attention inside out: how working memory serves behavior. *Annu Rev Psychol* 2023;**74**:137–65. <https://doi.org/10.1146/annurev-psych-021422-041757>
- Virtanen P, Gommers R, Oliphant TE et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020;**17**:261–72. <https://doi.org/10.1038/s41592-019-0686-2>
- Wolff MJ, Jochim J, Akyürek EG et al. Drifting codes within a stable coding scheme for working memory. *PLoS Biol* 2020;**18**:e3000625. <https://doi.org/10.1371/journal.pbio.3000625>