

# The Bodily Appearance of a Virtual Partner Affects the Activity of the Action Observation and Action Monitoring Systems in a Minimally Interactive Task

 Ugo Giulio Pesci,<sup>1,2\*</sup> Quentin Moreau,<sup>1,2\*</sup> Vanessa Era,<sup>1,2</sup> and Matteo Candidi<sup>1,2</sup>

<sup>1</sup>Department of Psychology, Sapienza University, Rome 00185, Italy and <sup>2</sup>IRCCS Fondazione Santa Lucia, Rome 00179, Italy

## Abstract

One pending question in social neuroscience is whether interpersonal interactions are processed differently by the brain depending on the bodily characteristics of the interactor, i.e., their physical appearance. To address this issue, we engaged participants in a minimally interactive task with an avatar either showing bodily features or not while recording their brain activity using electroencephalography (EEG) in order to investigate indices of action observation and action monitoring processing. Multivariate results showed that bodily compared with nonbodily appearance modulated parieto-occipital neural patterns throughout the entire duration of the observed movement and that, importantly, such patterns differ from the ones related to initial shape processing. Furthermore, among the electrocortical indices of action monitoring, only the early observational positivity (oPe) was responsive to the bodily appearance of the observed agent under the specific task requirement to predict the partner movement. Taken together, these findings broaden the understanding of how bodily appearance shapes the spatiotemporal processing of an interactor's movements. This holds particular relevance in our modern society, where human–artificial (virtual or robotic) agent interactions are rapidly becoming ubiquitous.

**Key words:** action monitoring; action observation; biological motion; EEG; interpersonal interactions

Received Sept. 8, 2024; revised Jan. 31, 2025; accepted Feb. 2, 2025.

The authors declare no competing financial interests.

Author contributions: U.G.P., Q.M., V.E., and M.C. designed research; U.G.P. and Q.M. performed research; U.G.P. and Q.M. analyzed data; U.G.P., Q.M., V.E., and M.C. wrote the paper.

U.G.P. was supported by Sapienza Grant Avvio alla Ricerca, Tipo 1 (AR1221816C75DC62). V.E. was supported by Bando Giovani Ricercatori, Ministero italiano della Salute, GR-2021-12372923. M.C. was funded by Sapienza Grant Progetti Medi 2022 (RM1221816C827130) and by Grant Medie Attrezzature 2021 (MA22117A8A97CD42).

\*U.G.P. and Q.M. contributed equally to this work.

*Continued on next page.*

## Significance Statement

During interpersonal motor interactions, the observation and monitoring of other's actions are essential mechanisms depending on two interconnected brain networks. Whether the neurophysiological signatures of action observation and monitoring are modulated by the appearance of an interacting partner remains an open question of particular relevance in order to tackle how the brain interfaces with artificial agents. In the present study we used highly ecological virtual stimuli in a minimally interacting scenario as well as univariate and multivariate EEG analyses to broaden our understanding of the influence of bodily appearance on the spatiotemporal processing of biological movements in the AON and in the action monitoring system.

## Introduction

Observing and monitoring others' actions are essential functions for humans' social life and support interpersonal coordination (Ridderinkhof et al., 2004; Knoblich and Sebanz, 2008; Vesper et al., 2010; Cardellicchio et al., 2021; Sacheli et al., 2022; Musco et al., 2023). Neuroimaging studies have identified, among others, two brain networks involved

in the prediction and monitoring of observed actions, namely, the action observation network (AON, Grafton and Hamilton, 2007) and the action monitoring system (AMS, Miltner et al., 2004; Ridderinkhof et al., 2004; van Schie et al., 2004; De Bruijn et al., 2007; Malfait et al., 2010; Cohen, 2016).

The AON comprises parietal, premotor, motor, and visual areas (Hardwick et al., 2018). Within the visual nodes of the network, the extrastriate body area (EBA) is thought to be specifically involved in human body form processing (Downing et al., 2001; Urgesi et al., 2006; Pitcher et al., 2009; Moreau et al., 2019, 2020; Fusco et al., 2022) and the superior temporal sulcus (STS) is thought to be crucial for the processing of dynamical aspects of visual motion (i.e., biological motion; Allison et al., 2000; Puce and Perrett, 2003; Saygin et al., 2012; Lingnau and Downing, 2015; Pitcher and Ungerleider, 2021). EEG studies have identified the alpha/mu desynchronization (8–13 Hz) over centroparietal electrodes as the main EEG marker of the AON activity, since it occurs during action observation and is similar to the markers of action execution (Cochin et al., 1999; Pineda, 2005; Coll et al., 2017).

The AMS has been described extensively for its role in the control of executed actions and has been also associated to the detection of sudden and/or unexpected changes in the individual actions or in those of others (e.g., a change in the trajectory of an observed movement; Miltner et al., 2004; De Bruijn et al., 2007; Shane et al., 2008; Malfait et al., 2010; Moreau et al., 2020, 2022). This system is involved in a constant process of prediction and adaptation (Ullsperger et al., 2014; Ullsperger, 2017; Wilken et al., 2023) and comprises the posterior part of the medial frontal cortex, specifically the anterior cingulate cortex (ACC), the pre-supplementary motor area (pre-SMA), and the adjacent dorsomedial prefrontal cortex (dmPF; Ridderinkhof et al., 2004; Nee et al., 2011). Previous EEG studies identified two event-related potentials (ERPs) linked to action monitoring processes, the error-related negativity (ERN; Falkenstein et al., 1991; Gehring et al., 1993; Taylor et al., 2007; Pezzetta et al., 2018, 2023) and the error positivity (Pe; Falkenstein et al., 2000). These ERPs are associated with the detection of both errors committed by the self and errors observed in others. In the latter case, they are referred to as observation-ERN (oERN) and observation-Pe (oPe; van Schie et al., 2004; Spinelli et al., 2018). Furthermore, action monitoring is indexed in the time–frequency domain by the so-called midfrontal theta rhythm (Cohen, 2011), a synchronization in the theta band (4–8 Hz) over midfrontal sites (i.e., electrode FCz) that tends to last a few hundred milliseconds (200–400) after an error is detected and can be easily observed in individual subjects. Both time and time–frequency domain indices have been observed during interpersonal motor interactions (Moreau et al., 2020, 2022), when sudden changes (i.e., corrections) in the movement of a partner were to be detected by individual participants in order to fulfill a joint task (Moreau et al., 2020, 2022) or during musical duets (Paas et al., 2021). One pending question concerning the interaction between action observation/monitoring and interpersonal motor coordination is whether biological visual motion can induce activity in the AON and AMS independently from the visual appearance of the observed/monitored agent. Moreover, previous studies did not investigate the modulation of these processes in dynamic social contexts. This is crucial to better understand the neural system supporting social observation as well as dynamic social interaction processes where reciprocal and mutual adaptation are involved (Tidoni et al., 2022).

Since recent results have also shown that theta oscillatory activity in occipitotemporal nodes of the AON (i.e., LOTC) are in phase with the activation of the action monitoring system during social (human–avatar) interactions (Moreau et al., 2020), we also aim at investigating if and how the appearance of an interactor during online (minimal) interactions affect the behavioral ability to predict their movements (Zucchini et al., 2023) and the neural patterns related to action monitoring. In order to do so, we engaged human participants in a minimally interactive version of a visuomotor interpersonal interaction task (Sacheli et al., 2015; Moreau et al., 2020; Era et al., 2023) where subjects had to monitor the movement of a virtual partner (i.e., VP) at different levels of interactivity. The appearance of the VP and the movement trajectory were manipulated so that participants observed either a body or a set of moving dots (Appearance factor), while the trajectory of the body/dots movement changed or remaining the same (Correction factor). This design is similar to that used in previous studies investigating the activation of the AMS during interpersonal motor interactions. We focused on the modulation by the appearance of an interactor on (1) the neurophysiological indices of error-monitoring (i.e., oERN, oPE, and midfrontal

Correspondence should be addressed to Vanessa Era at [vanessa.era@uniroma1.it](mailto:vanessa.era@uniroma1.it).

Copyright © 2025 Pesci et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

theta observed in response to movement corrections), (2) the alpha/mu desynchronization over centroparietal electrodes, and (3) whole-scalp single-trial data (analyzed through a multivariate pattern analysis—MVPA) to examine if and at which point in time the brain distinguishes bodily from nonbodily movements.

## Materials and Methods

### Participants

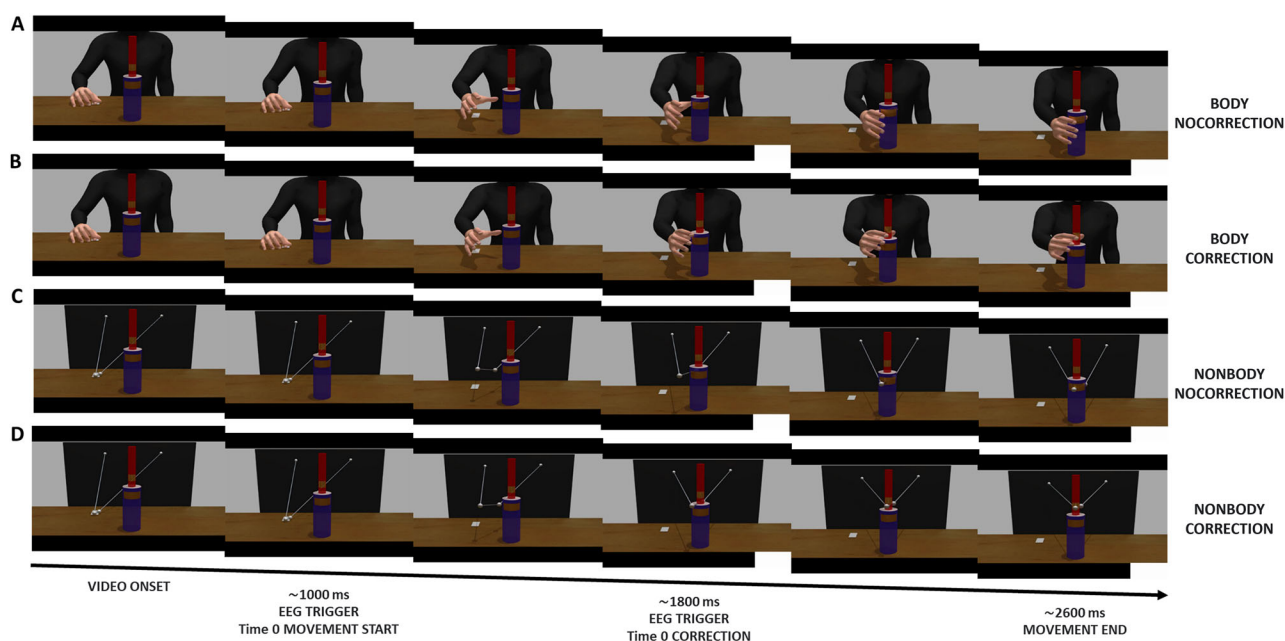
Twenty-one participants took part in the experiment. The sample size was based on a power analysis, performed with the software More Power (Campbell and Thompson, 2012). More specifically, we indicated as expected effect size the partial eta squared value obtained by Moreau et al. (2020) (i.e., 0.33). The output indicates that a  $2 \times 2 \times 2$  within-subject design, a power of 0.80 and a partial eta squared of 0.33, requires a sample size of 18 participants. All participants were right-handed with normal or corrected-to-normal vision. Participants were naive as to the aim of the experiment and were informed of the purpose of the study only after all the experimental procedures were completed. One participant was removed from all results due to technical issues during the recording. Thus, the final sample includes 20 participants (10 females; mean age, 24;  $SD = 2.5$ ). Human subjects were recruited at the Department of Psychology and IRCCS Fondazione Santa Lucia in Rome, and they signed informed consents approved along with all experimental procedures by the Ethics Committee of the Fondazione Santa Lucia (Rome, Italy), approval number Prot. CE/prog 835. The study was performed in accordance with the 2013 Declaration of Helsinki.

### Experimental stimuli and set-up

Participants sat in front of a desk and observed visual stimuli on a  $1.024 \times 768$  resolution LCD monitor placed at  $\sim 60$  cm from their eyes. Using a keyboard connected to the PC, participants were asked to press a button (up/down arrow) with the index/middle finger of their right hand at the exact same time when an virtual partner (VP—facing the participant) on the screen, or two dots in a control condition, touched a bottle-shaped object in front of him. The bottle-shaped object is constituted of two superimposed cylinders of different diameters (see below for details on task). The variability of the appearance of the interactor allowed us to explore the influence on the EEG correlates of the presence of a body on action observation.

The keyboard's press reaction times were recorded using E-Prime 2 Professional software (Psychology Software Tools). The VP's index–thumb grasping time was measured trial-by-trial by means of a photodiode placed on the screen sending a signal recorded by means of a TriggerStation (BrainTrends). The photodiode was triggered by a white dot displayed on the screen (not visible to the participants) during the clip frame corresponding to the instant when the avatar grasped (or the dots reached) the virtual object.

Video stimuli were adapted from previous studies (Gandolfo et al., 2019) in which reaching and grasping movements of a virtual character are based on the trajectories of real human actors performing a number of grasping movements (Fig. 1).



**Figure 1.** Examples of the sequence of frames for each combination of the Appearance factor with the Correction factor: (A) Body-NoCorrection; (B) Body-Correction; (C) NonBody-NoCorrection; and (D) NonBody-Correction.

As a control condition, the same motion kinematics of the human shoulders, right index and thumb were implemented in a set of four dots connected by lines in a manner aimed at reducing the perception of the shape of a human body and create a set of nonhumanoid stimuli (hereafter NonBody stimuli; Fig. 1). Kinematic features of the actors were recorded using a SMART-D motion capture system [MoCAP; Bioengineering Technology & Systems (B|T|S)] attaching infrared reflective markers (5 mm diameter) to an actor reaching and grasping with his right hand a cylinder on a table. Four infrared cameras with wide-angle lenses placed ~100 cm away from each of the four corners of the table captured at 100 Hz the movement of the markers in the 3D space (Tieri et al., 2015). The kinematics were eventually transferred on the virtual VP's bones by means of Motion Builder 2011 (Autodesk).

Stimuli were thus divided equally in two main categories representing (1) a humanoid partner (Fig. 1A,B, Body), applying the trajectories to a Caucasian male character, and (2) a nonhumanoid stimulus (Fig. 1C,D, NonBody), composed of four spheres representing the right and left shoulder, thumb, and index finger of the right hand, connected by white lines. Note that the character was presented only with the upper body from the shoulders down, since variables related to face were to be ignored by participants. The complete sample of clips comprised 12 different grasping movements (Grip factor), half of them ending with a precision grip (grasping the top part of the bottle-shaped object) and half ending with a power grip (on the bottom part). Each movement had a version with the Body and one with the NonBody (Appearance factor). The Appearance factor is fundamental for the present purposes as it represents the main manipulation of the stimuli to test the role of physical appearance of a human body in triggering behavioral and neural resources associated to action observation, prediction, and error processing.

Crucially for the hypothesis regarding action monitoring, in 30% of the trials the movements included an online correction of trajectory by the VP or the dots (Fig. 1B,D, Correction Factor), which switched from a precision to a power grip (or vice versa) during the reaching phase at different times for each video. In order to lock on the EEG signal the exact time of this correction, the photodiode was triggered by another white dot, identical to the abovementioned one, flashed during the clip frame corresponding to the start of the online correction, or in the corresponding time in the NoCorrection trials.

### Experimental task

Participants were asked to press a button corresponding either to the upper part of the bottle-shaped object (up arrow) or the bottom part of it (down arrow) as synchronously as possible with the touch time of the VP or of the dots on the bottle in the virtual environment. Participants had to follow an auditory instruction concerning the response to be executed, which was delivered prior to each trial via headphones. Specifically, interactions occurred in two different conditions (Interactivity factor): (1) a Cued condition, where participants knew in advance which key they had to press and therefore only had to track the timing of the VP's/dots' grasping time, and (2) an Interactive condition, where they did not know which key they had to press and had to adapt to the movement of the VP/dots, therefore having to anticipate and monitor it online to coordinate not only temporally but also spatially with it.

In detail, instructions for the Cued condition consisted of either a high pitch (indicating that participants had to press the upper arrow on their keyboard) or a low pitch sound (indicating that participants had to press the lower arrow on their keyboard). For the Interactive condition, instructions consisted in a sound saying "Ugua," shortened for Italian "Uguale" ("Same"), asking the subjects to press the key matching the VP's behavior or dots contact site (e.g., if the avatar grasped, or the dots reached, the upper part of the object, participant had to press the up arrow), or "Oppo," shortened for Italian "Opposto" ("Opposite"), asking the subjects to press the key opposite to the VP's behavior or dots contact site (e.g., if the avatar grasped, or the dots reached, the upper part of the object, participant had to press the down arrow). Instructions were presented 900 ms before the start of the VP's movement, and their duration was fixed to ensure they were all the same length, maintaining consistency across experimental conditions. This timing ensured that the evoked neural response to the instruction had fully dissipated before the movement began. As a result, no trigger corresponding to the instruction's onset was sent to the EEG. At the end of each trial, participants received feedback on their performance according to both synchrony (ms) and correct pressing site (according to the instruction received at the beginning).

The task presented above is a revision of an ecological and controlled human-avatar "Joint Grasping" task (Candidi et al., 2017; Sacheli et al., 2018; Era et al., 2020; Moreau et al., 2020), which has been shown to recruit similar behavioral and neural processes called into play during human-human interaction (Candidi et al., 2015; Era et al., 2018).

Participants performed in total eight 100-trial blocks (4 blocks of the Body condition, 4 for the NonBody condition, each divided in two successive Interactive and two successive Cued ones, presented in a counterbalanced order between participants). Stimuli presentation and randomization were controlled by E-Prime 2 Professional Software (Psychology Software Tools).

### Behavioral data

The Interpersonal Synchrony with the VP's/dots' grasping was considered as the main behavioral measure, computed as the absolute value of the time delay between the participant's pressing of the keyboard button and the VP's/dots' bottle touch time. This is believed to show the success of human-avatar interaction achieved by monitoring and predicting the movements of the stimulus. At the group level, we checked for the presence of outliers in our sample by investigating whether any of our participants' Interpersonal Synchrony across all conditions was 2.5 standard deviations above the group mean. According to this control, all 20 subjects were included in the final analyses.

## EEG recording and preprocessing

EEG signals were recorded and amplified using a Neuroscan SynAmps RT amplifiers system (Compumedics). These signals were acquired from 58 tin scalp electrodes embedded in a fabric cap (Electro-Cap International), arranged according to the 10–20 system. The EEG was recorded from the following channels: Fp1, Fpz, Fp2, AF3, AF4, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FC5, FC3, FC1, FCz, FC2, FC4, FC6, T7, C5, C3, C1, Cz, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, AF7, POz, AF8, PO4, PO8, O1, Oz, and O2. Horizontal electro-oculogram (HEOG) was recorded bipolarly from electrodes placed on the outer canthi of each eye and signals from the left earlobe were also recorded. All electrodes were physically referenced to an electrode placed on the right earlobe and were algebraically rereferenced off-line to the average of both earlobe electrodes. Impedance was kept below 5 K $\Omega$  for all electrodes for the whole duration of the experiment, amplifier hardware bandpass filter was 0.1–200 Hz and sampling rate was 1,000 Hz. To remove ocular artifacts (eyeblinks and saccades), a blind source separation method, independent component analysis (ICA; Jung et al., 2000), was applied on continuous raw signal. Artifactual components were removed based on the topography and the explained variance, and data were visually inspected to check the efficient removal of blinks and saccadic movement correlates from the EEG signal. The signal was then low-pass filtered at 45 Hz and segmented into epochs of 8,000 ms around the trigger corresponding to the start of the video (Go—trigger). To investigate the neural patterns related to action observation or action monitoring, the time 0 was manually moved to either the frame when the movement of the VP or of the dots started (monitor t0) or to the frame when a correction occurred (or the equivalent one in No-Correction trials—error t0), respectively. Segmented data were visually inspected and remaining trials containing ocular and muscular artifacts as well as trials with 0 behavioral accuracy were removed from further analysis. The average number of trials per participant after artifact rejection was 351 trials for Body (with 39 Interactive-Correction, 127 Interactive NoCorrection, 59 Cued-Correction, 126 Cued-NoCorrection) and 347 trials for NonBody (with 39 Interactive-Correction, 121 Interactive NoCorrection, 59 Cued-Correction, 128 Cued-NoCorrection). Importantly, in the analysis of action monitoring activity, the epochs were restricted to –2,000 to +2,000 ms around the trigger corresponding to the VP/dots' Correction/NoCorrection frame. All preprocessing and further univariate analyses were performed using the FieldTrip (version 2019-04-03) routines (Donders Institute; Oostenveld et al., 2011) in Matlab R2019a (The MathWorks).

## EEG analysis

### Univariate analysis

In order to investigate whether the neural correlates of action monitoring would be modulated by the appearance of the interacting partner (Appearance factor), we analyzed the event-related potentials (ERPs) over frontocentral electrodes (i.e., FCz) time locked to the correction of the observed movement's trajectory (error t0), as well as the activity in the time-frequency domain focused on midfrontal theta. We decided to focus these analyses on the activity over electrode FCz based on previous studies employing the same (and similar) paradigm (Moreau et al., 2020; Spinelli et al., 2022) as well as following visual inspection of the topography of the time and time–frequency responses to the correction of the observed movements (Fig. 3, bottom panel). Moreover, in order to investigate whether the appearance of the interacting partner would modulate earlier neural processes related to action observation, we also analyzed the alpha/mu desynchronization over centroparietal electrodes (Cochin et al., 1999; Pineda, 2005; Tucciarelli et al., 2015) time locked to the start of the observed movement (movement t0).

**ERPs.** Before ERP averaging across subjects for each condition, EEG time-series were high-pass filtered at 1 Hz to remove slow drifts from the data, hence reducing the contribution of slow potentials. Since we found no oERN in response to the observation of a partner's correction in any condition (see Discussion), only the early oPe component was further analyzed being quantified as the mean amplitude in the time window between 250 and 400 ms over electrode FCz.

**ERD/S.** After preprocessing, as in the time-lock analysis, we high-pass filtered the segmented EEG signal at 1 Hz. Moreover, to remove the effects of ERPs in the time–frequency domain we subtracted the mean evoked response from each single trial, thus removing phase-locked activity (Sauseng et al., 2007). Each epoch was then put in the frequency domain using Hanning-tapered window (Cohen, 2014) with a 50 ms time resolution, obtaining single-trial power representations for frequencies in a range from 1 to 40 Hz in steps of 1. The obtained induced power was averaged over trials for each condition and each subject, and the grand averages across subjects were displayed as event-related desynchronization/synchronization (ERD/ERS) with respect to a baseline period ranging from –0.2 to 0 ms before the correction occurred (for the error-locked analysis), and from –1,300 to –1,000 ms before the start of the movement (for the action observation analysis). Only the epochs when the virtual partner did not correct its trajectory (NoCorrection trials) were included in the action observation analysis.

For midfrontal theta, we extracted the ERD/ERS for the theta band (3–7 Hz) between 200 and 500 ms after the correction (or the corresponding frame in the NoCorrection trials) and analyzed the modulation of power over FCz. For the action observation analysis, we extracted the ERD/ERS for the alpha (8–13 Hz) band from 0 to 1,600 ms (average movement time) over centroparietal electrodes (CPz/1/2/3/4). As a control analysis (Coll et al., 2017), we also extracted the ERD/ERS in the same frequency band and time window over an equal number of occipital electrodes (Oz/1/2/PO7/8).

### Multivariate analysis

To better identify the neural patterns of action observation that were sensitive to the appearance of the observed interactor (i.e., a Body or a NonBody), we adopted a multivariate approach. Multivariate analyses are more sensitive than univariate analyses (Haxby et al., 2001; Tucciarelli et al., 2015), since they assume that the processing of different stimulus category has different neural patterns associated to be exploited; thus they treat whole-brain sensor/source-level data as response patterns rather than investigating changes between average responses for each single sensor/source (as in the univariate approach; Grootswagers et al., 2017). Analyses were performed using the MVPA-Light toolbox (Treder, 2020) in Matlab R2019a. Prior to every classification, data were resampled at 250 Hz and normalized using z-scores to center and scale the training data, providing numerical stability (Treder, 2020). Across all classification analyses, if a class had fewer trials than another, we corrected the imbalance by undersampling the over-represented condition (i.e., randomly removing trials). Then, using the preprocessing option `mv_preprocess_average_samples`, training data (i.e., EEG trials) from the same class (i.e., Body/NonBody) were randomly split into five groups and averaged, so that the single-trial dataset was replaced by five averaged trials (i.e., samples). The classifications were then run using these averaged samples, as this step is known to increase the signal-to-noise ratio (Grootswagers et al., 2017; Smith and Smith, 2019; Treder, 2020).

**Classification across time.** We first performed a binary classification analysis in time (−1,200 to 2,000 ms around movement start), using Body/NonBody samples. To do so, a linear discriminant analysis (LDA) classifier was implemented using the MVPA-Light toolbox. For linearly separable data, an LDA classifier divides the data space into  $n$  regions, depending on the number of classes, and finds the optimal separating boundary between them using a discriminant function to find whether the data fall on the decision boundary (i.e., 50% classification accuracy) or far from it (i.e., >50% classification accuracy). A  $k$ -fold cross-validation procedure in which samples ( $5 \times 2$  classes = 10 total samples) were divided into 10 folds was repeated 10 times so that each sample was either used for training or testing at least once, and the classifier's % accuracy (acc) was used to assess decoding performance over time in each comparison. Statistical significance (i.e., performance significantly above 0.5 chance level) was assessed through nonparametric cluster-based permutation tests (Maris and Oostenveld, 2007), using “maxsum” for cluster corrections and Wilcoxon test for significance assessment.

In order to further understand how the different patterns related to observing Body and NonBody motion evolve over time, we adopted the temporal generalization method (King and Dehaene, 2014). Specifically, we tested whether the neural representations carrying the classifier performance were stable (generalizable) across time or rather they differed. In order to do so, we trained our classifier on each single time point and tested it at all time points in the same time window previously selected for the classification across time. This provided a temporal generalization matrix (time  $\times$  time) with accuracy values tested for statistical significance.

Moreover, to investigate what electrodes contributed most to the classification over time, we performed a searchlight analysis using two neighboring matrices for time points and electrodes, respectively. Searchlight analysis is one approach to localize multivariate effects, as it strikes a balance between localization and statistical power (Kriegeskorte et al., 2006; Treder, 2020). Thus, in this analysis each electrode/time point and its direct neighbors acted as features for the classification, resulting in a channels  $\times$  time points matrix of accuracy scores tested for statistical significance. By plotting the results from this matrix also on a spatial topography averaged in specific time windows of interest, we then visualized which electrodes carried the most weight in the temporal decoding.

Lastly, we were interested in finding whether the neural patterns related to Body and NonBody motion were classifiable under different interactive scenarios (i.e., in the Interactive condition compared with the Cued one). Thus, we performed a multiclass LDA in time to classify the EEG activity in the time window from movement onset (0 s) to its average end (1,600 ms) for each of the four classes (Body–Interactive, Body–Cued, NonBody–Interactive, NonBody–Cued). The accuracy values across testing folds of all repetitions were then averaged and presented on a confusion matrix to assess the probability of the classifier to in/correctly assign classes.

### Data handling and statistics

Interpersonal Synchrony values, error-locked ERP mean amplitudes in the oPe time window (250–400 ms), and midfrontal theta synchronization values were analyzed through a  $2 \times 2 \times 2$ , within-subject, repeated-measures ANOVA with Appearance (Body/NonBody), Interactivity (Interactive/Cued), and Correction (Correction/NoCorrection) as within-subject factors. Conversely, while our primary interest concerning movement-locked time–frequency induced ERD/ERS focused on the interaction between the factors Interactivity and Appearance, we initially analyzed it by also including a factor Cluster (Occipital/Centro-Parietal) through a  $2 \times 2 \times 2$ , within-subject, repeated-measures ANOVA with Appearance (Body/NonBody), Interactivity (Interactive/Cued), and Cluster (Occipital/Centro-Parietal) as within-subject factors. After this first exploratory analysis aimed at testing the topographical distribution of the effects of Appearance and Interactivity, movement-locked time–frequency induced ERD/ERS values were analyzed through two separate  $2 \times 2$ , within-subject, repeated-measures ANOVAs with Appearance (Body/NonBody) and Interactivity (Interactive/Cued) as within-subject factors focusing, respectively, on the activity over five centroparietal electrodes and on the activity over five control occipital electrodes (similarly to Coll et al., 2017). Since our main hypothesis did not focus on Interaction Type (Complementary/Imitative) and Movement Type (Up/Down), these factors were collapsed in order to have a higher

number of trials for each condition. Nevertheless, in order to verify that the different instructions in the Interactive condition (i.e., “Oppo” or “Ugua”) and thus different interactive modalities (i.e., “Opposite” or “Same” movements) did not affect behavioral performance, we checked for any significance of the Interaction Type factor, which confirmed that subjects’ performance did not differ based on whether they pressed the final button in the same or opposite direction compared with the final target of the VP/dots ( $F_{(1,19)} = 1.752$ ,  $p = 0.205$ ,  $\eta_p^2 = 0.105$ ). Normality and homoscedasticity assumptions were checked using the Shapiro–Wilk test, revealing that, only for behavioral data, NonBody data were not normally distributed compared with the Body ones.

All statistical analyses were performed in Matlab R2019a, JASP software, and R using the DABEST Package (Ho et al., 2019). Post hoc correction for multiple comparisons (Ryan, 1959) was conducted applying Bonferroni’s method (Maris and Oostenveld, 2007) to the interactions that resulted as significant ( $p < 0.05$ ) from the ANOVA.

## Results

### Behavioral

The 2 Appearance (Body/NonBody)  $\times$  2 Interactivity (Interactive/Cued)  $\times$  2 Correction (Correction/NoCorrection) ANOVA showed that Interpersonal Synchrony was significantly modulated by the Correction factor (Fig. 2A;  $F_{(1,19)} = 39.71$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.676$ ). Specifically, Interpersonal Synchrony was worse when the stimulus (either virtual-partner/dots-pattern) did not correct its trajectory ( $M = 127.40$  ms,  $SD = 41.83$ ) rather than when a correction occurred ( $M = 106.22$  ms,  $SD = 29.05$ ). The analysis also showed a significant interaction between the Appearance and Correction factors ( $F_{(1,19)} = 14.32$ ,  $p = 0.001$ ,  $\eta_p^2 = 0.43$ ). Post hoc tests indicated that Interpersonal Synchrony was worse in the NonBody-NoCorrection condition compared with the NonBody-Correction one ( $p < 0.001$ ) and the Body-NoCorrection one ( $p = 0.020$ ). Also the interaction between the Interactivity and the Correction factor resulted to be significant ( $F_{(1,15)} = 13.07$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.41$ ), with interpersonal Synchrony being worse in NoCorrection trials compared with Correction trials both in the Interactive ( $p < 0.001$ ) and in the Cued ( $p = 0.020$ ) condition (Fig. 2B). These results suggest that, independently from the level of interactivity implied in the condition (i.e., Interactive or Cued), a correction in the trajectory of the observed movement might have prompted subjects’ attention, improving their ability to predict VP’s/dots’ reaching time. Regarding the Appearance factor, we observed that the humanoid appearance of the interactor seemed to improve the performance of the subjects only during NoCorrection trials, i.e., when subjects’ performance was generally worse.

### EEG results

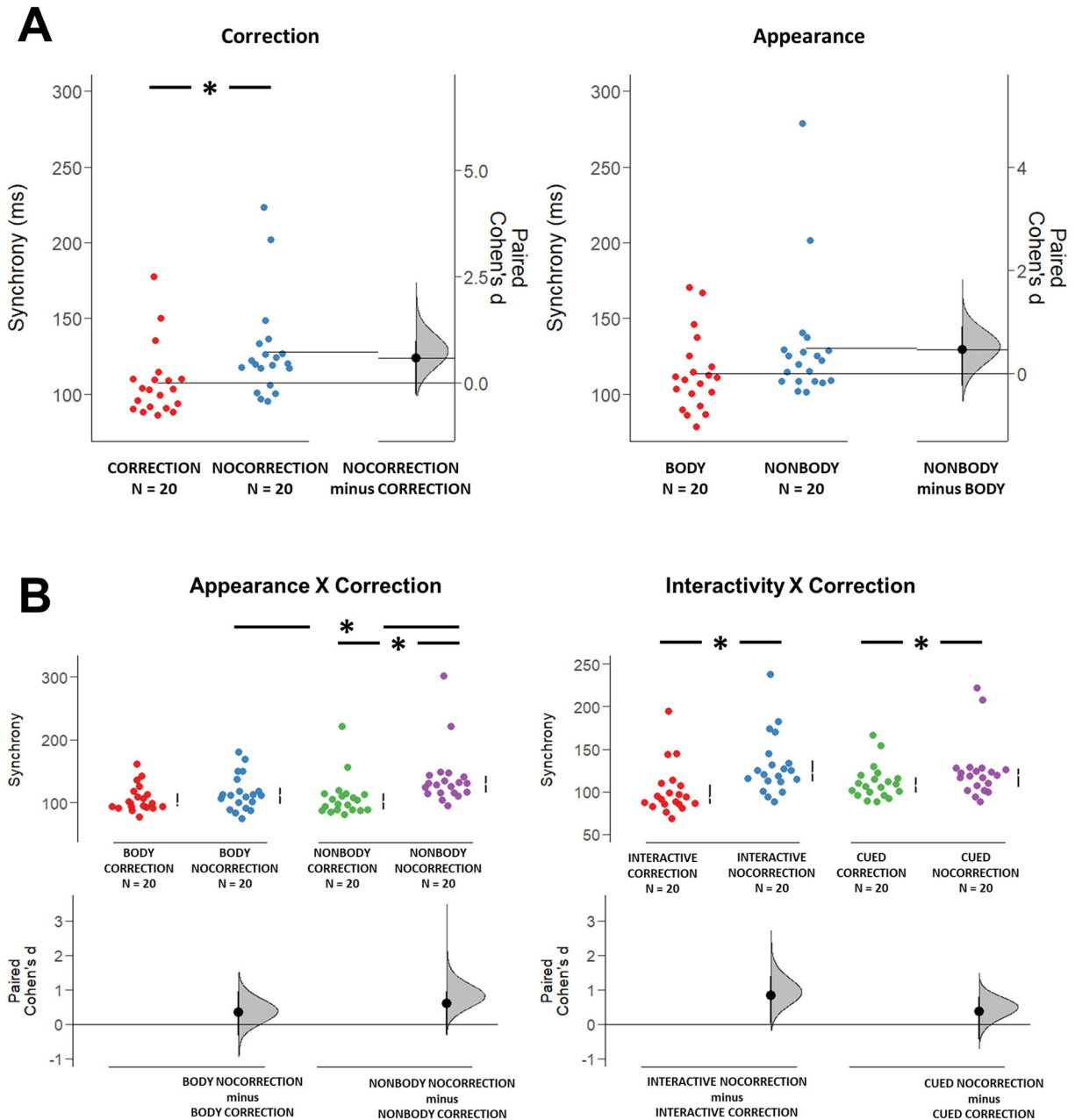
#### ERPs—oPe

The 2 Appearance (Body/NonBody)  $\times$  2 Interactivity (Interactive/Cued)  $\times$  2 Correction (Correction/NoCorrection) ANOVA on oPe amplitudes revealed a main effect of the Correction factor ( $F_{(1,19)} = 32.119$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.628$ ) showing that the early oPe amplitude was larger for Correction ( $M = 1.015$ ,  $SD = 0.86$ ) compared with NoCorrection trials ( $M = 0.127$ ,  $SD = 0.192$ ). There was also a significant main effect of Interactivity ( $F_{(1,19)} = 18.438$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.492$ ), with larger oPe for Interactive ( $M = 0.745$ ,  $SD = 0.95$ ) compared with Cued trials ( $M = 0.398$ ,  $SD = 0.46$ ). oPe amplitude was also larger for Body ( $M = 0.626$ ,  $SD = 0.86$ ) compared with NonBody trials ( $M = 0.517$ ,  $SD = 0.66$ ), as the Appearance factor approached significance threshold ( $F_{(1,19)} = 4.298$ ,  $p = 0.052$ ,  $\eta_p^2 = 0.184$ ). The three factors interacted significantly ( $F_{(1,19)} = 8.787$ ,  $p = 0.008$ ,  $\eta_p^2 = 0.316$ ). Post hoc tests indicated that the early oPe amplitude was significantly larger during Body-Interactive-Correction trials than during Body-Interactive-NoCorrection trials ( $p < 0.001$ ), Body-Cued-Correction trials ( $p < 0.001$ ), and NonBody-Interactive-Correction trials ( $p < 0.001$ ). Moreover, the early oPe amplitude recorded during NonBody-Interactive-Correction trials was larger than during NonBody-Interactive-NoCorrection trials ( $p < 0.001$ ), as well as than during NonBody-Cued-Correction trials ( $p = 0.010$ ; Fig. 3). This pattern of results is in line with previous results (Moreau et al., 2020) showing a modulation of the Pe amplitude by higher-order task-related factors (Interactivity factor), and they add new information regarding the electrophysiological indices of action monitoring when observing a bodily movement compared with a nonbodily one (Appearance factor).

Time–frequency analyses—midfrontal theta and centroparietal alpha/mu de/synchronization.

#### Correction-locked midfrontal theta over central electrodes

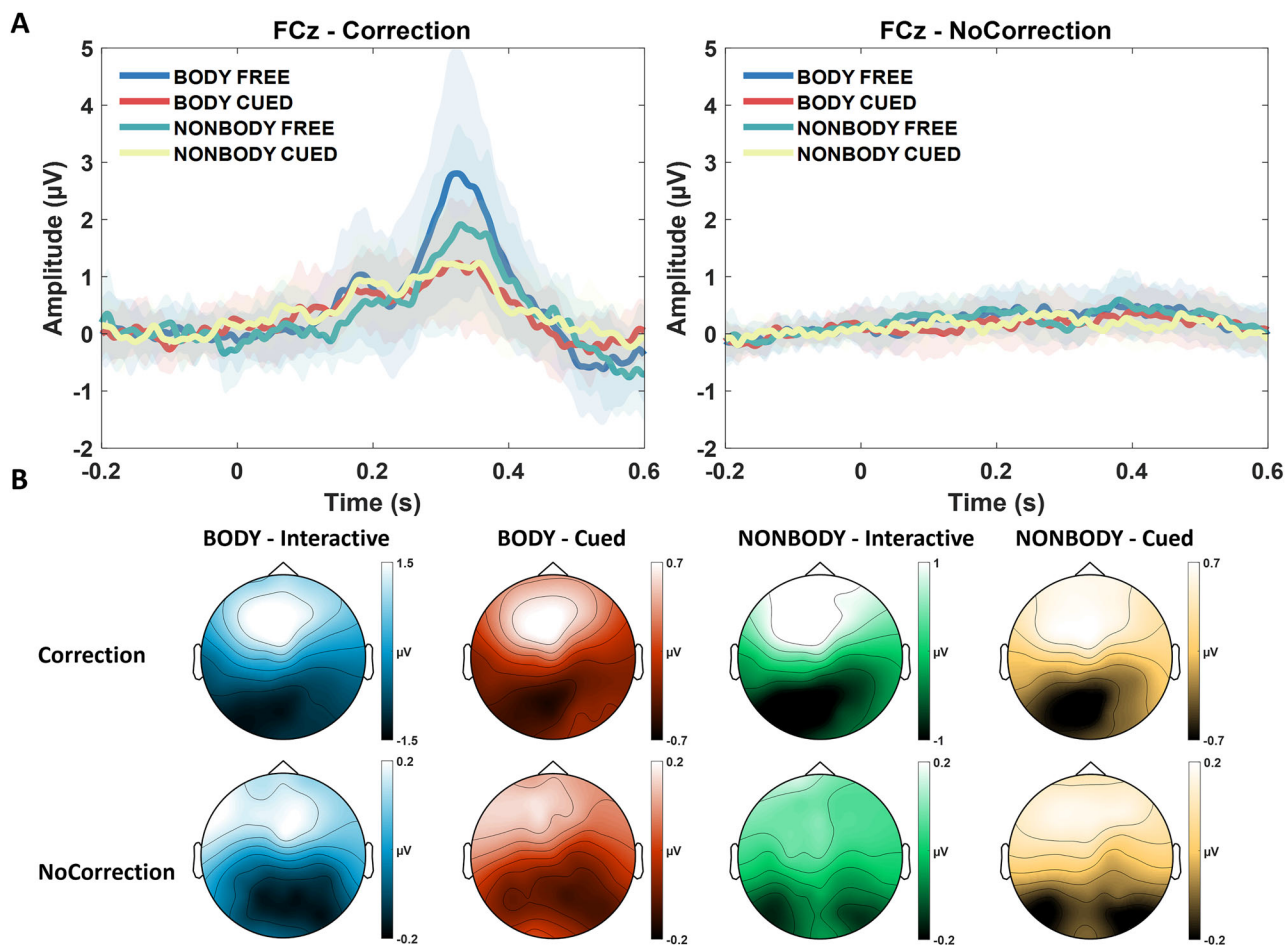
The 2 Appearance (Body/NonBody)  $\times$  2 Interactivity (Interactive/Cued)  $\times$  2 Correction (Correction/NoCorrection) ANOVA showed a main effect of Correction, with higher theta synchronization for Correction ( $M = 0.564$ ,  $SD = 0.479$ ) compared with NoCorrection trials ( $M = 0.046$ ,  $SD = 0.138$ ;  $F_{(1,19)} = 59.822$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.759$ ). Moreover, the Interactivity factor reached statistical significance, with larger theta synchronization during Interactive ( $M = 0.417$ ,  $SD = 0.528$ ) compared with Cued trials ( $M = 0.193$ ,  $SD = 0.283$ ;  $F_{(1,19)} = 19.790$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.510$ ). The Correction and Interactivity factors interacted significantly ( $F_{(1,19)} = 12.365$ ,  $p < 0.002$ ,  $\eta_p^2 = 0.394$ ). Post hoc tests indicated that theta synchronization was larger for Interactive-Correction trials compared with all other condition (all  $ps < 0.001$ ) as well as for Cued-Correction trials compared with Cued-NoCorrection trials ( $p < 0.001$ ). These results are in line with previous ones from our group (Moreau et al., 2020), but contrary to our hypothesis, no statistical difference in theta synchronization was found between Body and NonBody trials (Fig. 4).



**Figure 2.** Top row, left panel, Significant modulation of Interpersonal Synchrony values by the Correction Factor (left panel,  $F_{(1,19)} = 39.71$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.676$ ). Right panel, Nearly significant modulation of Interpersonal Synchrony by the Appearance factor (right panel,  $F_{(1,19)} = 3.61$ ,  $p = 0.07$ ,  $\eta_p^2 = 0.160$ ). Bottom row, left panel, Significant interaction between Appearance and Correction factor ( $F_{(1,19)} = 14.32$ ,  $p = 0.001$ ,  $\eta_p^2 = 0.43$ ). The post hoc test indicated that Interpersonal Synchrony was worse in the NonBody-NoCorrection condition compared with the NonBody-Correction one ( $p < 0.001$ ) and the Body-NoCorrection one ( $p = 0.020$ ). Right panel, Significant interaction between Interactivity and Correction factor ( $F_{(1,15)} = 13.07$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.41$ ). The post hoc test indicated that Interpersonal Synchrony was worse in NoCorrection trials compared with Correction trials both in the Free ( $p < 0.001$ ) and in the Cued ( $p = 0.020$ ) condition.

*Movement-monitoring power over centroparietal electrodes*

To investigate the alpha/mu (8–13 Hz) desynchronization related to action observation over centroparietal electrodes, compared with visual alpha desynchronization over occipital electrodes as control, we run a 2 Cluster (Occipital/Centro-Parietal)  $\times$  2 Appearance (Body/NonBody)  $\times$  2 Interactivity (Interactive/Cued), within-subjects, repeated-measures ANOVA. Both the Cluster ( $F_{(1,19)} = 38.324$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.668$ ) and the Interactivity factor ( $F_{(1,19)} = 5.574$ ,  $p = 0.029$ ,  $\eta_p^2 = 0.227$ ) reached statistical significance but their interaction did not ( $F_{(1,19)} = 2.586$ ,  $p = 0.124$ ,  $\eta_p^2 = 0.120$ ). Thus, we further conducted two separate 2 Appearance  $\times$  2 Interactivity ANOVAs focusing on alpha/mu desynchronization over, respectively, only the centroparietal cluster or the occipital cluster. As shown in Figure 5A, over centroparietal electrodes,



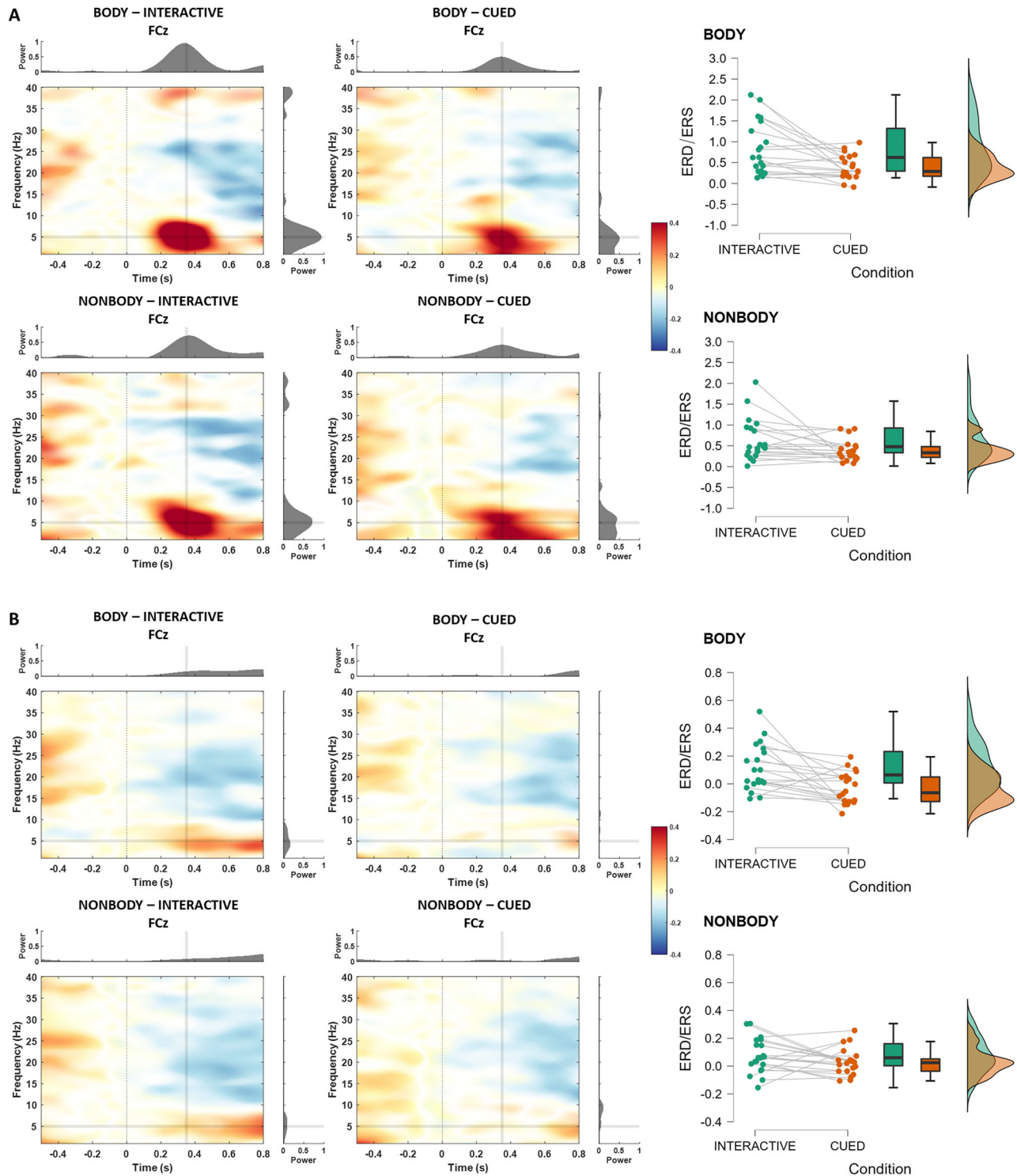
**Figure 3.** **A**, Grand averages of the early oPe component over FCz in all experimental conditions with (left panel) and without (right panel) correction. Data are time-locked to the correction of the Body/dots (or the equivalent frame when no correction occurred). **B**, Topographies of the early oPe component (from 250 to 400 ms after the correction of the observed movement's trajectory) for each condition.

the Interactivity factor reached statistical significance, with stronger alpha desynchronization during Interactive ( $M = -0.287$ ,  $SD = 0.171$ ) compared with Cued trials ( $M = -0.224$ ,  $SD = 0.147$ ;  $F_{(1,19)} = 5.233$ ,  $p = 0.034$ ,  $\eta_p^2 = 0.216$ ). Similar to the midfrontal theta results, no significant difference was found between the alpha desynchronization during the observation of Body and NonBody movements ( $F_{(1,19)} = 0.198$ ,  $p = 0.662$ ,  $\eta_p^2 = 0.010$ ) and no factor reached significance in the same analysis over occipital electrodes.

### Multivariate analysis

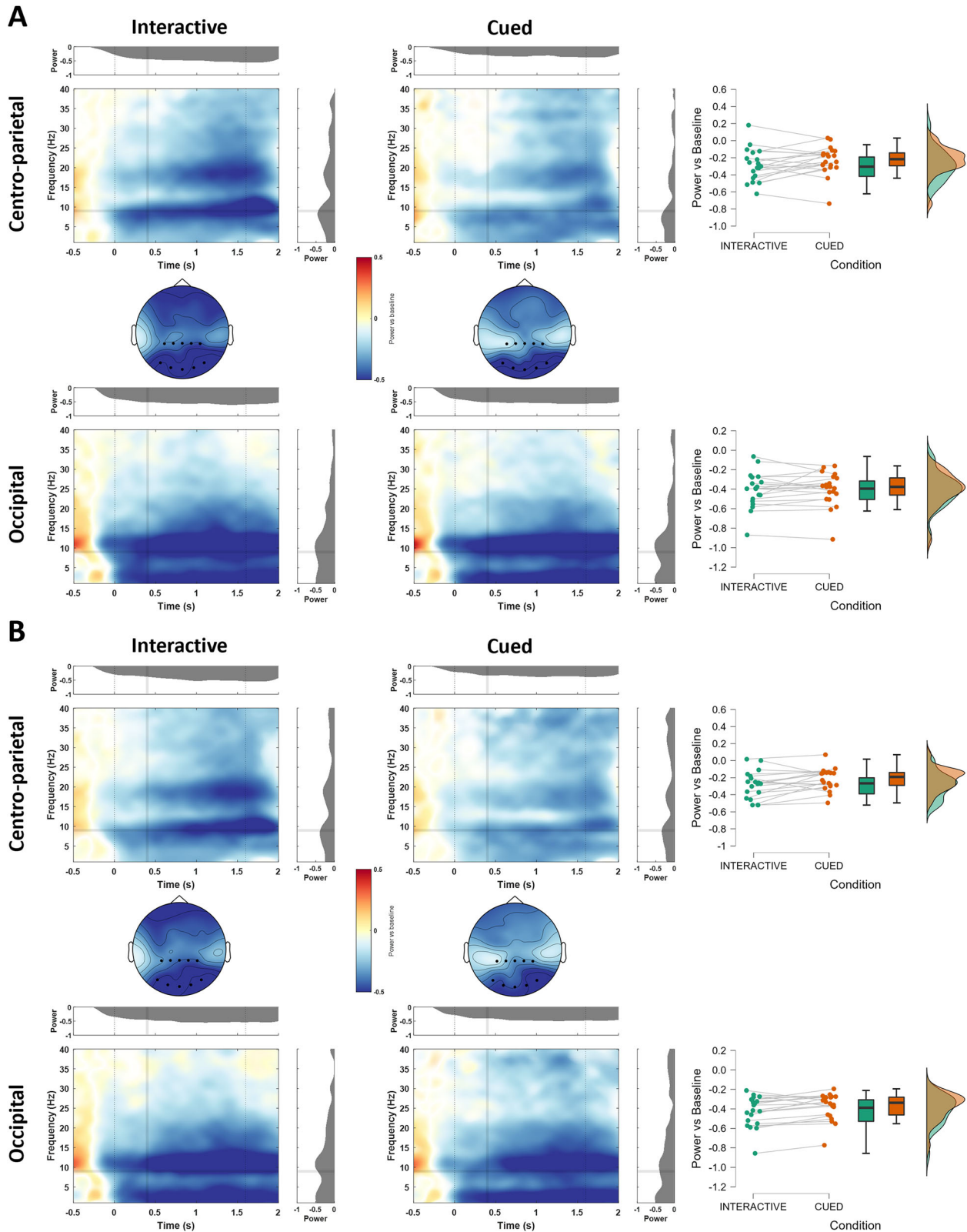
An MVPA was performed to assess if and how a classifier would be able to distinguish the EEG patterns related to the observation and prediction of a movement performed by either a Body or a NonBody. The results are shown in Figures 6 and 7. First, to investigate whether the neural patterns related to action observation were classifiable in terms of both appearance and the level of interactivity (Interactivity factor), we performed a multiclass classification across four experimental conditions (i.e., Body-Interactive, Body-Cued, NonBody-Interactive, NonBody-Cued). A multiclass LDA classifier was trained and tested on the EEG data in the time window between movement start (0 s) and movement end (1,600 ms) for each condition. Figure 6A shows the resulting confusion matrix, allowing the assessment of the probability for the classifier to in/correctly discriminate between classes of data. Interestingly, the classifier showed a performance well above chance level (i.e.,  $>0.25$ ) in distinguishing the four conditions (see the diagonal from top left to bottom right in the confusion matrix), as well as a higher tendency to confuse patterns related to same-appearance rather than same-condition factors (see lighter diagonal from bottom left to top right).

In Figure 6B, the results from the classification across time are reported, plotted in terms of percentage of accuracy, with statistically significant values of accuracy ( $p < 0.05$ , classification chance level 0.5) values of accuracy indicated by a bold line. The classifier was able to discriminate the appearance of the monitored interactor as early as  $\sim 500$  ms before movement start (significant clusters time window:  $[-600 -580$  ms;  $-560 -410$  ms]), with a first peak of classification accuracy  $\sim 250$  ms before the stimuli (either body or dots) moved (i.e., time 0). Since all the videos began with the interactor being still for  $\sim 1,000$  ms before starting to move, classification accuracy before 0 was likely driven by the visual processing of the

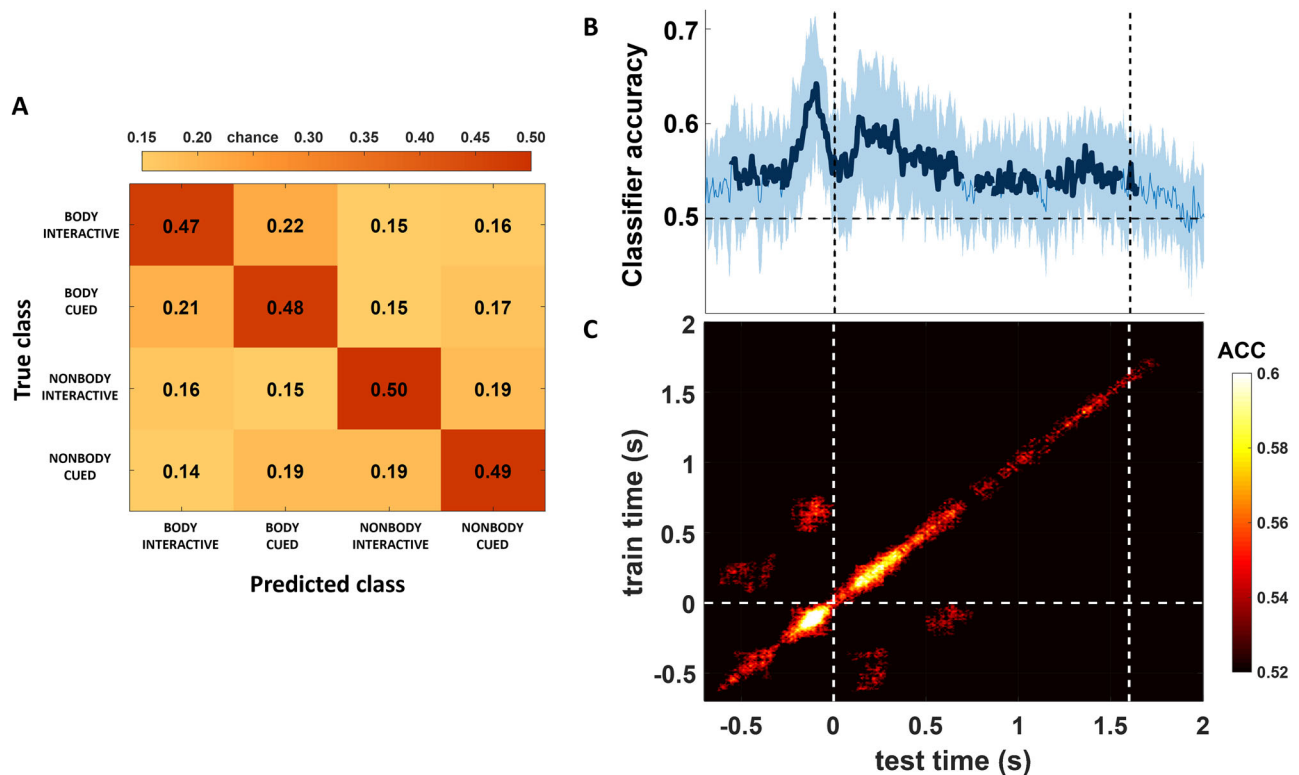


**Figure 4.** Power change relative to baseline over electrode FCz for all conditions. **A**, Time–frequency representations for all the conditions where a correction occurred. **B**, Time–frequency representations for all the conditions with no correction (time 0 represents the frame equivalent to the one when the movement changed during the correction trials). Raincloud plots on the right show the mean theta ERS over FCz in the time window between 200 and 500 ms.

interactor’s body shape. After the first peak, classification performance had a second peak ~100 ms after movement start and remained always significant until 700 ms after movement onset (significant cluster time window: [–390 700 ms]). The second peak was followed by a slow decay that remained significant for most of the movement period (significant clusters time window: [760 1,090 ms; 1,210 1,550 ms; 1,590 1,640 ms]).



**Figure 5.** Power change relative to baseline over centroparietal and occipital electrodes for all conditions (only NoCorrection trials). Panel **A** and **B** distinguish Body and NonBody trials, respectively. Topoplots in the middle and raincloud plots on the right show the mean alpha (8–13 Hz) ERD in the time window between 0 (movement start) and 1,600 ms (average movement end).



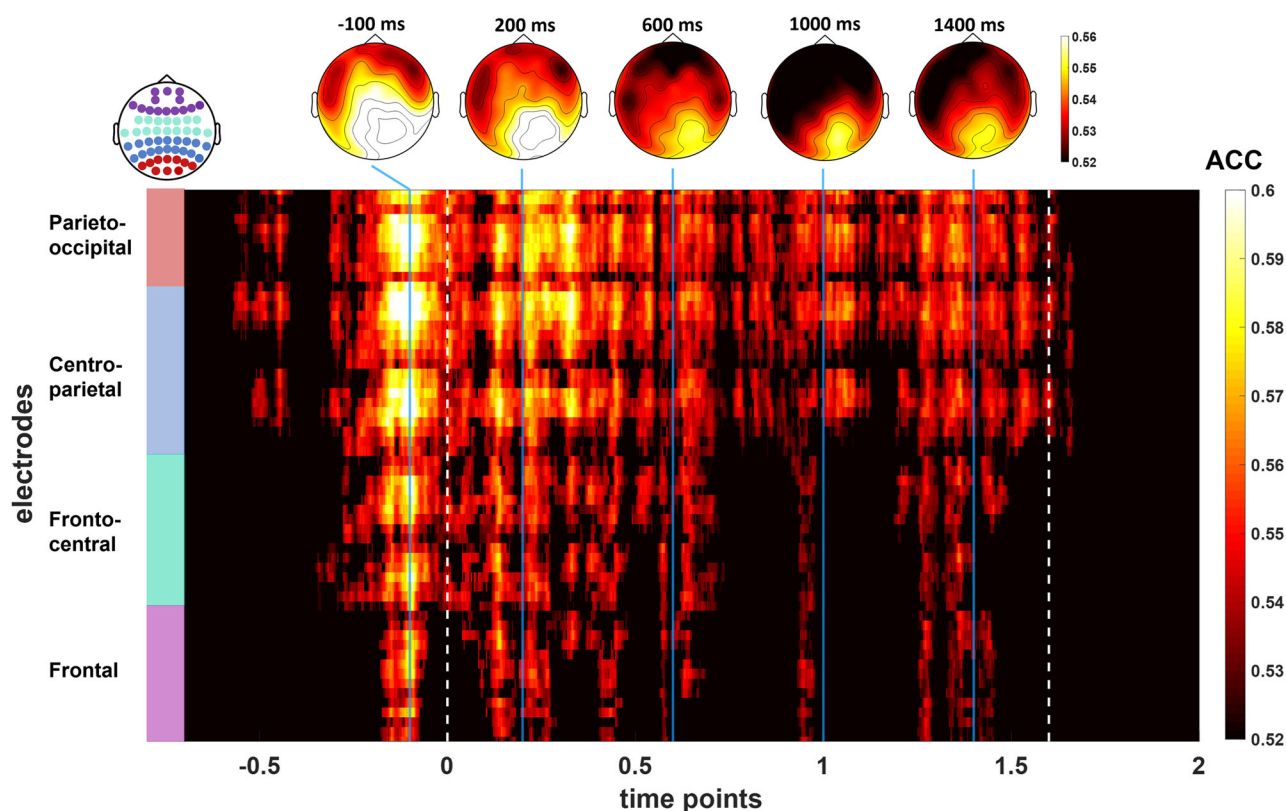
**Figure 6.** **A**, Confusion matrix showing the probability of the classifier (chance level 25%) to correctly predict the class of each sample in the movement time window (from 0 to 1,600 ms). **B**, Classification scores across time, with significant ( $p < 0.05$ ) accuracy clusters highlighted in bold. **C**, Temporal generalization matrix, with only significant accuracy scores ( $p < 0.05$ ) plotted. Dashed lines on the x-axis at 0 and 1.6 represent, respectively, movement start and average end time.

Since we were interested in the neural patterns related to the observation of the movement (from 0 ms on), in order to investigate whether they differed from the patterns related to visual shape discrimination (occurring even before movement start), we ran a temporal generalization analysis. As proposed by King and Dehaene (2014), if a classifier generalizes from one time window to another, then the underlying processes for the two time windows would be similar, and the values in the resulting time  $\times$  time plot would be represented as a single area spreading continuously out of the diagonal. If this is not the case, the results of the classification would be based on different information that travel across a chain of transient neural representations, and the results in the time  $\times$  time matrix would yield a diagonal generalization pattern. The results of this approach are reported in Figure 6C. The time  $\times$  time matrix, masked in order to show only significant results ( $p < 0.05$ ), highlights a diagonal generalization pattern, suggesting that the neural processes related to the visual processing of a body form (likely occurring as early as the video started and before 0) are not the same as the ones related to the monitoring of a movement performed by a Body or a NonBody. Interestingly, however, we found two significant blobs at symmetrical points above and below the diagonal ( $\sim 200$  and 600 ms). This suggests the recruitment of a chain of transient representations (represented along the diagonal) where later generators involved in movement monitoring reactivate earlier ones (leading to symmetrical blobs outside the diagonal) involved in the processing of the interactor's shape (King and Dehaene, 2014).

Lastly, in order to test the weight of different electrodes to classification performance, we ran a searchlight analysis across time and electrodes with the same parameters and statistical analyses adopted for the classification across time points (i.e., LDA classifier, cluster-based permutation test correcting for multiple comparisons). The results are plotted in Figure 7 both on a 2D matrix of electrodes  $\times$  time points and as topographies in five time windows of interest (i.e.,  $-100$ , 200, 600, 1,000, and 1,400 ms from movement start). Overall, the searchlight decoding is in line with the results across time, highlighting a sustained contribution of right parieto-occipital electrodes to the classification from  $\sim 300$  ms before movement start until its average end (1.6 s).

## Discussion

In the present study we recorded EEG in participants engaged in a minimally interactive version of an ecological human–avatar joint-grasping task that has been previously adopted to study interactors' action monitoring processes in different interactive contexts (Moreau et al., 2020). Specifically, we asked participants to observe a movement performed by either



**Figure 7.** Searchlight analysis results plotted as a 2D matrix showing accuracy values for each electrode and time point (from  $-700$  to  $2,000$  ms around movement start), masked in order to show only points with significantly above chance level ( $p < 0.05$ ) scores. The topographies above show accuracy values averaged in five time windows of interest. White dashed lines indicate the start of the movement ( $0$  ms) and its average end ( $1,600$  ms).

a Body or a (NonBody) control stimulus representing the same motion information (i.e., a set of dots with the very same human kinematics) and to synchronize their responses either only in time with the observed movement (i.e., Cued condition—subjects had to press a predefined key as synchronously as possible with the time when the body/dots touched to object) or in time and space (i.e., Interactive condition—subjects had to synchronize in time and to base their response depending on where the observed movement was directed, implying the necessity to continuously monitor the interactor’s movement).

Seminal studies have investigated the role of the physical appearance of an human figure in triggering activations in the AON (Perani et al., 2001; Tai et al., 2004; Gazzola et al., 2007; Shimada, 2010) and in triggering the emergence of visuo-motor interference in noninteractive conditions (Kilner et al., 2004). One previous study tested whether visuo-motor interference effects were present when human participants were coordinating their actions either with a virtual avatar or with a stimulus conveying no visual information of the form of the body but complying with biological motion rules (Gandolfo et al., 2019). Measuring the participants’ kinematic, they found no effect of the interactor’s body shape on visuo-motor interference. While this result did not highlight any specific difference for bodily versus nonbodily appearance, a study where transcranial magnetic stimulation (TMS) was used to investigate the neurophysiological indices of motor excitability during motion observation (Agosta et al., 2016) found that, although the observation of both human movement and abstract motion modulated corticospinal excitability, only the former’s kinematics were significantly correlated with the reactivity of the observer’s corticospinal tract. Moreover, previous studies highlighted the role of the AMS in detecting discrepancies between sensory predictions and action outcomes during action observation and learning (Malfait et al., 2010) and during online interpersonal motor interactions (Moreau et al., 2020, 2022; Boukarras et al., 2022), as well as the functional relation between this system and the posterior nodes of the AON, depending on the interactivity of the context (Moreau et al., 2020). Here, we aim at further understanding whether the neural patterns associated with action observation (alpha/mu ERD/ERS) and action monitoring (i.e., oPe, oERN) are modulated by the bodily appearance of an interactor. Our results replicate previous ones on the activity of the action monitoring system and shed new light on the neural patterns related to action observation. Specifically, by combining univariate and multivariate analyses, we show that (1) among the electrocortical indices of action monitoring, only the early oPe was modulated by the bodily appearance of the interactor, whereas we found no trace of the oERN; (2) the classical desynchronization in alpha/mu ( $8$ – $13$  Hz) over central sites related to action observation was stronger when subjects had to predict in space and time the observed movement (i.e., Interactive condition), compared with when they only had to synchronize their choice in time (i.e., Cued

condition); and (3) posterior nodes of the AON continuously discriminate movements performed by a Body compared with a NonBody, and the neural patterns associated to such a process are also modulated by the degree of prediction required in order to fulfill the task requirements, even in a minimally interactive scenario.

### **Interacting with a humanoid VP modulates the early error positivity (Pe)**

Previous studies have shown how the time-dependent neural responses related to the activity of the action monitoring system are elicited both when people perform errors and observe another agent making errors (Spinelli et al., 2018; Zubarev and Parkkonen, 2018; Somon et al., 2019). Nevertheless, the oERN and the oPe have been associated with different processes occurring in the error-detection cascade. Specifically, the oERN is thought to index a fast comparison between an internally generated prediction and incoming sensory inputs (Botvinick et al., 2001; Quilodran et al., 2008). Two subcomponents of the oPe have been defined, an early oPe and a late oPe, with the former occurring over fronto-central sites following the oERN and being associated predominantly to a reorientation response driven by the accumulation of evidence that an error has been committed/observed (Ridderinkhof et al., 2009; Steinhäuser and Yeung, 2010) and the latter having a centroparietal topography and being related mainly to error awareness (Endrass et al., 2007; Ullsperger et al., 2014). Interestingly, Di Gregorio et al. (2018) have studied how the Pe can occur in the absence of the ERN, thus pointing toward a functional dissociation between the systems underlying these two components.

In our study we report no trace of the oERN in any condition, whereas we report the occurrence of the early oPe only when a correction of the observed trajectory took place (Correction trials). Moreover, the amplitude of the early oPe was modulated by both the Appearance and the Interactivity factor, with larger early oPe in the Interactive condition compared with the Cued one as well as an increased amplitude for changes in the trajectory of a Body compared with a NonBody (i.e., only during Interactive trials).

With regard to the absence of the oERN in the present experiment, we hypothesize that this may be due to the minimal need for the subject to constantly anticipate and predict the observed movement, since participants observed the VP's movement from a third-person perspective in order to eventually react to it. Thus, we hypothesize that the oERN is elicited when an early mismatch between an internally generated prediction and a change in the environment occurs, whereas in our paradigm subjects formed a minimal internal prediction leading to a complete absence of oERN.

Previous studies have shown how the amplitude of the early oPe is linearly modulated as a function of the observed error's magnitude (Spinelli et al., 2018). Adding to this, our study highlights that such a top-down mechanism is modulated by the degree of prediction required during a minimally interactive interpersonal prediction task (i.e., larger amplitude for Interactive compared with Cued trials). Moreover, we report that, when a stronger attentional reorientation is needed (i.e., when subjects had to adapt their response based on the change in the movement's trajectory in Interactive trials), the appearance of the interactor plays a role, with an increase in the early oPe amplitude for corrections of bodily movements compared with nonbodily ones. This result, together with our multivariate results, suggests that bodily movements are actually encoded differently than nonbodily ones at specific stages of the action observation and monitoring process.

With regard to the error-locked activity in the time–frequency domain, although previous results have shown that the rhythms in some nodes of the AON (i.e., LOTC) are in phase with the activation of the action monitoring system during social (human–avatar) interactions (Moreau et al., 2020, 2022), we found no evidence for a modulation of the midfrontal theta rhythm by the appearance of the interacting partner (i.e., VP vs dots). Nevertheless, we replicated previous results highlighting a modulation of midfrontal theta by top-down task-related processes (Interactivity factor).

### **Centroparietal alpha/mu desynchronization is modulated by top-down processes**

The sensorimotor alpha/mu (8–13 Hz) desynchronization over central electrodes has often been the focus of EEG studies investigating action perception and understanding in interpersonal contexts (Yun et al., 2012; Fitzpatrick et al., 2019; Scanlon et al., 2022). Although our set-up was purely observational, lacking an action execution condition, our results broaden the literature on the role of alpha/mu suppression in the observation of actions in (minimally) interactive scenarios. Indeed, we report a significant difference between the alpha/mu desynchronization over central electrodes depending on the interactivity of the task. Specifically, we found a stronger decrease in power during the observation of a movement in the Interactive compared with the Cued condition. This result suggests a modulation of the activity in somatomotor areas by higher order, task-related processes and is in line with previous studies showing a modulation of sensorimotor oscillations by the observation of different type of actions (Streltsova et al., 2010), as well as studies where online interactions were investigated by means of dual-EEG (Ménoiret et al., 2014). Notably, in line with previous studies (Urgen et al., 2013), we found no modulation of the alpha/mu suppression by the Appearance factor, suggesting that such rhythm, as well as midfrontal theta, does not show any modulation for human appearances.

### **The appearance of an interactor is coded continuously throughout its movement**

Recent studies have shed light on the mechanisms underlying the perception of biological motion, showing how several cortical areas take part in this process at different hierarchical levels (Albertini et al., 2021; Puce and Perrett, 2003; Van Overwalle, 2009; Grosbras et al., 2012; Saygin et al., 2012; Hirai and Senju, 2020; Troje and Chang, 2023). For example, Duarte et al. (2022) reported that during the observation of point lights representing the main joints of a person and moving in accordance to biological motion kinematics rules, at least two different stages of motion processing must be

differentiated. Specifically, the brain must distinguish motion patterns at the level of each individual dot (i.e., local motion) and integrate them into the percept of a coherent body where these multiple joints bear together (i.e., global motion). The authors investigated how these two patterns are serially processed in specific areas of the LOTC by means of functional magnetic resonance imaging (fMRI) and multivariate analyses. Their results highlighted a two-stage framework for the neural processing of biological motion, with early (i.e., hMT+ and V3A) and higher-level [i.e., fusiform gyrus (FFG) and EBA] visual areas processing local motion patterns, whereas the posterior STS (pSTS) was involved in global motion processing. Their results provide strong, highly spatially resolved evidence that pairs with another recent study from [Chang et al. \(2021\)](#), where the authors characterized the spatiotemporal characteristics of the perception of biological motion by means of magnetoencephalography (MEG). In detail, through a combination of univariate and multivariate analyses, they showed that the neural patterns associated with global motion processing can be discriminated from those related to local kinematics as early as 100 ms after the beginning of the point-light display movement and that such discrimination remains relatively stable throughout the majority of the movement period.

Our results broaden the abovementioned ones, by investigating whether the neural activity generated by the predictive observation of bodily movements performed by a virtual partner (VP; instead of point lights) may be distinguished by the neural activity associated to the observation of the movements of a set of dots and lines showing the very same kinematics ([Hofree et al., 2015](#); [Kupferberg et al., 2018](#); [Miyamoto et al., 2023](#)). Our stimuli allowed to further understand the influence of top-down processes over these activities, as participants were asked to observe the movements while being asked to predict their spatiotemporal development at different degrees in cued and interactive conditions. We report that a LDA classifier applied on participants' EEG signal can successfully decode bodily from nonbodily movements as early as ~100 ms after movement start (in line with Chang and colleagues' results) and that such significant decoding, after a slight decay, remains relatively stable throughout the entire movement period. Moreover, by performing a temporal generalization analysis, we were able to show that the neural patterns contributing to such classification are not the same related to the processing of a still body shape, although the generators contributing to such discrimination are recalled in early time windows during bodily movement decoding. Furthermore, our searchlight classification analysis revealed the electrodes the classifier relied the most on to decode the different stimuli, highlighting a sustained contribution of mainly right occipitoparietal electrodes throughout the entire movement period, suggesting that the neural patterns from these areas are crucial for the discrimination of global bodily movements. Our results thus broaden the understanding of the role of the occipitoparietal node of the AON, following previous studies that highlighted how it is crucial for discriminating between the meaning and the effector of an action ([Lingnau and Petris, 2013](#)), recognizing different actions ([Urgesi et al., 2014](#)), while also having an early access to abstract action representations ([Tucciarelli et al., 2015](#); [Moreau et al., 2023](#)). Our results are also in line with previous fMRI ones reporting right lateral occipitotemporal areas as a crucial node for global motion processing ([Sokolov et al., 2018](#); [Duarte et al., 2022](#)).

Lastly, the results from our multiclass classification analysis strengthened the ones from the binary (i.e., Body vs NonBody) classification, showing that our classifier was able to distinguish the appearance of the observed agent while also classifying correctly trials in the Interactive condition from trials in the Cued one. Interestingly, the classifier tended to confuse more same-appearance samples rather than same-condition ones, suggesting that the appearance of the observed agent was clearly distinguishable even across different tasks.

To conclude, by means of EEG we investigated if and how biological movements are processed differently based on their bodily/nonbodily appearance in a minimally interactive scenario. Evolutionarily, biological motion perception and discrimination has been suggested to be part of a group of specialized neurocognitive mechanisms tuned broadly for "life detection" of living beings ([Troje and Westhoff, 2006](#); [Deen et al., 2023](#)), and many brain areas functionally related to such mechanisms have been found also in nonhuman primates ([Mantini et al., 2011](#); [Cléry et al., 2021](#)), highlighting a common evolutionary origin of the neural circuitry supporting complex social cognitive functions ([Roumazeilles et al., 2021](#)). Moreover, the study of biological motion is a central theme also in developmental neuroscience ([Frankenhuis et al., 2012](#)), with a growing body of literature suggesting that the perception of this specific type of movement follows the development of infants' control of their own body movements ([Blake and Shiffrar, 2006](#); [Todorova et al., 2019](#)), serving as a hallmark of social cognition ([Pavlova, 2011](#)). Targeting the neural correlates of the activity in different nodes of the AON as well as in the action monitoring system, we found that the sensorimotor alpha/mu rhythm typically related to action observation is modulated by top-down, task-related processes, with a stronger desynchronization when subjects had to rely on spatial and temporal features of the observed movement (i.e., predict and monitor the target of the observed movements) compared with when they only had to synchronize their response to the stimuli, whereas we found no modulation by the appearance of the agent moving. Rather, our multivariate results showed that such a feature is related to specific, parieto-occipital neural patterns that remain stable throughout the entire duration of an observed movement. We showed that such patterns differ from the ones related to initial shape processing (i.e., differentiating a still body shape from a non-body one), although the two overlap at specific stages during movement observation, suggesting a reactivation of the generators related to body-shape processing at later times. Lastly, we replicated previous results showing a modulation of the electrocortical indices of error monitoring by top-down, task-related processes, while only the early oPe was responsive also to the appearance of the observed agent, depending on the task. Taken together, these findings broaden our understanding of the influence of bodily appearance on the spatiotemporal processing of biological movements in the AON and in the action monitoring system during (minimally) interactive scenarios. Lastly, it must be highlighted that our results

replicate and broaden our understanding of human–avatar interactions. While further evidence on the differences between the observation and monitoring of a human compared with a nonhuman partner’s actions during real-life scenarios is still needed, our results provide important foundations to better inform artificial agent design, suggesting that the online monitoring of movements performed by a human-shaped body differs from that of a nonbodily agent. The present results may contribute to the development of virtual avatars and humanoid robots designed to evoke neural responses associated with error processing (oPe). This may be particularly valuable for applications in recreational and rehabilitative observational learning. Additionally, our findings show that the neural patterns involved in processing dynamic humanoid stimuli differ from those triggered by static body shapes. Specifically, moving humanoid figures seem to activate larger neural networks in the right occipitotemporal regions, compared with those engaged in shape processing alone. This suggests that dynamic humanoid virtual partners integrating motion and shape information may enhance the effectiveness of rehabilitation interventions compensating for lower-level deficits. Reactivity to real-person observation is different from that triggered by avatars (uncanny valley effect; Mori, 1970; Dimitrova et al., 2024) or robots (Turner-Cobb et al., 2019), but the number of contexts in which human–avatar interactions are expected in the future is increasing and ranges from education to rehabilitation (e.g., medical avatar), as well as in cultural and recreative environments (Laroche et al., 2021). Potential studies might, for example, investigate the neural patterns evoked by the monitoring of a real person compared with a robot moving in front of the participant in different interactive scenarios (e.g., active vs passive observation, joint vs separate actions). These results provide the first evidence that when individuals must predict an agent’s movements, the biological appearance of the agent’s body modulates both the activity of the AON, and the early responses of the action monitoring system. This finding also lays the foundation for future studies with larger sample sizes or involving groups of individuals with varying susceptibility to others’ body appearance, enabling a more comprehensive understanding of individual differences and enhancing the generalizability of the results. Lastly, these results may potentially inform advancements in the field of human–artificial agent interaction, fostering the development of technologies that align with the complexities of neural processing during social interactions.

## References

- Agosta S, Battelli L, Casile A (2016) Human movements and abstract motion displays activate different processes in the observer’s motor system. *Neuroimage* 130:184–193.
- Albertini D, Lanzilotto M, Maranesi M, Bonini L (2021) Largely shared neural codes for biological and nonbiological observed movements but not for executed actions in monkey premotor areas. *J Neurophysiol* 126:906–912.
- Allison T, Puce A, McCarthy G (2000) Social perception from visual cues: role of the STS region. *Trends Cogn Sci* 4:267–278.
- Blake R, Shiffrar M (2006) Perception of human motion. *Annu Rev Psychol* 58:47–73.
- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. *Psychol Rev* 108:624–652.
- Boukarras S, Özkan DG, Era V, Moreau Q, Tieri G, Candidi M (2022) Midfrontal theta transcranial alternating current stimulation facilitates motor coordination in dyadic human–avatar interactions. *J Cogn Neurosci* 34:897–915.
- Campbell JID, Thompson VA (2012) Morepower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behav Res Methods* 44:1255–1265.
- Candidi M, Curioni A, Donnarumma F, Sacheli LM, Pezzulo G (2015) Interactional leader–follower sensorimotor communication strategies during repetitive joint actions. *J R Soc Interface* 12:20150644.
- Candidi M, Sacheli LM, Era V, Canzano L, Tieri G, Aglioti SM (2017) Come together: human–avatar on-line interactions boost joint-action performance in apraxic patients. *Soc Cogn Affect Neurosci* 12:1793–1802.
- Cardellucchio P, Dolfini E, D’Ausilio A (2021) The role of dorsal premotor cortex in joint action stopping. *iScience* 24:103330.
- Chang DHF, Troje NF, Ikegaya Y, Fujita I, Ban H (2021) Spatiotemporal dynamics of responses to biological motion in the human brain. *Cortex* 136:124–139.
- Cléry JC, Hori Y, Schaeffer DJ, Menon RS, Everling S (2021) Neural network of social interaction observation in marmosets. *Elife* 10:e65012.
- Cochin S, Barthelemy C, Roux S, Martineau J (1999) Observation and execution of movement: similarities demonstrated by quantified electroencephalography: qEEG of observation and execution of movement. *Eur J Neurosci* 11:1839–1842.
- Cohen MX (2011) Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *Neuroimage* 55:1373–1383.
- Cohen MX (2014) *Analyzing neural time series data: theory and practice*. Cambridge, Massachusetts, USA: The MIT Press.
- Cohen MX (2016) Midfrontal theta tracks action monitoring over multiple interactive time scales. *Neuroimage* 141:262–272.
- Coll M-P, Press C, Hobson H, Catmur C, Bird G (2017) Crossmodal classification of mu rhythm activity during action observation and execution suggests specificity to somatosensory features of actions. *J Neurosci* 37:5936–5947.
- De Bruijn ERA, Schubotz RI, Ullsperger M (2007) An event-related potential study on the observation of erroneous everyday actions. *Cogn Affect Behav Neurosci* 7:278–285.
- Deen B, Schwiedrzik CM, Sliwa J, Freiwald WA (2023) Specialized networks for social cognition in the primate brain. *Annu Rev Neurosci* 46:381–401.
- Di Gregorio F, Maier ME, Steinhauser M (2018) Errors can elicit an error positivity in the absence of an error negativity: evidence for independent systems of human error monitoring. *Neuroimage* 172:427–436.
- Dimitrova M, Chehlarova N, Madzharov A, Krastev A, Chavdarov I (2024) Psychophysics of user acceptance of social cyber-physical systems. *Front Robot AI* 11:1.
- Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A cortical area selective for visual processing of the human body. *Science* 293:2470–2473.
- Duarte JV, Abreu R, Castelo-Branco M (2022) A two-stage framework for neural processing of biological motion. *Neuroimage* 259:119403.
- Endrass T, Reuter B, Kathmann N (2007) ERP correlates of conscious error recognition: aware and unaware errors in an antisaccade task: ERP correlates of conscious error recognition. *Eur J Neurosci* 26:1714–1720.
- Era V, Candidi M, Gandolfo M, Sacheli LM, Aglioti SM (2018) Inhibition of left anterior intraparietal sulcus shows that mutual adjustment

- marks dyadic joint-actions in humans. *Soc Cogn Affect Neurosci* 13:492–500.
- Era V, Aglioti SM, Candidi M (2020) Inhibitory theta burst stimulation highlights the role of left aIPS and right TPJ during complementary and imitative human–avatar interactions in cooperative and competitive scenarios. *Cereb Cortex* 30:1677–1687.
- Era V, et al. (2023) The dopaminergic system supports flexible and rewarding dyadic motor interactive behaviour in Parkinson's disease. *Soc Cogn Affect Neurosci* 18:nsac040.
- Falkenstein M, Hohnsbein J, Hoormann J, Blanke L (1991) Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalogr Clin Neurophysiol* 78:447–455.
- Falkenstein M, Hoormann J, Christ S, Hohnsbein J (2000) ERP components on reaction errors and their functional significance: a tutorial. *Biol Psychol* 51:87–107.
- Fitzpatrick P, Mitchell T, Schmidt RC, Kennedy D, Frazier JA (2019) Alpha band signatures of social synchrony. *Neurosci Lett* 699:24–30.
- Frankenhuis WE, House B, Barrett HC, Johnson SP (2012) Infants' perception of chasing. *Cognition* 126:224–233.
- Fusco G, Fusaro M, Aglioti SM (2022) Midfrontal-occipital  $\theta$ -tACS modulates cognitive conflicts related to bodily stimuli. *Soc Cogn Affect Neurosci* 17:91–100.
- Gandolfo M, Era V, Tieri G, Sacheli LM, Candidi M (2019) Interactor's body shape does not affect visuo-motor interference effects during motor coordination. *Acta Psychol* 196:42–50.
- Gazzola V, Rizzolatti G, Wicker B, Keysers C (2007) The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage* 35:1674–1684.
- Gehring WJ, Goss B, Coles MGH, Meyer DE, Donchin E (1993) A neural system for error detection and compensation. *Psychol Sci* 4:385–390.
- Grafton ST, Hamilton AF (2007) Evidence for a distributed hierarchy of action representation in the brain. *Hum Mov Sci* 26:590–616.
- Grootswagers T, Wardle SG, Carlson TA (2017) Decoding dynamic brain patterns from evoked responses: a tutorial on multivariate pattern analysis applied to time series neuroimaging data. *J Cogn Neurosci* 29:677–697.
- Grosbras M, Beaton S, Eickhoff SB (2012) Brain regions involved in human movement perception: a quantitative voxel-based meta-analysis. *Hum Brain Mapp* 33:431–454.
- Hardwick RM, Caspers S, Eickhoff SB, Swinnen SP (2018) Neural correlates of action: comparing meta-analyses of imagery, observation, and execution. *Neurosci Biobehav Rev* 94:31–44.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430.
- Hirai M, Senju A (2020) The two-process theory of biological motion processing. *Neurosci Biobehav Rev* 111:114–124.
- Ho J, Tumkaya T, Aryal S, Choi H, Claridge-Chang A (2019) Moving beyond P values: data analysis with estimation graphics. *Nat Methods* 16:565–566.
- Hofree G, Urgen BA, Winkielman P, Saygin AP (2015) Observation and imitation of actions performed by humans, androids, and robots: an EMG study. *Front Hum Neurosci* 9:1.
- Jung T-P, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ (2000) Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin Neurophysiol* 111:1745–1758.
- Kilner JM, Paulignan Y, Boussaoud D (2004) Functional connectivity during real vs imagined visuomotor tasks: an EEG study. *Neuroreport* 15:637–642.
- King J-R, Dehaene S (2014) Characterizing the dynamics of mental representations: the temporal generalization method. *Trends Cogn Sci* 18:203–210.
- Knoblich G, Sebanz N (2008) Evolving intentions for social interaction: from entrainment to joint action. *Philos Trans R Soc Lond B Biol Sci* 363:2021–2031.
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868.
- Kuperberg A, Iacoboni M, Flanagin V, Huber M, Kasparbauer A, Baumgartner T, Hasler G, Schmidt F, Borst C, Glasauer S (2018) Fronto-parietal coding of goal-directed actions performed by artificial agents. *Hum Brain Mapp* 39:1145–1162.
- Laroche J, Vuarnesson L, Endaltseva A, Dumit J, Bachrach A (2021) [Re]moving bodies – a shared diminished reality installation for exploring relational movement. *Front Psychol* 12:1.
- Lingnau A, Petris S (2013) Action understanding within and outside the motor system: the role of task difficulty. *Cereb Cortex* 23:1342–1350.
- Lingnau A, Downing PE (2015) The lateral occipitotemporal cortex in action. *Trends Cogn Sci* 19:268–277.
- Malfait N, Valey KF, Culham JC, Anton J-L, Brown LE, Gribble PL (2010) fMRI activation during observation of others' reach errors. *J Cogn Neurosci* 22:1493–1503.
- Mantini D, et al. (2011) Default mode of brain function in monkeys. *J Neurosci* 31:12954–12962.
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190.
- Ménoret M, Varnet L, Fargier R, Cheylus A, Curie A, des Portes V, Nazir TA, Paulignan Y (2014) Neural correlates of non-verbal social interactions: a dual-EEG study. *Neuropsychologia* 55:85–97.
- Miltner WHR, Brauer J, Hecht H, Trippe R, Coles MGH (2004) Parallel brain activity for self-generated and observed errors. Errors, conflicts, and the brain: current opinions on performance monitoring, 124–129. Scopus.
- Miyamoto Y, Uchitomi H, Miyake Y (2023) Effects of avatar shape and motion on mirror neuron system activity. *Front Hum Neurosci* 17:1173185.
- Moreau Q, Candidi M, Era V, Tieri G, Aglioti SM (2020) Midline frontal and occipito-temporal activity during error monitoring in dyadic motor interactions. *Cortex* 127:131–149.
- Moreau Q, Parrotta E, Era V, Martelli ML, Candidi M (2019) Role of the occipito-temporal theta rhythm in hand visual identification. *J Neurophysiol* 123:167–177.
- Moreau Q, Tieri G, Era V, Aglioti SM, Candidi M (2022) The performance monitoring system is attuned to others' actions during dyadic motor interactions. *Cereb Cortex* 33:222–234.
- Moreau Q, Parrotta E, Pesci UG, Era V, Candidi M (2023) Early categorization of social affordances during the visual encoding of bodily stimuli. *Neuroimage* 274:120151.
- Mori M (1970) The uncanny valley. *Energy* 7:33–35.
- Musco MA, Zazzera E, Paulesu E, Sacheli LM (2023) Error observation as a window on performance monitoring in social contexts? A systematic review. *Neurosci Biobehav Rev* 147:105077.
- Nee DE, Kastner S, Brown JW (2011) Functional heterogeneity of conflict, error, task-switching, and unexpectedness effects within medial prefrontal cortex. *Neuroimage* 54:528–540.
- Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) Fieldtrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
- Paas A, Novembre G, Lappe C, Keller PE (2021) Not all errors are alike: modulation of error-related neural responses in musical joint action. *Soc Cogn Affect Neurosci* 16:512–524.
- Pavlova MA (2011) Biological motion processing as a hallmark of social cognition. *Cereb Cortex* 22:981–995.
- Perani D, Fazio F, Borghese NA, Tettamanti M, Ferrari S, Decety J, Gilardi MC (2001) Different brain correlates for watching real and virtual hand actions. *Neuroimage* 14:749–758.
- Pezzetta R, Nicolardi V, Tidoni E, Aglioti SM (2018) Error, rather than its probability, elicits specific electrocortical signatures: a combined EEG-immersive virtual reality study of action observation. *J Neurophysiol* 120:1107–1118.
- Pezzetta R, Ozkan DG, Era V, Tieri G, Zabberoni S, Taglieri S, Costa A, Peppe A, Caltagirone C, Aglioti SM (2023) Combined EEG and

- immersive virtual reality unveil dopaminergic modulation of error monitoring in Parkinson's disease. *NPJ Parkinsons Dis* 9:3.
- Pineda JA (2005) The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing". *Brain Res Rev* 50:57–68.
- Pitcher D, Ungerleider LG (2021) Evidence for a third visual pathway specialized for social perception. *Trends Cogn Sci* 25:100–110.
- Pitcher D, Charles L, Devlin JT, Walsh V, Duchaine B (2009) Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Curr Biol* 19:319–324.
- Puce A, Perrett D (2003) Electrophysiology and brain imaging of biological motion. *Philos Trans R Soc Lond B Biol Sci* 358:435–445.
- Quilodran R, Rothé M, Procyk E (2008) Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron* 57:314–325.
- Ridderinkhof KR, Ramautar JR, Wijnen JG (2009) To P<sub>E</sub> or not to P<sub>E</sub>: a P3-like ERP component reflecting the processing of response errors. *Psychophysiology* 46:531–538.
- Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004) The role of the medial frontal cortex in cognitive control. *Science* 306:443–447.
- Roumazeilles L, et al. (2021) Social prediction modulates activity of macaque superior temporal cortex. *Sci Adv* 7:1–11.
- Ryan TA (1959) Multiple comparison in psychological research. *Psychol Bull* 56:26–47.
- Sacheli LM, Candidi M, Era V, Aglioti SM (2015) Causative role of left aIPS in coding shared goals during human–avatar complementary joint actions. *Nat Commun* 6:1–11.
- Sacheli LM, Tieri G, Aglioti SM, Candidi M (2018) Transitory inhibition of the left anterior intraparietal sulcus impairs joint actions: a continuous theta-burst stimulation study. *J Cogn Neurosci* 30:737–751.
- Sacheli LM, Musco MA, Zazzera E, Banfi G, Paulesu E (2022) How shared goals shape action monitoring. *Cereb Cortex* 32:4934–4951.
- Sauseng P, Klimesch W, Gruber WR, Hanslmayr S, Freunberger R, Doppelmayr M (2007) Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. *Neuroscience* 146:1435–1444.
- Saygin AP, Chaminade T, Ishiguro H, Driver J, Frith C (2012) The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Soc Cogn Affect Neurosci* 7:413–422.
- Scanlon JEM, Jacobsen NSJ, Maack MC, Debener S (2022) Stepping in time: alpha-mu and beta oscillations during a walking synchronization task. *Neuroimage* 253:119099.
- Shane MS, Stevens M, Harenski CL, Kiehl KA (2008) Neural correlates of the processing of another's mistakes: a possible underpinning for social and observational learning. *Neuroimage* 42:450–459.
- Shimada S (2010) Deactivation in the sensorimotor area during observation of a human agent performing robotic actions. *Brain Cogn* 72:394–399.
- Smith FW, Smith ML (2019) Decoding the dynamic representation of facial expressions of emotion in explicit and incidental tasks. *Neuroimage* 195:261–271.
- Sokolov AA, Zeidman P, Erb M, Ryvlin P, Friston KJ, Pavlova MA (2018) Structural and effective brain connectivity underlying biological motion detection. *Proc Natl Acad Sci U S A* 115:E12034–E12042.
- Somon B, Campagne A, Delorme A, Berberian B (2019) Human or not human? Performance monitoring ERPs during human agent and machine supervision. *Neuroimage* 186:266–277.
- Spinelli G, Tieri G, Pavone EF, Aglioti SM (2018) Wronger than wrong: graded mapping of the errors of an avatar in the performance monitoring system of the onlooker. *Neuroimage* 167:1–10.
- Spinelli G, Pezzetta R, Canzano L, Tidoni E, Aglioti SM (2022) Brain dynamics of action monitoring in higher-order motor control disorders: the case of apraxia. *eNeuro* 9:ENEURO.0334-20.2021.
- Steinhauser M, Yeung N (2010) Decision processes in human performance monitoring. *J Neurosci* 30:15643–15653.
- Streltsova A, Berchio C, Gallese V, Umiltà MA (2010) Time course and specificity of sensory-motor alpha modulation during the observation of hand motor acts and gestures: a high density EEG study. *Exp Brain Res* 205:363–373.
- Tai YF, Scherfler C, Brooks DJ, Sawamoto N, Castiello U (2004) The human premotor cortex is "mirror" only for biological actions. *Curr Biol* 14:117–120.
- Taylor SF, Stern ER, Gehring WJ (2007) Neural systems for error monitoring: recent findings and theoretical perspectives. *Neuroscientist* 13:160–172.
- Tidoni E, Holle H, Scandola M, Schindler I, Hill L, Cross ES (2022) Human but not robotic gaze facilitates action prediction. *iScience* 25:104462.
- Tieri G, Tidoni E, Pavone EF, Aglioti SM (2015) Mere observation of body discontinuity affects perceived ownership and vicarious agency over a virtual hand. *Exp Brain Res* 233:1247–1259.
- Todorova GK, Hatton REM, Pollick FE (2019) Biological motion perception in autism spectrum disorder: a meta-analysis. *Mol Autism* 10:49.
- Treder MS (2020) MVPA-light: a classification and regression toolbox for multi-dimensional data. *Front Neurosci* 14:289.
- Troje NF, Chang DHF (2023) Life detection from biological motion. *Curr Dir Psychol Sci* 32:26–32.
- Troje NF, Westhoff C (2006) The inversion effect in biological motion perception: evidence for a "life detector"? *Curr Biol* 16:821–824.
- Tucciarelli R, Turella L, Oosterhof NN, Weisz N, Lingnau A (2015) MEG multivariate analysis reveals early abstract action representations in the lateral occipitotemporal cortex. *J Neurosci* 35:16034–16045.
- Turner-Cobb JM, Asif M, Turner JE, Bevan C, Fraser DS (2019) Use of a non-human robot audience to induce stress reactivity in human participants. *Comput Human Behav* 99:76–85.
- Ullsperger M (2017) Neural bases of performance monitoring. In: *The Wiley handbook of cognitive control* (Egner T, ed), pp 292–313. Hoboken, New Jersey, USA: John Wiley & Sons, Ltd.
- Ullsperger M, Danielmeier C, Jocham G (2014) Neurophysiology of performance monitoring and adaptive behavior. *Physiol Rev* 94:35–79.
- Urgen BA, Plank M, Ishiguro H, Poizner H, Saygin AP (2013) EEG theta and mu oscillations during perception of human and robot actions. *Front Neurobot* 7:1.
- Urgesi C, Candidi M, Avenanti A (2014) Neuroanatomical substrates of action perception and understanding: an anatomic likelihood estimation meta-analysis of lesion-symptom mapping studies in brain injured patients. *Front Hum Neurosci* 8:1.
- Urgesi C, Candidi M, Ionta S, Aglioti SM (2006) Representation of body identity and body actions in extrastriate body area and ventral premotor cortex. *Nat Neurosci* 10:30–31.
- Van Overwalle F (2009) Social cognition and the brain: a meta-analysis. *Hum Brain Mapp* 30:829–858.
- van Schie HT, Mars RB, Coles MGH, Bekkering H (2004) Modulation of activity in medial frontal and motor cortices during error observation. *Nat Neurosci* 7:549–554.
- Vesper C, Butterfill S, Knoblich G, Sebanz N (2010) A minimal architecture for joint action. *Neural Netw* 23:998–1003.
- Wilken S, Böttcher A, Adelhöfer N, Raab M, Hoffmann S, Beste C (2023) The neurophysiology of continuous action monitoring. *iScience* 26:106939.
- Yun K, Watanabe K, Shimojo S (2012) Interpersonal body and neural synchronization as a marker of implicit social interaction. *Sci Rep* 2:959.
- Zubarev I, Parkkonen L (2018) Evidence for a general performance-monitoring system in the human brain. *Hum Brain Mapp* 39:4322–4333.
- Zucchini E, Borzelli D, Casile A (2023) Representational momentum of biological motion in full-body, point-light and single-dot displays. *Sci Rep* 13:10488.