



## Neurocomputations on dual-brain signals underlie interpersonal prediction during a natural conversation

Tengfei Zhang<sup>a,1</sup>, Siyuan Zhou<sup>b,1</sup>, Xialu Bai<sup>a,1</sup>, Faxin Zhou<sup>a</sup>, Yu Zhai<sup>a</sup>, Yuhang Long<sup>a</sup>, Chunming Lu<sup>a,\*</sup>

<sup>a</sup> State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, PR China

<sup>b</sup> Institute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu 610066, PR China



### ARTICLE INFO

#### Keywords:

Interpersonal prediction  
Natural conversation  
Prediction hierarchy  
Dual-brain neurocomputation  
fNIRS

### ABSTRACT

Prediction on the partner's speech plays a key role in a smooth conversation. However, previous studies on this issue have been majorly conducted at the single-brain rather than dual-brain level, leaving the interpersonal prediction hypothesis untested. To fill this gap, this study combined a neurocomputational modeling approach with a natural conversation paradigm in which two salespersons persuaded a customer to buy their product with their haemodynamic signals being collected using functional near-infrared spectroscopy hyperscanning. First, the results showed a cognitive hierarchy in a natural conversation, with the lower-level process (i.e., pragmatic representation of the persuasion) in the salesperson interacting with the higher-level process (i.e., value representation of the product) in the customer. Next, we found that the right dorsal lateral prefrontal cortex (rDLPFC) and temporoparietal junction (rTPJ) were associated with the representation of the product's value in the customer, while the right inferior frontal cortex (rIFC) was associated with the representation of the pragmatic processes in the salesperson. Finally, neurocomputational modeling results supported the prediction of the salesperson's lower-level brain activity based on the customer's higher-level brain activity. Moreover, the updating weight of the prediction model based on the neural computation between the rIFC of the salesperson and the rTPJ of the customer was closely associated with the interaction context, whereas that based on the rIFC-rDLPFC was not. In summary, these findings provide initial support for the interpersonal prediction hypothesis at the dual-brain level and reveal a hierarchy for the interpersonal prediction process.

### 1. Introduction

A high degree of coordination is crucial for a successful conversation. Previous studies have shown that this coordination requires not only an early plan of people's own speech but also a priori prediction of their partner's speech (Castellucci et al., 2022; Levinson, 2016). Specifically, while neural activity in the left caudal inferior frontal gyrus and middle frontal gyrus has been found to be involved in the speaker's own planning of the next words (Castellucci et al., 2022), interpersonal neural synchronization (INS) between interlocutors in the left temporoparietal junction (TPJ) has been found to be involved in the listener's predicting of the partner's next words (Dai et al., 2018). Most importantly, the early plan of an individual's own speech is suggested to largely depend on her/his prediction of the partner's inputs (Corps et al., 2018; Salazar

et al., 2021). Thus, understanding the process of how a listener predicts the upcoming inputs from the partner (i.e., interpersonal prediction) is the key to understanding both the planning process and the relationship between predicting and planning during a natural conversation. While most previous studies have behaviorally investigated the interpersonal prediction process and then demonstrated an association between the prediction and the neural response through a behavior-brain correlation at the single-brain level (Brodbeck et al., 2018; Donhauser and Baillet, 2020; Leonard et al., 2016), there is an explicit gap in directly testing the interpersonal prediction process at the dual-brain level; that is, whether and how does an individual's brain activity predict the brain activity of the other individual, based on which a successful conversation occurs? An examination of this question will provide direct evidence for the neural process of interpersonal prediction, whereby advancing our

\* Corresponding author at: State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, No. 19 Xinjiekouwai Street, Beijing 100875, PR China.

E-mail address: [luchunming@bnu.edu.cn](mailto:luchunming@bnu.edu.cn) (C. Lu).

<sup>1</sup> These authors Contributed equally to this work.

understanding on the hidden computational mechanisms of social interaction.

Previous studies on the prediction process at the single-brain level have suggested that the prediction process may occur at different linguistic levels including those of phonemics, semantics and pragmatics (Kuperberg and Jaeger, 2016; Pickering and Gambi, 2018). Specifically, it has been suggested that, during language comprehension, people continuously predict the upcoming syllables, whereby to anticipate the upcoming words (DeLong et al., 2005; Donhauser and Baillet, 2020; Heilbron et al., 2022). Additionally, people also make predictions on the grammatical forms of the upcoming words when the meaning of the context is ambiguous or with high cognitive demand (Altmann and Kamide, 1999; Wicha et al., 2004). Moreover, studies have suggested a para-linguistic prediction level in that people will infer others' communicative intentions based on the prediction of the linguistic information, which extends the prediction process beyond the mere linguistic components (Frank and Goodman, 2012; Grice, 1975). Therefore, it seems that there is a prediction hierarchy including both the linguistic and para-linguistic components. Correspondingly, previous evidence has indicated a cortical architecture supporting the prediction hierarchy. For example, studies have shown that the primary auditory and motor cortices are associated with the prediction of acoustic features produced by speech action (Donhauser and Baillet, 2020; Heilbron et al., 2022). The prediction related to the semantic and/or pragmatic processes is associated with the language network including the bilateral superior temporal cortices (STC) and left inferior frontal cortex (IFC) (Caucheteux et al., 2023; Goldstein et al., 2022; Heilbron et al., 2022; Willems et al., 2016). For the prediction of para-linguistic components, only a few studies have been conducted, which suggests the involvement of the default mode network (DMN) including the TPJ and medial prefrontal cortex (mPFC) (Mi et al., 2021; Yeshurun et al., 2021). Taken together, these findings suggest that the sensorimotor and language networks are more closely associated with the lower levels of the hierarchy such as the linguistic components, while the default mode network (e.g., TPJ, mPFC) is more closely associated with the higher levels of the hierarchy such as the para-linguistic components.

Within the prediction hierarchy, previous theories have suggested an interaction between the higher and lower levels, whereby the prediction process takes effect. Specifically, according to the intrapersonal predictive coding theory, higher-level prediction-related brain regions continuously generate top-down predictions on the upcoming inputs. Once the input and the prediction are different, the lower-level prediction-related brain regions will generate prediction errors (PE) and transmit the PE to higher-level brain regions, based on which the next prediction is updated (Friston, 2011, 2010; Mathys et al., 2011). At the interpersonal level, a few relevant theories have provided some implications on the interpersonal prediction process. For example, it has been proposed that the predictive coding mechanism may underlie social synchrony (Shamay-Tsoory et al., 2019). Most importantly, a recent theory on interpersonal verbal communication suggests that, during social interaction such as a natural conversation, individuals always seek to minimize the differences between them and their partner in different levels of the communication such as the levels of sensorimotor, semantics/syntax, and social mental states (Jiang et al., 2021). This will lead to an increase in similarity between individuals and further demonstrate high-level interpersonal synchronization of behaviors and physiological and neural activities. Moreover, the interpersonal prediction is made based on the past experience, which may result in a time lag between the prediction and the upcoming input (Fischer and Whitney, 2014; Kok et al., 2017). Correspondingly, the higher-level neural activity of the prediction hierarchy in one individual will temporally precede that of lower-level inputs from the other individual, showing a time-lagged INS. This neural pattern has been found both between human communicators such as that of listener-speaker (Dai et al., 2018) or teacher-student (Zheng et al., 2018) and non-human animals

(Ferrari-Toniolo et al., 2019). However, no neurocomputational modeling studies have been conducted on the time-lagged neural signals to test the interpersonal prediction hypothesis during a natural conversation, i.e., the prediction from the higher level of the prediction hierarchy in one individual to the lower level of the prediction hierarchy in the other individual.

This study aimed to test this hypothesis through combining a neurocomputational modeling approach with a customized persuasion task. Specifically, in a three-member group, one member was randomly assigned as the customer, whereas the other two were randomly assigned as the salespeople. In the task, two brands were randomly assigned to the two salespeople who were requested to persuade the customer in sequence to purchase their product. During the persuasion, a natural conversation occurred between the customer and one salesperson, while the other salesperson remained silent and served as a control for the shared physical and social environments. This paradigm has advantages compared to other two-member conversation paradigms in that the silent pair can serve as an ideal control for the shared physical and social environments. This is important because previous work has indicated that staying in the shared physical and social environments can induce a shared inter-brain activity (Zhang and Yartsev, 2019). Moreover, compared to other three- or more-member paradigms, a requirement for the persuasion will induce more reciprocal interactions and activate more levels of the prediction hierarchy (Jiang et al., 2021) so that the prediction between different levels of social interaction in different individuals can be effectively tested. In our task, it was assumed that the customer's decision on choosing the product was based on her/his value representation of the product, which was a para-linguistic component at the higher level of the prediction hierarchy. Additionally, the customer's value representation should be made based on the visual, acoustic, semantic or pragmatic inputs from the salesperson, which were linguistic components at the lower levels of the prediction hierarchy. Therefore, the interpersonal prediction should occur between the brain activity associated with the customer's value representation and that associated with the salesperson's linguistic inputs rather than vice versa.

To build the dual-brain interpersonal prediction model, haemoglobin concentration changes in the outer cortex of the brain were simultaneously measured in the two salespersons and the customer using functional near-infrared spectroscopy (fNIRS) hyperscanning. We mainly proposed two types of neurocomputational models. According to the interpersonal prediction hypothesis (Jiang et al., 2021), we first proposed that people should hold a priori expectations about the upcoming inputs of the partner. By dynamically calculating the difference between the expectation and the actual input (i.e., the interpersonal prediction error, IPE), the expectation would be dynamically updated and used to decide the subsequent action. Additionally, the prediction and the upcoming input would occur at different levels of the prediction hierarchy. While the prediction should be made at the higher level, the upcoming input and the IPE should be associated with the lower level. Moreover, the neural signals associated with different levels of different individuals would demonstrate a computational process consistent with the predictive coding mechanism. To validate the prediction model, we additionally proposed two control models in that people may not need much in the way of a priori expectations about the partner's upcoming inputs due to high cognitive demand from the prediction. Thus, people may only track the immediate prior inputs from the partner and calculate the difference between the prior inputs and their current state. The difference was then used to update the representation of their current state and decide on their subsequent action (Hoerl and McCormack, 2019). A significantly better performance in the prediction model compared to the control models would support the interpersonal prediction hypothesis.

## 2. Materials and methods

### 2.1. Participants

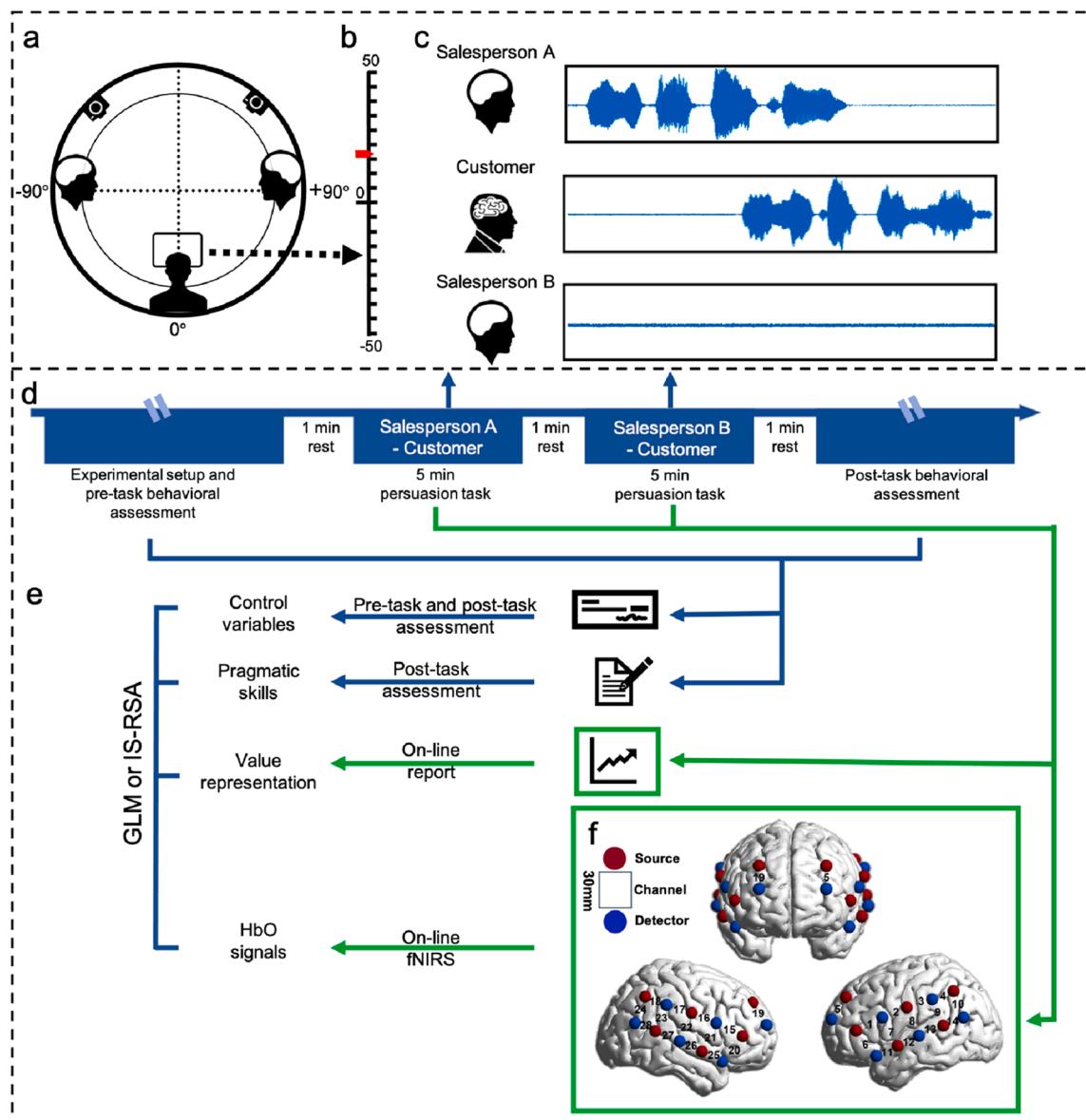
In total, 357 healthy adults (177 women, age =  $22.32 \pm 2.43$ ) were recruited in Beijing through advertisement. All participants were right-handed and had normal or corrected-to-normal vision. None of them had any language, neurological, or psychiatric disorders. These participants were randomly assigned to 119 3-member same-gender groups. No participants were excluded.

The study protocol was approved by the Institutional Review Board of the State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University. Written informed consent was received from all participants.

### 2.2. Task and materials

In the task, the salespersons were provided with two brands of a virtual product (i.e., electric toothbrush, brand A and B). To select a suitable product for the task, the following procedures were conducted. First, experimenters selected five products that are commonly used in college life, including mobile phones, sports shoes, earphones, shared bikes, and electric toothbrushes. Next, 194 college students were asked to rate their familiarity and frequency of use (1 represents the most familiar and frequent, and 5 represents the least familiar and frequent) of the five products on a 5-point Likert scale. As shown in Table S1, the electric toothbrush was the least familiar and frequently used product for college students, suggesting that many college students do not have much knowledge about the product. Finally, the electric toothbrush was selected for use in this study.

Additionally, to avoid the potential influence of the real brands on the results, two virtual brands (brand A and B) and their corresponding



**Fig. 1.** Experimental setup. (a) Three members of a group sit in a circle. (b) The customer was requested to report the value of the product on a 100-point scale. (c) An example of the persuasion process. (d) The experimental procedures. (e) Data obtained from the experiment were analyzed using either the general linear model (GLM) method or the intersubject representation similarity analysis (IS-RSA). (f) The optode probes were placed on the bilateral frontal, temporal, and parietal cortices. The numbers indicate the measurement channels. The exact Montreal Neurological Institute (MNI) coordinates are provided in Table S2.

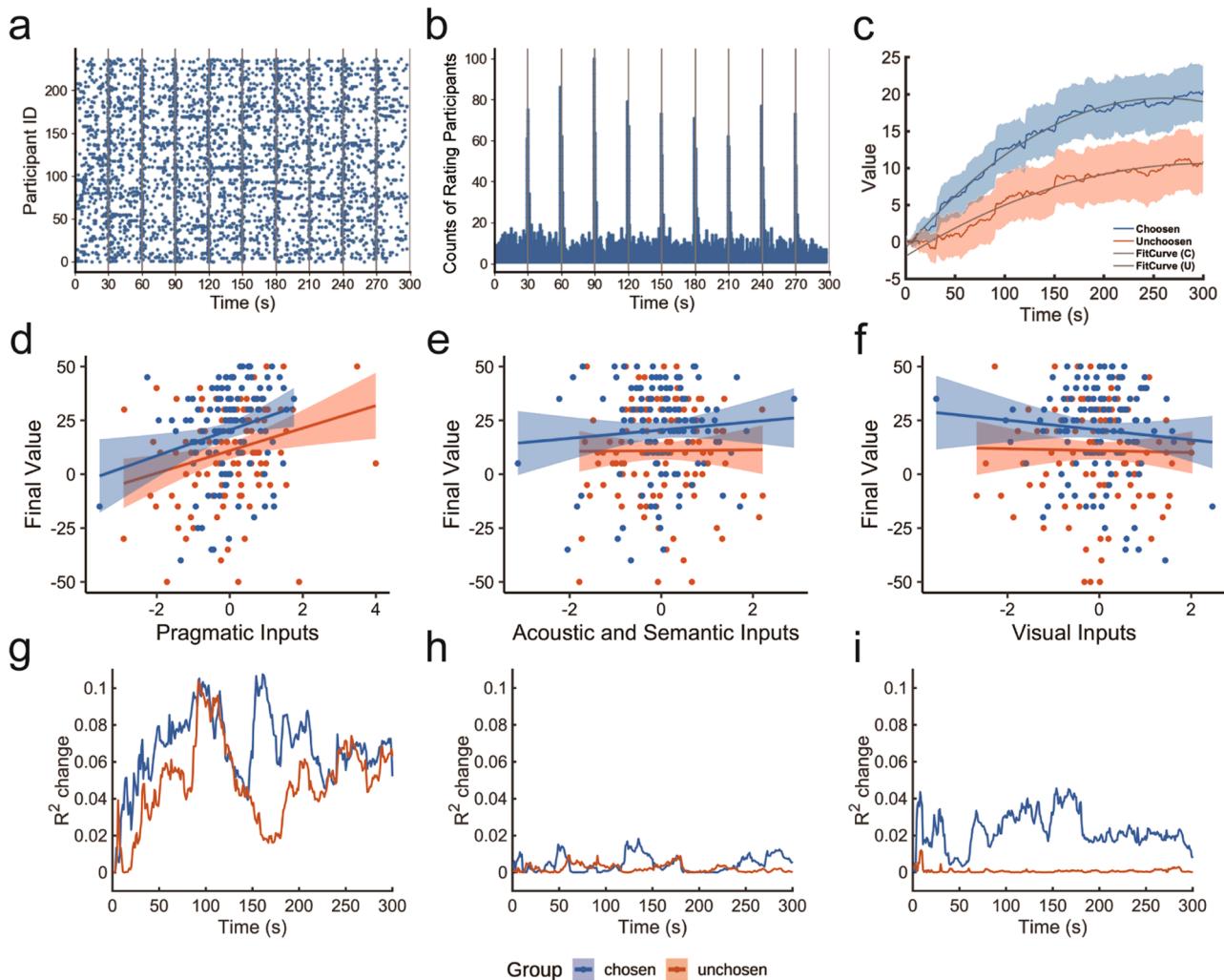
detailed product brochures were created by the experimenters. The two brands were randomly assigned to the two salespersons. To investigate whether the two brands and the products' brochures were matched, 5 participants with more than 6 months of electric toothbrush experience were recruited to conduct a qualitative assessment on the matching degree of the two brands based on the product brochures. All participants reported that the two brochures and the brands were highly matched. Furthermore, 17 participants without experience in using electric toothbrushes were also recruited. They first carefully read the two brochures and then were asked to choose one electric toothbrush they preferred to buy. The results showed that 8 participants chose brand A and 9 participants chose brand B. The chosen probability of two brands was nearly 50 %, indicating that two brochures and their corresponding brands were highly matched (brand A: 47.06 %, Table S1).

During the experiment, the three members of a group sat in a circle (Fig. 1a). The locations of the left and right salespersons were  $-90^\circ$  (left) and  $+90^\circ$  (right) relative to the customer, respectively. The two brands were randomly assigned to the two salespersons. At the beginning of the experiment, each salesperson received her/his brochure and had 15 min to read it. Meanwhile, the customer was requested to report her/his first impression of each salesperson on a 10-point Likert scale. Next, the two salespersons were requested to persuade the customer in sequence to

purchase their product. That is, while one salesperson was selling, the other salesperson was asked to wear earplugs, close their eyes, and remain motionless. During the persuasion, a natural conversation occurred between the active salesperson and the customer (Fig. 1a and c). Each persuasion process lasted for 5 min. The persuasion order and the salesperson's location were counterbalanced across the 3-member groups.

### 2.3. Assessment of the hierarchy in the conversation

First, during the persuasion, the customer was requested to rate the product's value in real-time on a 100-point scale, with  $-50$  representing the lowest value and  $50$  representing the highest value of the product (Fig. 1b). The scores were used to index the para-linguistic components of the hierarchy in the customer. The report was recorded by the customer moving a mouse wheel. Both the mouse and the customer's hand were covered with cloth, blocking the salesperson from observing the reporting, to avoid any potential confounding on the persuasion. An LCD monitor was placed in front of the customer and used to present the scale. The initial score of the scale was always set to  $0$  (Fig. 1b). While the customer was allowed to make their reports at any time, a bell buzzed every  $30$  s to remind the customer to make a report. The buzzing



**Fig. 2.** Behavioral results. (a) The distribution of the customer's on-line rating at each time point. The y-axis indicates each participant. The gray vertical line indicates the time when the bell sounds. (b) The counts of participants who made reports at the specific time points. (c) The customer's dynamic on-line report on the product's. The area around the curve represents the 95 % Confidence Interval. (d)–(f) The unique contribution of the salesperson's inputs to the customer's final on-line report on the product's value. The shadow indicates the 95 % Confidence Interval. (g)–(i) The unique contribution of the salesperson's inputs to the customer's on-line dynamic report on the product's value.

sound could be heard by both the customer and the salesperson. The rating pattern is shown in Fig. 2a and b. The mean rating times were  $21.038 \pm 14.782$ . There were  $86.18 \pm 8.55\%$  of customers rating the value around the required time points (i.e., from 1 s before to 3 s after the buzzing). The reliability of the customer's report was assessed using Cronbach's alpha coefficient, which reached 0.998. Immediately after the end of the experiment (i.e., two rounds of the persuasion), the customer was additionally requested to make a final decision on which brand of the product she/he preferred to buy in the near future.

Second, the customer was requested to assess each salesperson's linguistic inputs including the visual inputs, acoustic and semantic inputs and pragmatic inputs on a 10-point Likert scale (1 represents the lowest level, while 10 represents the highest level, Table S3). These assessments were used to index the linguistic components of the salesperson. Here the assessments of acoustic and semantic inputs were combined because it was difficult to rate them separately.

#### 2.4. fNIRS data acquisition

A LABNIRS system (Shimadzu Corporation) was used to simultaneously collect the haemodynamic concentration changes of each participant. Each participant had 28 measurement channels (CH). For each participant, three sets of customized optode probes were used. Two sets covered the bilateral inferior frontal, temporal, and parietal cortices, and the third set covered the prefrontal cortex (Fig. 1f). The anatomical location of each CH was determined by the international 10–20 system. The positions of the probe sets were checked and adjusted before the experiment to ensure consistency among participants.

To further confirm the anatomical locations of each CH, MRI data were obtained from a woman and a man. They wore plastic caps on which the CHs' true positions had been marked using Vitamin E balls. A high-resolution T1-weighted magnetization-prepared rapid gradient-echo sequence was employed (TR = 2530 ms; TE = 3.30 ms; flip angle = 7°; slice thickness = 1.3 mm; in-plane resolution =  $1.3 \times 1.0 \text{ m}^2$ ; number of interleaved sagittal slices = 128). Here a woman and a man were involved because previous evidence has suggested different brain sizes and head circumferences between women and men. Thus, to avoid bias, we determined the precise positions of the CHs in the woman and the man through scanning their brains, respectively, and combined the data in the final report. Statistical Parametric Mapping 12 (SPM12, Wellcome Department of Cognitive Neurology, London, UK) was used to normalize the MRI data to the standard Montreal Neurological Institute (MNI) coordinate space (Ashburner and Friston, 2005). The MNI coordinates of the CHs were generated according to the Automated Anatomical Labelling template (Tzourio-Mazoyer et al., 2002) using the NIRS\_SPM toolbox (Ye et al., 2009). Based on this information, we were able to check the consistency between the true positions and the expected anatomical positions and then adjust the true positions. This procedure was repeated several times until the true positions and the expected positions reached a high consistency. The MNI coordinates of the CHs are provided in Table S2.

The optical density of near-infrared light (780, 805, and 830 nm) was measured at a sampling rate of 47.6 Hz. Then, the oxyhaemoglobin (HbO), deoxyhaemoglobin (HbR), and total haemoglobin (HbT) concentration changes were calculated based on the modified Beer-Lambert law. Because previous studies showed that HbO was the most sensitive indicator of changes in regional cerebral blood flow and had the highest signal-to-noise ratio in fNIRS measurements (Hoshi, 2007), this study only focused on the changes in the HbO concentration.

#### 2.5. Behavioral data analyses

Immediately after the end of the experiment, the customers were requested to decide which brand of the product they would like to buy. Based on this decision, the two salespersons were assigned either to the chosen group or the unchosen group. Both logistic and linear regression

modeling approaches were used to determine whether it was the visual inputs, acoustic and semantic inputs, or pragmatic inputs of the salespersons that had contributed to the customers' representation of the product's value and their final decision. First, in a logistic regression model, the customer's binary choice of the product was treated as the dependent variable. Second, in a linear regression model, the customer's final representation of the product's value (i.e., the last time of on-line report before the end of the experiment) was treated as the dependent variable. Finally, a linear regression model was also conducted by incorporating each time point of the customer's on-line report on the product's value. The last two analyses were conducted within the chosen and unchosen groups separately. For all analyses, the salesperson's location, the persuasion order, the brand of the product, and the score of salesperson's first impression were treated as control variables.

#### 2.6. fNIRS data analyses

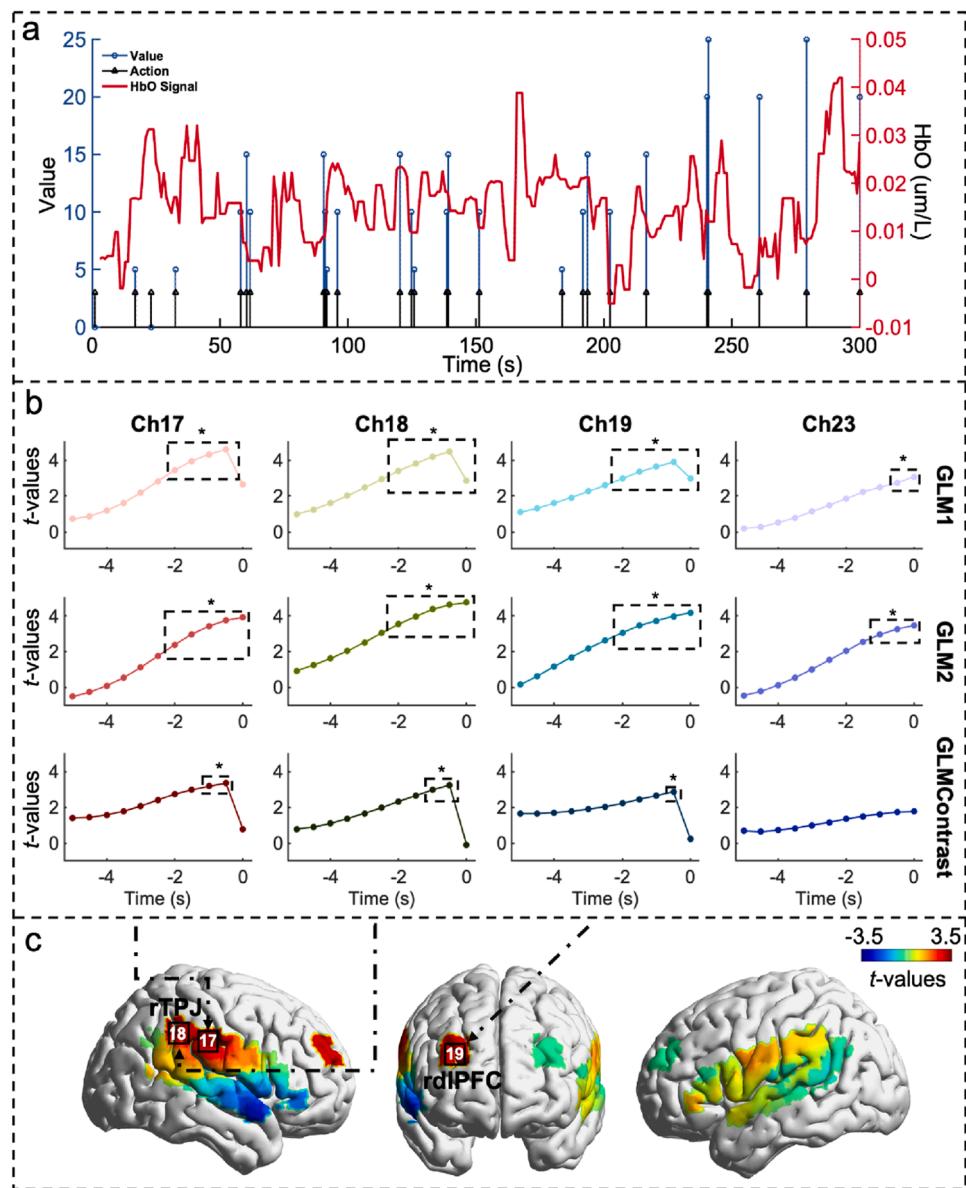
##### 2.6.1. Preprocessing

During preprocessing, functions in Homer3 (Huppert et al., 2009) were used to preprocess the data. Specifically, motion artefacts were detected and corrected using the discrete wavelet transformation filter procedure (Molavi and Dumont, 2012), and global physiological noises (e.g., skin blood flow) were removed via the principal component analysis (PCA) with an 80 % threshold of variance (Zhang et al., 2005). Finally, a bandpass filtering procedure was performed to remove the high- and low-frequency noises (0.01–0.5 Hz).

##### 2.6.2. The HbO signals associated with the customer's value representation

CH-wise general linear model (GLM) analyses were conducted to locate the brain activity that was associated with the customer's representation of the product value (Fig. 1e). In our experiment, the customers reported their rating of the value by scrolling the mouse wheel. Therefore, brain activities induced by the customer's scrolling actions and value representations might be mixed up. To address this issue, two GLMs were built to disentangle the effects of scrolling actions and value representations. In GLM 1, each score of on-line report on the product's value was considered as an event and modelled as a stick function (i.e., duration = 0) with the value score as a parametric modulator. This procedure generated a reference time course of events (Fig. 3a). Additionally, it was assumed that value representation preceded the action of the on-line report. Thus, the reference time course was shifted forward by 1–5 s relative to the time point of the on-line report action (step = 0.5 s). According to the previous literature, a rapid event-related design with a randomly jittered interval can capture the neural response to stimuli with a short interval (< 1 s) (Burrock et al., 1998). The time course was then convolved with the canonical haemodynamic response function and related to the preprocessed HbO signals of each customer. A higher *beta* value indicated a closer relationship between the reference time course and the HbO signals. Thus, GLM 1 would identify the brain activity associated with the representation of the product's value as well as the potential confoundings of the report action. In GLM 2, the reference time course of the on-line report actions was used without the scores of the rating value as the parametric modulator. All other procedures were the same as those in the GLM 1. GLM 2 would only identify the brain activity associated with the report action.

At the group level, first, to exclude the potential impacts from the times of the customer's on-line report, this variable was regressed out from CH-wise *beta* value. The residuals were used in the following analyses. Second, one sample *t*-test was conducted on the residuals of CHs across all customers to find which brain region was significantly associated with the representation of the product's value or report action. Multiple comparisons across CHs were corrected using the false discovery rate (FDR) method ( $p < 0.05$ ). Third, a paired-sample *t*-test was conducted on the residuals from different GLMs to determine which brain region was specifically associated with the representation of the product's value while that associated with report action was controlled



**Fig. 3.** Neural response in the customer. (a) An example of the GLM analysis. The blue line is the customer's on-line report of the product's value, while the black line is the customer's on-line report action. The continuous red line is the preprocessed HbO signal in the customer. (b) The brain activity of different GLMs. The parts highlighted by the dashed rectangles reached significance. (c) The brain activity associated with the representation of the product's value in the customer reached a peak when the reference time course was shifted forward by 0.5 s. \* indicates  $p < 0.05$ .

(FDR corrected,  $p < 0.05$ ).

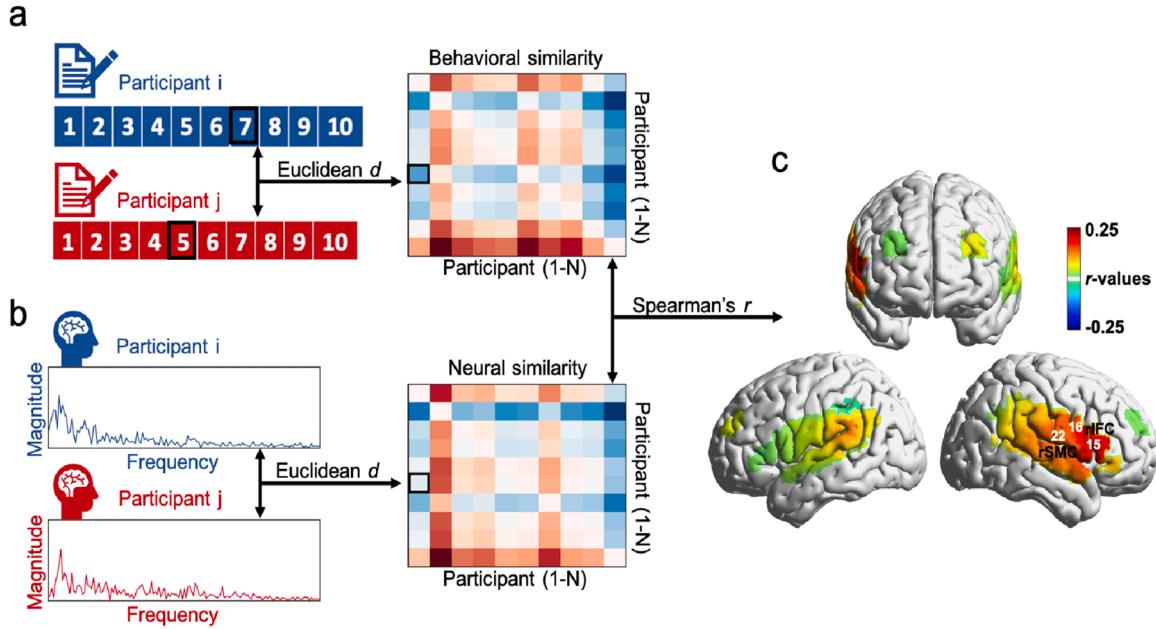
#### 2.6.3. The HbO signals associated with the salesperson's inputs

To examine the HbO signals associated with the salesperson's inputs, a CH-wise intersubject representational similarity analysis (IS-RSA, Finn et al., 2020; Zhou et al., 2023) was conducted (Fig. 4). It should be noted that in this step the GLM method was not suitable because the salespersons' inputs were measured as a stable rather than a dynamically changing variable (i.e., not a time course). To this end, first, a behavioral similarity matrix was constructed by calculating differences between each pair of salespersons in their behavioral scores. Second, the preprocessed HbO signal was fast Fourier-transformed to obtain the power spectrum. Then, the square root of the power spectrum was calculated as the signal's magnitude of the fluctuation and averaged across 0.01–0.5 Hz. The brain similarity matrix was constructed by calculating the difference (Zang et al., 2007) between each pair of salespersons in the averaged magnitude of the signal's fluctuation

(Fig. 4b). Spearman correlation was calculated between the behavioral and brain similarity matrices (Fig. 4c).

A permutation procedure was used to determine the statistical significance of the correlation following the methods described in a previous study (Kriegeskorte et al., 2008). First, the rows and columns of the behavioral similarity matrix were randomly shuffled, generating a random matrix. Second, the Spearman correlation was calculated based on the shuffled behavioral similarity matrix and the original brain matrix for each CH. Third, this procedure was repeated 1000 times to generate the null distributions of the correlation coefficients for each CH. Finally, the original Spearman correlation coefficients were compared with this distribution and corrected at the CH level (2-tailed,  $p < 0.05$ , FDR corrected).

To further validate the reliability of the IS-RSA method, the participant sample was split into two halves and the same procedures as above were repeated in each half.



**Fig. 4.** Neural response in the salesperson. (a) The behavioral matrix was calculated based on saleperson's pragmatic skills. (b) The neural matrix was calculated based on the signal's magnitude of fluctuation. It should be noted that the element in row  $i$  and column  $j$  represents the behavioral or neural difference between participant  $i$  and  $j$ . (c) HbO signals associated with the salesperson's pragmatic inputs.

## 2.7. Neurocomputational modeling

### 2.7.1. The prediction model

According to the interpersonal prediction hypothesis, the customer always holds expectations about the value of the product and makes the comparison between her/his expectation and the upcoming input from the salesperson. The difference between her/his expectation and the upcoming input, i.e., IPE, was used to guide the customer's subsequent value representation. Because we had already identified the HbO signals associated with the customer's value representation and the salesperson's input, the prediction process could be described using the z-scored HbO signals according to the equations below:

$$C_{(t+1,CH_{cus})} = C_{(t,CH_{cus})} + \alpha_{(CH_{cus})} \times \delta_{Predict(t,CH_{cus})} \quad (1)$$

$$\delta_{Predict(t, CH_{cus})} = (S_{real(t,CH_{sal})} - C_{real(t-\tau,CH_{cus})}) \quad (2)$$

where  $\alpha$  is a free parameter,  $CH_{cus}$  and  $CH_{sal}$  represent the CHs of the customer and the salesperson, respectively. Eqs. (1) and (2) describe the updating process of the customer's HbO signal based on her/his current HbO signal (i.e.,  $C_{(t,CH_{cus})}$ ) and the IPE (i.e.,  $\delta_{Predict(t, CH_{cus})}$ ). The IPE is indexed by the difference between the salesperson's current HbO signal  $S_{real(t,CH_{sal})}$  and the customer's HbO signal that preceded that of the salesperson  $C_{real(t-\tau,CH_{cus})}$  by  $\tau$  seconds ( $\tau \sim 6$  s, corresponding to the time of the HbO signal to reach a peak).  $\alpha$  ( $0 < \alpha < 1$ ) is the updating weight, which reflects the updating degree of the customer's HbO signal. The higher the  $\alpha$  value is, the greater the customer's HbO signal is updated by the upcoming input.  $t$  ( $0 \leq t \leq endtime - \tau$ ) is the time point of the signal.

### 2.7.2. The control model 1 and 2

To validate the prediction process, we proposed two control models. First, information previously transmitted by the salesperson was dynamically tracked by and integrated with the current state of the customer, and then the difference between previous information and the current input, i.e., the interpersonal tracking error (ITE), was used to guide her/his subsequent value representation. This process can be described by equations below (the control model 1):

$$C_{(t+1,CH_{cus})} = C_{(t,CH_{cus})} + \alpha_{(CH_{cus})} \times \delta_{Track1(t,CH_{cus})} \quad (3)$$

$$\delta_{Track(t,CH_{cus})} = (S_{real(t-\tau,CH_{sal})} - C_{real(t,CH_{cus})}) \quad (4)$$

where  $\delta_{Track(t,CH_{cus})}$ , i.e., the ITE, is the difference between the customer's current HbO signal (i.e.,  $C_{real(t,CH_{cus})}$ ) and the salesperson's HbO signal (i.e.,  $S_{real(t-\tau,CH_{sal})}$ ) that preceded that of the customer by  $\tau$  seconds ( $\tau \sim 6$  s).  $t$  ( $\tau \leq t \leq endtime$ ) is the time point of the signal. All other parameters are the same as those of the prediction model.

Besides, the customer may also track the current input of information from the salesperson and integrate it with her/his current state (ITE). This process can be described by the equations below (the control model 2):

$$C_{(t+1,CH_{cus})} = C_{(t,CH_{cus})} + \alpha_{(CH_{cus})} \times \delta_{Track2(t,CH_{cus})} \quad (5)$$

$$\delta_{Track2(t,CH_{cus})} = (S_{real(t,CH_{sal})} - C_{real(t,CH_{cus})}) \quad (6)$$

where  $\delta_{Track2(t,CH_{cus})}$  is the difference between the customer's current HbO signal (i.e.,  $C_{real(t,CH_{cus})}$ ) and the salesperson's HbO signal (i.e.,  $S_{real(t,CH_{sal})}$ ) at the same time point ( $\tau = 0$  s).  $t$  ( $0 \leq t \leq endtime$ ) is the time point of the signal. All other parameters are the same as those of the prediction model.

### 2.7.3. The performance of models

Each model was fitted to the preprocessed HbO signals associated with the salespersons' inputs and the customers' value representation (interaction dyads, Fig. 1c). Based on the results of the GLM and IS-RSA (see below), we obtained several brain regions (or CHs) either associated with the customer's value representation or the salesperson's inputs (see below). Before building the model, the HbO signals of CH17 and 18 were averaged because the two CHs covered the same brain regions according to the Brodmann Atlas and Automated Anatomical Labelling Atlas (Table S2). Thus, there were 1 brain region (i.e., rIFC, CH15 and CH16) involved in the salesperson's pragmatic processes and 2 brain regions (i.e., rTPJ and rdlPFC) in the customer's value representation. In addition, the preprocessed HbO signal was downsampled to one point around every six seconds to reduce the effect of autocorrelation. Thus, a total of

four pairs of HbO signals (i.e., rIFC (CH15)-rTPJ (CH17 and 18), rIFC (CH16)-rTPJ (CH17 and 18), rIFC(CH15)-rdlPFC (CH19) and rIFC (CH16)-rdlPFC (CH19)) were tested. The same tests were conducted between the customer and the silent salesperson who did not interact with the customer (no-interaction dyads, Fig. 1c). The latter was used to control the potential confounding from the shared social and physical environments.

All model fitting procedures were run in MATLAB 2020b with the "fmincon" function using maximum likelihood estimation. To ensure the validity of the fitting procedure, we randomly set 100 different initial values for the free parameters to avoid being stuck in the local minima.

In the model comparison, first, correlations between real HbO signals and the reconstructed signals across the four pairs of HbO signals were used as a criterion for the model selection. In previous studies, the correlation between real brain signal (e.g., BOLD signal) and the reconstructed signal has been widely used for selecting the best models (Caucheteux et al., 2023; Goldstein et al., 2022; Heilbron et al., 2022). Specifically, two aspects are commonly considered for model evaluations in previous studies: (1) the model's absolute performance (e.g., correlations, R-squared coefficients, and likelihood) and (2) the model's complexity (often measured by the number of free parameters in the model; that is, the more parameters there are, the more complex the model is). In this study, each model only had one free parameter, resulting in similar model complexities across models. Therefore, in this study, the absolute performance of the models (i.e., the correlation) was selected as the indicator for the model comparison. The correlation coefficients were Fisher-Z transformed. Then, the paired two-sample *t*-test was conducted to compare the performance of the three models. The *p* values were corrected by the FDR method.

Additional analysis was conducted to further validate the contribution of the IPE or the ITE in the models' performance by excluding the potential confoundings from the cross-correlation between the salesperson's and customer's HbO signals or the autocorrelation in the customer's own HbO signal. According to Eqs. (1) and (2), the cross-correlation was calculated by the Pearson's correlation between  $C_{real(t+1,CH_{cus})}$  and  $S_{real(t,CH_{sal})}$ , while the autocorrelation was calculated by the Pearson's correlation between  $C_{real(t+1,CH_{cus})}$  and  $-C_{real(t-\tau,CH_{cus})}$ . According to Eqs. (3) and (4), the cross-correlation was calculated by the Pearson's correlation between  $C_{real(t+1,CH_{cus})}$  and  $S_{real(t-\tau,CH_{sal})}$ , while the autocorrelation was calculated by the Pearson's correlation between  $C_{real(t+1,CH_{cus})}$  and  $-C_{real(t,CH_{cus})}$ . According to Eqs. (5) and (6), the cross-correlation was calculated by the Pearson's correlation between  $C_{real(t+1,CH_{cus})}$  and  $S_{real(t,CH_{sal})}$ , while the autocorrelation was calculated by the Pearson's correlation between  $C_{real(t+1,CH_{cus})}$  and  $-C_{real(t,CH_{cus})}$ . The calculation was performed for each dyad at the CH level and then combined across all dyads and CHs. Finally, the paired two-sample *t*-test was conducted to compare the performance of the IPE or the ITE model, autocorrelation and cross-correlation in pairs. FDR method was used to correct multiple comparisons.

Moreover, to examine whether the interpersonal prediction process was specific to the interaction dyads, the correlation coefficients were Fisher-Z transformed and then compared between the interaction dyads and no-interaction dyads using paired two-sample *t*-tests. A paired two-sample *t*-test was also conducted between the interaction and no-interaction dyads on the free parameter  $\alpha$  of the prediction model. Here  $\alpha$  reflects the updating weight of the customer's HbO signal.

Finally, to test whether the prediction was only made from the higher level of the prediction hierarchy in the customer rather than from the lower level of the prediction hierarchy in the salesperson, an alternative model (i.e., a prediction model on the salesperson's side) was built (see Fig. S1). In this model, we assume that the salesperson makes predictions about the value representation of customers based on her/his pragmatic processes. The same fitting and computational procedures were conducted on this model. Moreover, the performance of this model was compared with that of the prediction model on the customer side.

## 2.8. Validation on the cognitive significance of the IPE

According to the hypothesis of the prediction model, the IPE reflects the difference between the customer's expectation and the salesperson's input and was used to guide the customer's decision on the subsequent value representation. To further validate the association between the IPE and the customer's value representation, first, IPE was calculated based on the brain signals used in the prediction model. Previous work has shown that the IPE's magnitude (regardless of its sign) can boost learning and memory at both behavior and neural levels (Antony et al., 2021; Rouhani et al., 2018). Thus, the absolute value of the IPE was used in our analysis. Then, a subject-by-subject IPE similarity matrix was constructed by calculating the Euclidean distance of the absolute value of the IPE time series between each pair. Third, the value representation similarity matrix was also constructed by calculating the Euclidean distance of the time series of standardized values between each pair of participants. Finally, the Spearman correlation was applied to calculate the relationship between the IPE similarity matrix and the value representation similarity matrix. To determine the statistical significance of the Spearman correlation coefficient, a permutation test was conducted. Specifically, rows and columns of the IPE similarity matrix were randomly shuffled and then Spearman correlation was calculated again. This procedure was repeated 1000 times to generate the null distributions of the correlation coefficients. Only the original Spearman correlation coefficients that exceeded the top or bottom 2.5 % of the null distribution (2-tailed,  $p < 0.05$ ) were considered as significance.

Next, to further test whether the association between the IPE and the customer's value representation was modulated by the customer's final decision, IS-RSA was repeated within the chosen and unchosen groups, respectively.

## 2.9. Code availability

All analyses were performed using MATLAB R2020b with standard functions and toolboxes. All codes used are available upon request.

## 2.10. Data availability

The data that support the findings of this study are available from the corresponding author upon request.

## 3. Results

### 3.1. Behavioral results

The regression results showed that only the salesperson's pragmatic inputs had a significant contribution to the customer's final decision ( $\chi^2 = 8.033$ ,  $\Delta R^2 = 0.039$ ,  $p = 0.005$ , Table S4), while the other linguistic inputs, salesperson's location, the persuasion order, the brand of products, and the score of salespersons' first impression were controlled. The contribution of visual inputs and the acoustic and semantic inputs did not reach significance when pragmatic inputs, as well as the other variables, were controlled (visual inputs:  $\chi^2 = 0.345$ ,  $\Delta R^2 = 0.002$ ,  $p = 0.557$ ; acoustic and semantic inputs:  $\chi^2 = 0.550$ ,  $\Delta R^2 = 0.003$ ,  $p = 0.458$ ; Table S4).

Next, the results of the hierarchical linear regression model showed that only the pragmatic inputs of the salesperson had a significant contribution to the customer's final on-line report on the product's value (chosen group:  $F(1, 111) = 7.460$ ,  $\Delta R^2 = 0.052$ ,  $p = 0.007$ ; unchosen group:  $F(1, 111) = 9.420$ ,  $\Delta R^2 = 0.063$ ,  $p = 0.003$ , Fig. 2d and Table S4). The contribution of the other aspects did not reach significance when pragmatic inputs, as well as the other variables, were controlled (Fig. 2e-f).

Finally, the results showed that the pragmatic inputs had a significant contribution to the on-line dynamic report on the product's value

along the course of conversation ( $ps < 0.05$ ) after correction for multiple comparisons across time points using the FDR method (Fig. 2g). Moreover, the contribution of pragmatic inputs rapidly increased as the conversation progressed and then stayed on a plateau after 11 s for the chosen group and 32 s for the unchosen group, when the other variables were controlled (chosen group at 11 s:  $F(1, 111) = 4.522, \Delta R^2 = 0.037, p = 0.036$ ; unchosen group at 32 s:  $F(1, 111) = 5.998, \Delta R^2 = 0.044, p = 0.016$ ). However, the contribution of the other aspects didn't reach significance at any time points when pragmatic inputs, as well as the other variables, were controlled ( $ps > 0.05$ , FDR correction across time points, Fig. 2h-i). Together, these results indicated that the pragmatic inputs of the salesperson were closely associated with the value representation of the product in the customer. Therefore, the HbO signals were identified based on either the pragmatic inputs on the salesperson's side or the value representation on the customer's side.

### 3.2. The HbO signals associated with the customer's value representation

In GLM 1, the activities in the right TPJ (rTPJ, CH17 and 18) and dorsal lateral PFC (rdlPFC, CH19) reached the peak at time  $-0.5$ , while that of the rTPJ (CH23) peaked at time 0 (Fig. 3b). Results from one-sample  $t$ -tests further indicated that the brain activities of these two brain regions were significantly higher than 0 (rTPJ, CH17,  $t(235) = 4.587, p < 0.001$ ; CH18,  $t(235) = 4.472, p < 0.001$ ; CH23,  $t(235) = 3.083, p = 0.036$ ; rdlPFC, CH19,  $t(235) = 3.896, p = 0.001$ , FDR corrected).

In GLM 2, the rTPJ (CH17, 18 and 23) and rdlPFC (CH19) activities peaked at time 0 (Fig. 3b), and the results of one-sample  $t$ -tests further revealed that their brain activities were significantly higher than 0 (rTPJ, CH17,  $t(235) = 3.916, p = 0.001$ ; CH18,  $t(235) = 4.741, p < 0.001$ ; CH23,  $t(235) = 3.461, p = 0.004$ ; rdlPFC, CH19,  $t(235) = 4.167, p < 0.001$ , FDR corrected).

The direct comparison between GLM 1 and 2 showed that, comparing to the report actions, value representations were associated with the rTPJ (CH17,  $t(235) = 3.338, p = 0.019$ ; CH18,  $t(235) = 3.241, p = 0.019$ , FDR corrected) and rdlPFC (CH19,  $t(235) = 2.862, p = 0.043$ , FDR corrected) in the customer when the reference time course was shifted forward by 0.5 s (Fig. 3c). The brain activities of the other brain regions did not reach significance ( $ps > 0.05$ , Fig. 3b, Table S5). This result indicated that the HbO signals in the rTPJ and rdlPFC were more closely associated with the representation of the product's value in the customer.

### 3.3. The HbO signals associated with the salesperson's pragmatic inputs

The IS-RSA results showed significant correlations between the salesperson's pragmatic inputs and HbO signals in the right IFC (rIFC, CH15,  $r = 0.204, p = 0.028$ ; CH16,  $r = 0.148, p = 0.037$ , FDR corrected) and sensorimotor cortex (rSMC, CH22,  $r = 0.136, p = 0.028$ , FDR corrected) (Fig. 4c). The results in other channels did not reach significance ( $ps > 0.05$ , Table S6).

The replication results showed that the correlations between the salesperson's pragmatic inputs and HbO signals reached significance only in the rIFC (first half, CH15,  $r = 0.202, p = 0.006$ ; CH16,  $r = 0.147, p = 0.038$ ; second half, CH15,  $r = 0.203, p = 0.010$ ; CH16,  $r = 0.149, p = 0.026$ , no correction). The correlations in the other channels did not reach significance ( $ps > 0.05$ ). Thus, HbO signals of the rIFC as well as those of the rTPJ and dIPFC were entered into the neurocomputational modeling analysis below.

### 3.4. The neurocomputational model

First, the correlation results showed that, compared to the two control models, the prediction model had a better performance among the interaction dyads (prediction model vs. control model 1:  $t(942) =$

$89.891, p < 0.001$ ; prediction model vs. control model 2:  $t(942) = 91.696, p < 0.001$ ; control model 1 vs. control model 2:  $t(942) = 1.954, p = 0.051$ , FDR corrected). This result confirmed the neurocomputations on the interpersonal prediction process in the dual-brain signals during a natural conversation.

Second, the validation results indicated that the performance of the IPE model was significantly better than the cross-correlation ( $t(942) = 28.020, p < 0.001$ ) and autocorrelation approaches ( $t(942) = 28.947, p < 0.001$ ), but no significant difference was found between the cross-correlation and autocorrelation approaches ( $t(942) = 0.148, p = 0.883$ , FDR corrected). The ITE model's performance was significantly lower than that of the cross-correlation approach (control model 1:  $t(942) = -36.699, p < 0.001$ ; control model 2:  $t(942) = -26.974, p < 0.001$ ) but higher than that of the autocorrelation approach (control model 1:  $t(942) = 16.137, p < 0.001$ ; control model 2:  $t(942) = 15.778, p < 0.001$ , Fig. 5a).

Third, the comparison between the interaction and no-interaction dyads using paired two-sample  $t$ -tests showed that the correlation was significantly higher in the interaction dyads than in the non-interaction dyads in the rIFC<sub>salesperson</sub>-rTPJ<sub>customer</sub> (CH16-17 and CH16-18,  $t(237) = 3.120, p = 0.012$ ) and rIFC<sub>salesperson</sub>-rdlPFC<sub>customer</sub> (CH16-CH19,  $t(237) = 3.068, p = 0.014$ , Fig. 5c and d) after FDR correction ( $p < 0.05$ ). Additionally, the updating weight of interaction dyads was significantly higher than that of no-interaction dyads in the rIFC<sub>salesperson</sub>-rTPJ<sub>customer</sub> ( $t(237) = 2.567, p = 0.022$ , FDR corrected). No significant difference was found in the rIFC<sub>salesperson</sub>-rdlPFC<sub>customer</sub> ( $t(237) = 1.5754, p = 0.117$ , FDR corrected, Fig. 5b).

Finally, to test whether the prediction was only made from the higher level of the prediction hierarchy in the customer, an alternative model from the salesperson's side was built. However, the results showed that, although the prediction model from the salesperson's side also outperformed the control models as did the prediction model from the customer's side (Fig. S1a), the prediction model from the customer's side performed significantly better than that from the salesperson's side ( $t(237) = 2.087, p = 0.038$ , Fig. S1b).

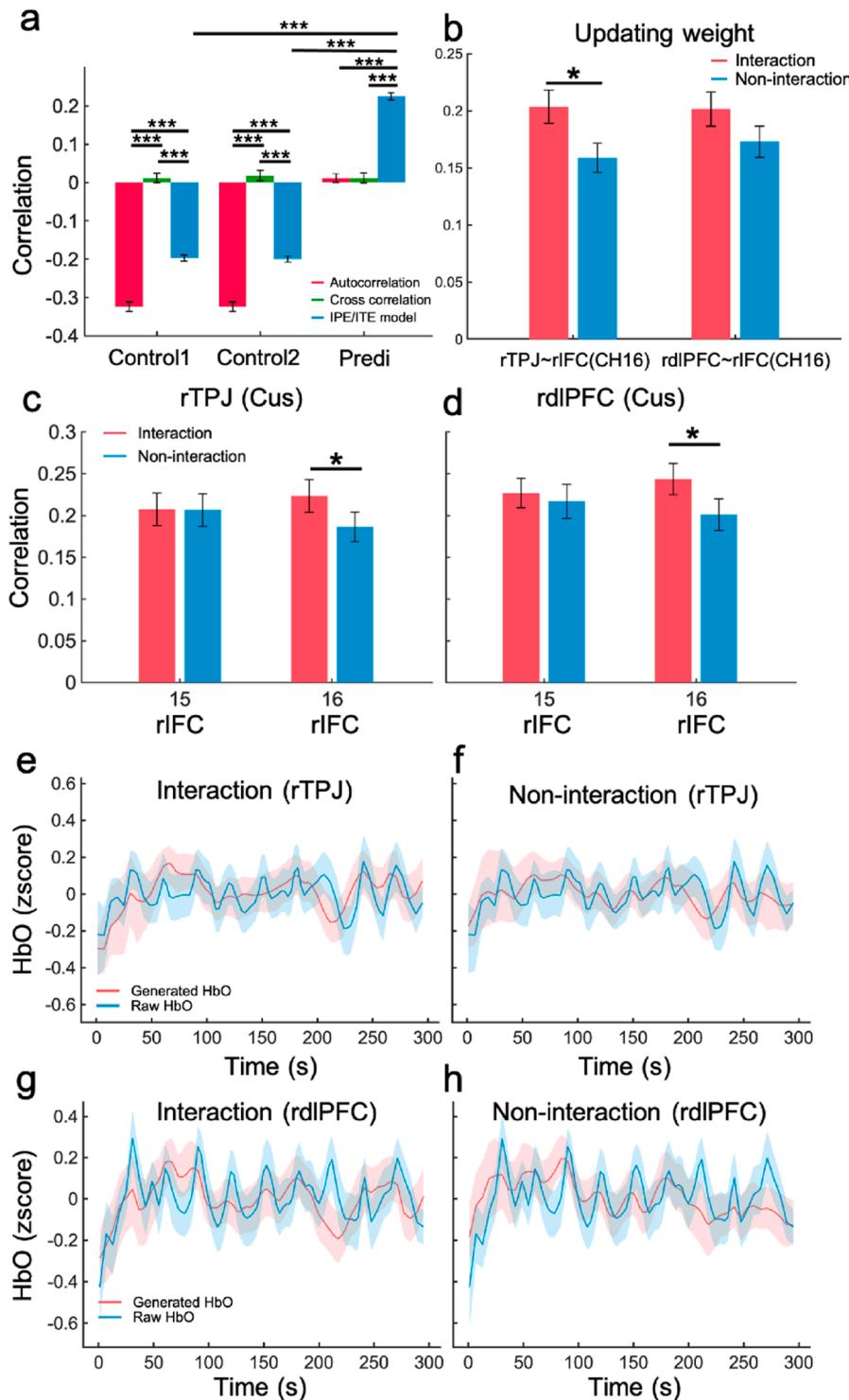
### 3.5. Validation on the cognitive significance of the IPE

As shown in Fig. 6, the Spearman correlation between the IPE and the value representation reached significance ( $r = 0.0211, p < 0.001$ ). Additionally, the results showed that in the chosen group, the IPE was significantly correlated with the value representation ( $r = 0.0246, p = 0.018$ ), but the correlation did not reach significance in the unchosen group ( $r = 0.0155, p = 0.087$ ). Thus, the IPE-based prediction process seemed to play a key role in the customer's final decision.

## 4. Discussion

Interpersonal prediction is a ubiquitous feature of human conversations. Most previous studies, however, have only tested this process at the behavioral level and from the perspective of the listener, leaving an explicit gap in directly understanding the interpersonal prediction process at the neural level and from the perspective of interlocutors. To fill this gap, the present study employed a customized persuasion task, based on which the dual-brain HbO signals associated with the different levels of a conversation were identified. Then, we tested the prediction model as compared to the control models on these HbO signals. The results supported the prediction model, suggesting that the prediction occurs not only at the behavioral level within an individual (i.e., intrapersonal prediction) but also at the neural level between individuals (i.e., interpersonal prediction).

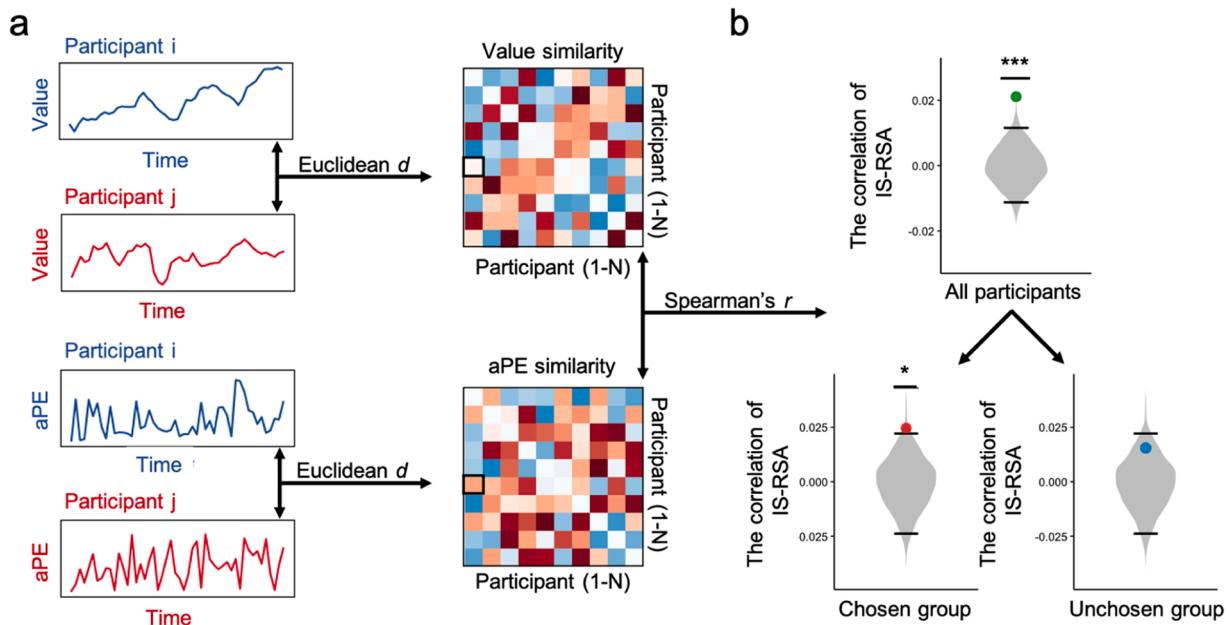
Specifically, first, we found that the prediction occurred between the rTPJ of the customer and the rIFC of the salesperson. According to the intrapersonal predictive coding theory, higher-level brain regions of the prediction hierarchy continuously generate top-down predictions on the upcoming inputs and lower-level brain regions of the prediction



**Fig. 5.** Performance of the prediction model on the customer's side. (a) The prediction model performed better than the two control models. (b) Parameter comparison between the interaction and no-interaction dyads. (c) and (d) Correlation between the model-generated and real HbO signals. (e–h) Examples of the pre-processed HbO signals and HbO signals that were generated by the prediction model. \* indicates  $p < 0.05$ , \*\*\* indicates  $p < 0.001$ . The black error bars represent standard errors. Cus, customer.

hierarchy transmit prediction errors to higher-level brain regions by comparing the difference between the prediction and inputs. Previous studies employing the behavior-brain correlation approach have illustrated a prediction hierarchy and a cortical architecture for the social prediction including both linguistic and para-linguistic levels in a non-interaction context. In those studies, the left frontal and temporal cortices are more involved in the prediction at the linguistic level,

showing an activation prior to the occurrence of the target phoneme or word during the speech comprehension (Brodbeck et al., 2018; Donhauser and Baillet, 2020; Goldstein et al., 2022; Heilbron et al., 2022; Leonard et al., 2016; Willems et al., 2016), while the DMN including the TPJ is more involved in the prediction at the para-linguistic level such as the social intention (Mi et al., 2021; Yeshurun et al., 2021). The present findings extend these findings to an interaction context, showing an



**Fig. 6.** The relationship between the IPE and the customer's value representation. (a) The value change similarity matrix was calculated based on the time series of on-reported value in the customer. The IPE similarity matrix was calculated based on the brain signals of the rIFG-rTPJ. (b) The IS-RSA results for all participants, chosen group and unchosen groups, respectively. \* indicates  $p < 0.05$ , \*\*\* indicates  $p < 0.001$ .

interaction between the rTPJ at the para-linguistic level and the frontal and temporal cortices at the linguistic level. This result is consistent with recent evidence showing that INS in the rTPJ usually peaked prior to the occurrence of turn-taking and is closely associated with the selective conversation in a noisy context (Dai et al., 2018). According to previous studies, the rTPJ is an important region for predicting the other's intention by integrating information from low-level brain regions (Schurz et al., 2014; Van Overwalle, 2009). Thus, the present findings suggest that the customer might dynamically evaluate the product's value by predicting the underlying meaning and intention of the salespersons' input and then use the IPEs to update her/his subsequent evaluation.

Second, we showed that it was the salesperson's pragmatic inputs rather than the basic visual inputs or acoustic and semantic inputs that contributed to the customer's value representation. Previous linguistic theories have suggested that effective communication is based on the ability of pragmatic interpretation, which means that the speaker and listener need to encode and decode each other's intended meaning rather than only perceiving the visual or acoustic inputs (Grice, 1975). The IFC is thought to be the key hub for both the execution and observation of pragmatic processes. Many studies have found stronger activation in the IFC during both the comprehension and production of the intended meaning through the pragmatic process (Bašnáková et al., 2014; Beatty et al., 2017; Egorova et al., 2016). Additionally, studies have found that compared to the left IFC, the rIFC plays a more unique role in forming pragmatic representations (Ferstl et al., 2005; Menenti et al., 2009). Most importantly, previous predictive coding theory on the single brain as well as the interpersonal prediction theory indicated that the interpersonal prediction should occur between nearby levels of the hierarchy. Thus, our findings support the conclusion that interpersonal prediction also occurs in a hierarchical structure.

Third, when comparing the updating weight of the prediction model between the interaction and no-interaction dyads, only the result of the rIFC-rTPJ reached significance. This finding suggests that the rIFC-dlPFC is not directly associated with the dynamic process of the customer's linguistic prediction. Although both the rTPJ and dlPFC are important for social decision-making (Rilling and Sanfey, 2011; Seo and Lee, 2012), previous studies have shown that, while the rTPJ is more

closely associated with the dynamic prediction of the other's mental states based on the current context (Konovalov et al., 2021), the dlPFC is more closely associated with the cognitive control and long-term context (Bhatt et al., 2010; Busemeyer et al., 2019; Yamagishi et al., 2016). Therefore, compared to the dlPFC, the rTPJ-related prediction might be more closely associated with the context of a conversation.

Finally, we found that the dynamic change of the IPE calculated based on the brain signals of the rIFC-rTPJ is closely associated with the customer's dynamic value representation, which provides further evidence for the prediction model; that is, the customers dynamically update their subsequent decision based on the IPE. Previous single-brain studies have found that PE plays an important role in driving behavior change in decision-making (Hackel et al., 2015; Hayden et al., 2011; Rouhani and Niv, 2021; Vassena et al., 2020). A recent study also found that in a social interaction context, PE of the observed action of others drives the change of value estimation (Jiang et al., 2023), which is encoded in the lateral frontal cortex. However, these studies have examined the dynamic relationship among PE, decision behaviors and brain functions only at the single-brain level. Our findings extend these conclusions to the dual-brain level in the interaction context.

In summary, while the interpersonal prediction process is a key process in a natural conversation and has been considered a crucial feature of the human brain, the neurocomputational mechanisms of the interpersonal prediction hypothesis have never been directly tested at the dual-brain level. The present study filled this gap by directly testing the interpersonal prediction hypothesis between neural signals of two individuals during a natural conversation. The results confirmed, for the first time, that the interpersonal prediction process between the HbO signals of the rTPJ and dlPFC on the prediction side and the HbO signal of the rIFC on the target side. Moreover, compared to the dlPFC, the prediction of the rTPJ seems to be more closely associated with the conversation context. These findings extend the social prediction from the single-brain level to the dual-brain level, establishing a robust relationship between the interpersonal prediction process and natural conversations.

## CRediT authorship contribution statement

**Tengfei Zhang:** Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Siyuan Zhou:** Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Xialu Bai:** Investigation, Formal analysis. **Faxin Zhou:** Data curation. **Yu Zhai:** Data curation. **Yuhang Long:** Conceptualization, Investigation. **Chunming Lu:** Conceptualization, Writing – original draft, Writing – review & editing, Funding acquisition, Supervision.

## Declaration of Competing Interest

The authors declare no competing financial interests.

## Data availability

Data will be made available on request.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (62293550, 62293551, 61977008) and the Lingang Laboratory (LG-TKN-202205-01).

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2023.120400](https://doi.org/10.1016/j.neuroimage.2023.120400).

## References

- Altmann, G.T.M., Kamide, Y., 1999. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73, 247–264.
- Antony, J.W., Hartshorne, T.H., Pomeroy, K., Gureckis, T.M., Hasson, U., McDougle, S.D., Norman, K.A., 2021. Behavioral, physiological, and neural signatures of surprise during naturalistic sports viewing. *Neuron* 109, 377–390. <https://doi.org/10.1016/j.neuron.2020.10.029> e9.
- Ashburner, J., Friston, K.J., 2005. Unified segmentation. *Neuroimage* 26, 839–851. <https://doi.org/10.1016/j.neuroimage.2005.02.018>.
- Bašnáková, J., Weber, K., Petersson, K.M., van Berkum, J., Hagoort, P., 2014. Beyond the language given: the neural correlates of inferring speaker meaning. *Cereb. Cortex* 24, 2572–2578. <https://doi.org/10.1093/cercor/bht112>.
- Beatty, R.E., Silvia, P.J., Benedek, M., 2017. Brain networks underlying novel metaphor production. *Brain Cogn.* 111, 163–170. <https://doi.org/10.1016/j.bandc.2016.12.004>.
- Bhatt, M.A., Lohrenz, T., Camerer, C.F., Montague, P.R., 2010. Neural signatures of strategic types in a two-person bargaining game. *Proc. Natl. Acad. Sci. U. S. A.* 107, 19720–19725. <https://doi.org/10.1073/pnas.1009625107>.
- Brodbeck, C., Hong, L.E., Simon, J.Z., 2018. Rapid transformation from auditory to linguistic representations of continuous speech. *Curr. Biol.* 28, 3976–3983. <https://doi.org/10.1016/j.cub.2018.10.042> e5.
- Burock, M.A., Buckner, R.L., Woldorff, M.G., Rosen, B.R., Dale, A.M., 1998. Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *Neuroreport* 9, 3735–3739. <https://doi.org/10.1097/00001756-19981116-00030>.
- Busemeyer, J.R., Gluth, S., Rieskamp, J., Turner, B.M., 2019. Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends Cogn. Sci.* 23, 251–263. <https://doi.org/10.1016/j.tics.2018.12.003>.
- Castellucci, G.A., Kovach, C.K., Howard, M.A., Greenlee, J.D.W., Long, M.A., 2022. A speech planning network for interactive language use. *Nature* 602, 117–122. <https://doi.org/10.1038/s41586-021-04270-z>.
- Caucheteux, C., Gramfort, A., King, J.R., 2023. Evidence of a predictive coding hierarchy in the human brain listening to speech. *Nat. Hum. Behav.* 21–24. <https://doi.org/10.1038/s41562-022-01516-2>.
- Corps, R.E., Crossley, A., Gambi, C., Pickering, M.J., 2018. Early preparation during turn-taking: listeners use content predictions to determine what to say but not when to say it. *Cognition* 175, 77–95. <https://doi.org/10.1016/j.cognition.2018.01.015>.
- Dai, B., Chen, C., Long, Y., Zheng, L., Zhao, H., Bai, X., Liu, W., Zhang, Y., Liu, L., Guo, T., Ding, G., Lu, C., 2018. Neural mechanisms for selectively tuning in to the target speaker in a naturalistic noisy situation. *Nat. Commun.* 9 <https://doi.org/10.1038/s41467-018-04819-z>.
- DeLong, K.A., Urbach, T.P., Kutas, M., 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nat. Neurosci.* 8, 1117–1121. <https://doi.org/10.1038/nn1504>.
- Donhauser, P.W., Baillet, S., 2020. Two distinct neural timescales for predictive speech processing. *Neuron* 105, 385–393. <https://doi.org/10.1016/j.neuron.2019.10.019> e9.
- Egorova, N., Shtyrov, Y., Pulvermüller, F., 2016. Brain basis of communicative actions in language. *Neuroimage* 125, 857–867. <https://doi.org/10.1016/j.neuroimage.2015.10.055>.
- Ferrari-Toniolo, S., Visco-Comandini, F., Battaglia-Mayer, A., 2019. Two brains in action: joint-action coding in the primate frontal cortex. *J. Neurosci.* 39, 3514–3528. <https://doi.org/10.1523/JNEUROSCI.1512-18.2019>.
- Ferstl, E.C., Rinck, M., Cramon, D.Y.V., 2005. Emotional and temporal aspects of situation model processing during text comprehension: an event-related fMRI study. *J. Cogn. Neurosci.* 17, 724–739.
- Finn, E.S., Glerean, E., Khajandi, A.Y., Nielson, D., Molfese, P.J., Handwerker, D.A., Bandettini, P.A., 2020. Idiosynchrony: from shared responses to individual differences during naturalistic neuroimaging. *Neuroimage* 215, 116828. <https://doi.org/10.1016/j.neuroimage.2020.116828>.
- Fischer, J., Whitney, D., 2014. Serial dependence in visual perception. *Nat. Publ. Gr.* 17 <https://doi.org/10.1038/nrn.3689>.
- Frank, M.C., Goodman, N.D., 2012. Predicting pragmatic reasoning in language games. *Science* 336. <https://doi.org/10.1126/science.1218633> (80–)998–998.
- Friston, K., 2011. What is optimal about motor control? *Neuron* 72, 488–498. <https://doi.org/10.1016/j.neuron.2011.10.018>.
- Friston, K., 2010. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. <https://doi.org/10.1038/nrn2787>.
- Goldstein, A., Zada, Z., Buchnik, E., Schain, M., Price, A., Aubrey, B., Nastase, S.A., Feder, A., Emanuel, D., Cohen, A., Jansen, A., Gazula, H., Choe, G., Rao, A., Kim, C., Casto, C., Fanda, L., Doyle, W., Friedman, D., Dugan, P., Melloni, L., Reichtart, R., Devore, S., Flinner, A., Hasenfratz, L., Levy, O., Hassidim, A., Brenner, M., Matias, Y., Norman, K.A., Devinsky, O., Hasson, U., 2022. Shared computational principles for language processing in humans and deep language models. *Nat. Neurosci.* 25, 369–380. <https://doi.org/10.1038/s41593-022-01026-4>.
- Grice, H.P., 1975. Logic and conversation. In: Cole, P., Morgan, J. (Eds.), *Syntax and Semantics*. Academic Press, New York, pp. 43–58.
- Hackel, L.M., Doll, B.B., Amodio, D.M., 2015. Instrumental learning of traits versus rewards: dissociable neural correlates and effects on choice. *Nat. Neurosci.* 1–8. <https://doi.org/10.1038/nn.4080>.
- Hayden, B.Y., Heilbronner, S.R., Pearson, J.M., Platt, M.L., 2011. Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J. Neurosci.* 31, 4178–4187. <https://doi.org/10.1523/JNEUROSCI.4652-10.2011>.
- Heilbron, M., Armeni, K., Schoffelen, J.M., Hagoort, P., de Lange, F.P., 2022. A hierarchy of linguistic predictions during natural language comprehension. *Proc. Natl. Acad. Sci.* 119 <https://doi.org/10.1073/pnas.2201968119>, 2020.12.03.410399.
- Hoerl, C., McCormack, T., 2019. Thinking in and about time: a dual systems perspective on temporal cognition. *Behav. Brain Sci.* 42 <https://doi.org/10.1017/S0140525X18002157>.
- Hoshi, Y., 2007. Functional near-infrared spectroscopy: current status and future prospects. *J. Biomed. Opt.* 12, 062106. <https://doi.org/10.1117/1.2804911>.
- Huppert, T.J., Diamond, S.G., Franceschini, M.A., Boas, D.A., 2009. Homer: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl. Opt.* 48, 0–33. <https://doi.org/10.1364/AO.48.00D280>.
- Jiang, J., Zheng, L., Lu, C., 2021. A hierarchical model for interpersonal verbal communication. *Soc. Cogn. Affect. Neurosci.* 16, 246–255. <https://doi.org/10.1093/scan/nsaa151>.
- Jiang, Y., Mi, Q., Zhu, L., 2023. Neurocomputational mechanism of real-time distributed learning on social networks. *10.1038/s41593-023-01258-y*.
- Kok, P., Mostert, P., Lange, F.P., 2017. Prior expectations induce prestimulus sensory templates 114. *10.1073/pnas.1705652114*.
- Konovalov, A., Hill, C., Daunizeau, J., Ruff, C.C., 2021. Dissecting functional contributions of the social brain to strategic behavior. *Neuron* 109, 3323–3337. <https://doi.org/10.1016/j.neuron.2021.07.025> e5.
- Kriegeskorte, N., Mur, M., Bandettini, P., 2008. Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2 <https://doi.org/10.3389/neuro.06.004.2008>.
- Kuperberg, G.R., Jaeger, T.F., 2016. What do we mean by prediction in language comprehension? *Lang. Cogn. Neurosci.* 31, 32–59. <https://doi.org/10.1080/23273798.2015.1102299>.
- Leonard, M.K., Baud, M.O., Sjerps, M.J., Chang, E.F., 2016. Perceptual restoration of masked speech in human cortex. *Nat. Commun.* 7, 1–9. <https://doi.org/10.1038/ncomms13619>.
- Levinson, S.C., 2016. Turn-taking in human communication – origins and implications for language processing. *Trends Cogn. Sci.* 20, 6–14. <https://doi.org/10.1016/j.tics.2015.10.010>.
- Mathys, C., Daunizeau, J., Friston, K.J., Stephan, K.E., 2011. A Bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* 5, 9. <https://doi.org/10.3389/fnhum.2011.00039>.
- Menenti, L., Petersson, K.M., Scheeringa, R., Hagoort, P., 2009. When elephants fly: differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. *J. Cogn. Neurosci.* 21, 2358–2368. <https://doi.org/10.1162/jocn.2008.21163>.
- Mi, Q., Wang, C., Camerer, C.F., Zhu, L., 2021. Reading between the lines: listener's vmPFC simulates speaker cooperative choices in communication games. *Sci. Adv.* 7 <https://doi.org/10.1126/sciadv.abe6276>.
- Molavi, B., Dumont, G.A., 2012. Wavelet-based motion artifact removal for functional near-infrared spectroscopy. *Physiol. Meas.* 33, 259–270. <https://doi.org/10.1088/0967-3334/33/2/259>.

- Pickering, M.J., Gambi, C., 2018. Predicting while comprehending language: a theory and review. *Psychol. Bull.* 144, 1002–1044. <https://doi.org/10.1037/bul0000158>.
- Rilling, J.K., Sanfey, A.G., 2011. The neuroscience of social decision-making. *Annu. Rev. Psychol.* 62, 23–48.
- Rouhani, N., Niv, Y., 2021. Signed and unsigned reward prediction errors dynamically enhance learning and memory. *Elife* 10, 1–28. <https://doi.org/10.7554/elife.61077>.
- Rouhani, N., Norman, K.A., Niv, Y., 2018. Dissociable effects of surprising rewards on learning and memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 44, 1430–1443. <https://doi.org/10.1037/xlm0000518>.
- Salazar, M., Shaw, D.J., Gajdo, M., Mare, R., Czekóová, K., Mikl, M., Brázdil, M., 2021. NeuroImage You took the words right out of my mouth: dual-fMRI reveals intra- and inter-personal neural processes supporting verbal interaction. *Neuroimage* 228.
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., Perner, J., 2014. Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neurosci. Biobehav. Rev.* 42, 9–34. <https://doi.org/10.1016/j.neubiorev.2014.01.009>.
- Seo, H., Lee, D., 2012. Neural basis of learning and preference during social decision-making. *Curr. Opin. Neurobiol.* 22, 990–995. <https://doi.org/10.1016/j.conb.2012.05.010>.
- Shamay-Tsoory, S.G., Saporta, N., Marton-Alper, I.Z., Gvirts, H.Z., 2019. Herding brains: a core neural mechanism for social alignment. *Trends Cogn. Sci.* 23, 174–186. <https://doi.org/10.1016/j.tics.2019.01.002>.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. <https://doi.org/10.1006/nimg.2001.0978>.
- Van Overwalle, F., 2009. Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 30, 829–858. <https://doi.org/10.1002/hbm.20547>.
- Vassena, E., Deraeve, J., Alexander, W.H., 2020. Surprise, value and control in anterior cingulate cortex during speeded decision-making. *Nat. Hum. Behav.* 4, 412–422. <https://doi.org/10.1038/s41562-019-0801-5>.
- Wicha, N.Y.Y., Moreno, E.M., Kutas, M., 2004. Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in spanish sentence reading. *J. Cogn. Neurosci.* 16, 1272–1288. <https://doi.org/10.1162/0898929041920487>.
- Willems, R.M., Frank, S.L., Nijhof, A.D., Hagoort, P., Van Den Bosch, A., 2016. Prediction during natural language comprehension. *Cereb. Cortex* 26, 2506–2516. <https://doi.org/10.1093/cercor/bhv075>.
- Yamagishi, T., Takagishi, H., De Souza Rodrigues Fermin, A., Kanai, R., Li, Y., Matsumoto, Y., 2016. Cortical thickness of the dorsolateral prefrontal cortex predicts strategic choices in economic games. *Proc. Natl. Acad. Sci. U. S. A.* 113, 5582–5587. <https://doi.org/10.1073/pnas.1523940113>.
- Ye, J.C., Tak, S., Jang, K.E., Jung, J., Jang, J., 2009. NIRS-SPM: statistical parametric mapping for near-infrared spectroscopy. *Neuroimage* 44, 428–447. <https://doi.org/10.1016/j.neuroimage.2008.08.036>.
- Yeshurun, Y., Nguyen, M., Hasson, U., 2021. The default mode network: where the idiosyncratic self meets the shared social world. *Nat. Rev. Neurosci.* 22, 181–192. <https://doi.org/10.1038/s41583-020-00420-w>.
- Zang, Y.F., Yong, H., Chao-Zhe, Z., Qing-Jiu, C., Man-Qiu, S., Meng, L., Li-Xia, T., Tian-Zi, J., Yu-Feng, W., 2007. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain Dev.* 29, 83–91. <https://doi.org/10.1016/j.braindev.2006.07.002>.
- Zhang, W., Yartsev, M.M., 2019. Correlated neural activity across the brains of socially interacting bats. *Cell* 178, 413–428. <https://doi.org/10.1016/j.cell.2019.05.023>.
- Zhang, Y., Brooks, D.H., Franceschini, M.A., Boas, D.A., 2005. Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging. *J. Biomed. Opt.* 10, 011014. <https://doi.org/10.1117/1.1852552>.
- Zheng, L., Chen, C., Liu, W., Long, Y., Zhao, H., Bai, X., Zhang, Z., Han, Z., Liu, L., Guo, T., Chen, B., Ding, G., Lu, C., 2018. Enhancement of teaching outcome through neural prediction of the students' knowledge state. *Hum. Brain Mapp.* 39, 3046–3057. <https://doi.org/10.1002/hbm.24059>.
- Zhou, S., Xu, X., He, X., Zhou, F., Zhai, Y., Chen, J., Long, Y., Zheng, L., Lu, C., 2023. Biasing the neurocognitive processing of videos with the presence of a real cultural other. *Cereb. Cortex* 33, 1090–1103. <https://doi.org/10.1093/cercor/bhac122>.