

EEG frequency tagging evidence of social interaction recognition

Danna Oomen,^{1,2} Emiel Cracco,^{1,2} Marcel Brass,^{3,4,2} and Jan R. Wiersema^{1,2}

¹Department of Experimental Clinical and Health Psychology, Ghent University, Ghent B-9000, Belgium

²EXPLORA, Ghent University, Ghent B-9000, Belgium

³Department of Experimental Psychology, Ghent University, Ghent B-9000, Belgium

⁴School of Mind and Brain/Department of Psychology, Humboldt Universität zu Berlin, Berlin 10099, Germany

Correspondence should be addressed to Danna Oomen, Department of Experimental Clinical and Health Psychology, Ghent University, Henri Dunantlaan 2, Ghent B-9000, Belgium. E-mail: danna.oomen@ugent.be.

Abstract

Previous neuroscience studies have provided important insights into the neural processing of third-party social interaction recognition. Unfortunately, however, the methods they used are limited by a high susceptibility to noise. Electroencephalogram (EEG) frequency tagging is a promising technique to overcome this limitation, as it is known for its high signal-to-noise ratio. So far, EEG frequency tagging has mainly been used with simplistic stimuli (e.g. faces), but more complex stimuli are needed to study social interaction recognition. It therefore remains unknown whether this technique could be exploited to study third-party social interaction recognition. To address this question, we first created and validated a wide variety of stimuli that depict social scenes with and without social interaction, after which we used these stimuli in an EEG frequency tagging experiment. As hypothesized, we found enhanced neural responses to social scenes with social interaction compared to social scenes without social interaction. This effect appeared laterally at occipitoparietal electrodes and strongest over the right hemisphere. Hence, we find that EEG frequency tagging can measure the process of inferring social interaction from varying contextual information. EEG frequency tagging is particularly valuable for research into populations that require a high signal-to-noise ratio like infants, young children and clinical populations.

Key words: EEG; frequency tagging; social interaction recognition

Introduction

Recognizing third-party social interactions is essential to navigate ourselves through the social world. This ability emerges early in development (Hamlin, 2013) and is shared with other primates (Sliwa and Freiwald, 2017). We use this ability to form impressions of others. When observing others interact, we are for instance able to tell with fair accuracy whether people are co-workers, friends or lovers (Costanzo and Archer, 1989), teasing or fighting (Sinke *et al.*, 2010; Cowell and Decety, 2015) and what their social status is (Mast and Hall, 2004). Recognizing third-party social interactions may also guide our own actions. For example, people tend to avoid passing through interacting people or social units (Efran and Cheyne, 1973; Knowles, 1973, 2015), and in accord with the impressions formed from third-party social interactions, we may in turn decide to avoid or approach someone (Quadflieg and Penton-Voak, 2017) or change our attitudes towards them (Christ *et al.*, 2014). Furthermore, interacting dyads receive preferential access to visual awareness (Su *et al.*, 2016), and dyad features as well as features of the individuals that make up a dyad are remembered better when the individuals interact

(Vestner *et al.*, 2019, but also see 2020 for an alternative explanation), further stressing the importance of third-party social interaction recognition.

Only recently, studies have started to investigate third-party social interaction recognition using neuroscientific methods to reveal the brain regions involved in, and the time course of, recognizing social interactions. Arioli and Canessa (2019) reported meta-analytic functional magnetic resonance imaging (fMRI) evidence for a ‘social interaction network’ including joint involvement of the action observation network for representing action meaning and the mentalizing network for representing others’ mental states, possibly in conjunction with an amygdala network for the evaluation of affective valence. Regarding the time course, Isik *et al.* (2020) used magnetoencephalography (MEG) decoding to investigate whether social interaction recognition is a primarily rapid feedforward process, like object recognition (about 150 ms), or a slower post-perceptual inference that requires iterative top-down computations. They found that social interaction recognition could only be read out from subjects’ MEG data 300 ms after image onset, well after feedforward visual processes. Their results

Received: 22 October 2021; Revised: 4 March 2022; Accepted: 20 April 2022

© The Author(s) 2022. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

suggest that even though social interaction recognition is spontaneous, it relies on slower post-perceptual inferences. Furthermore, event-related potential (ERP) research has focused on the differentiation of different types of social interaction and found distinct neural responses for, for instance, harm vs help in children (Cowell and Decety, 2015) and intentional vs unintentional harm in adults (Decety and Cacioppo, 2012).

In sum, prior neuroscience research has already provided valuable information on the processing of third-party social interaction recognition. Unfortunately, however, the methods used have an important limitation, that is, they are highly susceptible to noise (i.e. external noise from electronic equipment as well as movement and ocular artefacts) and therefore do not offer a high signal-to-noise ratio (SNR). Hence, long test sessions and large samples are required to somewhat mitigate this limitation and acquire robust data, which is not always ideal or feasible. This limitation and the accompanying requirements also make it particularly difficult to test certain populations like infants, young children or clinical populations. To overcome these challenges, here, we investigated whether a promising novel technique, electroencephalogram (EEG) frequency tagging, known for its high SNR (Norcia et al., 2015), can be used to measure social interaction recognition implicitly.

In short, EEG frequency tagging is a method based on the principle that stimuli presented periodically induce 'steady-state visual evoked potentials', also referred to as 'cyclical electrophysiological responses' (Retter and Rossion, 2016), in the EEG signal that are coupled to the stimuli (Norcia et al., 2015). For instance, stimuli presented two times per second (2 Hz) will induce a response in the EEG signal that is detectable in the frequency domain as a high amplitude 'peak' at exactly 2 Hz and at harmonic frequencies that are integer multiples of this fundamental frequency (4 Hz, 6 Hz, etc.; Regan, 1966). The key advantage of this approach over traditional ERP analyses is that brain responses can be identified objectively at a narrow, pre-defined frequency. Importantly, as a result, noise from other frequencies does not contaminate the response of interest, leading to a very high SNR. Furthermore, frequency tagging also does not require an explicit behavioural task or task comprehension, which makes this technique particularly interesting to use in young populations.

Although EEG frequency tagging was originally used to investigate low-level processes such as luminance information (e.g. Kamp et al., 1960), it has recently also been used to study more complex cognitive processes (Norcia et al., 2015), such as face processing (e.g. Alonso-Prieto et al., 2013) and perspective taking (Beck et al., 2018). Still, the stimulus complexity in these studies has been rather limited. More specifically, existing research has mainly used stimuli depicting objects/tools or (parts of) single agents. In contrast, research on social interaction processing requires stimuli that depict at least two agents. Adibpour et al. (2021) addressed this issue in their frequency tagging study by comparing facing and non-facing dyads. The results revealed stronger responses for facing agents than for non-facing agents. However, while two persons oriented towards each other may be a cue for social interaction, not all interactions involve facing and not all facing individuals interact. Indeed, recent work suggests that the facilitated detection of facing dyads may even reflect a general attentional mechanism not specific to social interaction (Vestner et al., 2022). Instead, social interaction is often inferred from the context. Although Adibpour et al. (2021) showed two agents, their stimuli were devoid of any contextual information and, therefore, did not resemble the complexity of

what we encounter in daily life. Hence, it remains unclear if frequency tagging can be used to measure the process of detecting social interaction from contextual information in variable situations, which, as mentioned above, is thought to require iterative top-down computations (Isik et al., 2020).

In the current study, we provide a direct test of whether EEG frequency tagging can be used to measure the process of inferring social interaction from context by measuring neural responses to rich social scenes depicting either social interaction or not. To this end, we first created a database of stimuli with and without social interaction, based on the PiSCES database (Teh et al., 2018). In study 1, we validated these stimuli and then used them in an EEG frequency tagging experiment in study 2. We had the following hypothesis: the neural response to scenes depicting interaction will be stronger than the neural response to scenes without interaction.

Validation study

Methods

Participants

Seventy participants were recruited using Prolific (www.prolific.co, accessed January 2020). The sample size was based on the sample size of the PiSCES database validation study (Teh et al., 2018). We restricted submission to West-European residents¹ aged 18–35 years without an autism spectrum disorder (ASD) diagnosis. The latter restriction was included as ASD is characterized by social interaction difficulties (American Psychiatric Association, 2013). The average age of our sample was 28 years old ($M = 27.97$, $s.d. = 4.81$), and our sample included 46 females (65%). The average years of education starting from year 1 (learning to read) was 16 years ($M = 16.17$, $s.d. = 4.23$), and the majority of participants were white ($n_{\text{White}} = 58$, $n_{\text{Hispanic or Latino}} = 3$, $n_{\text{Black}} = 3$, $n_{\text{Asian/ Pacific islander}} = 6$). Before the start of the survey, participants signed an online informed consent, after which they completed the online survey. The whole survey lasted about 60 min. Participants were reimbursed for their time. This study was conducted according to the ethical rules presented in the General Ethical Protocol of the Faculty of Psychology and Educational Sciences of Ghent University.

Stimuli

We aimed for a balance between experimental control and ecological validity in the stimuli. The database of images we created is based on the PiSCES database ('pictures with social context and emotional scenes'; Teh et al., 2018). The PiSCES database includes 203 black-and-white drawings, of which 100 depict multiple agents in interaction and 103 images depict a single agent. All images are normed on perceived valence, intensity and social engagement. Furthermore, all images include a situational context to resemble the complexity of real-life people and objects but are free from distractors that might influence interpretation (e.g. shadows and patterns on clothing).

In order for the control stimuli to meet our requirements (i.e. a set of images that depict two agents who are not in social interaction with each other), we merged (i.e. cross-paired) agents from the original 103 PiSCES single-agent images into one image. We crossed-paired only within valence and managed to create

¹ Defined as countries that are in the Western European and Other States Group, but not including Israel, Turkey, Australia, New Zealand, Canada and the USA.

47 control images. This set of control stimuli included 14 negatively valenced, 19 neutral valenced and 14 positively valenced images. For the social interaction stimuli, we erased agents from images that depicted more than two agents to match the number of agents of the control stimuli. One image was excluded because erasing an agent was not possible. Two other images were excluded because they depicted only one agent in interaction over the phone. We created five new interaction images from the erased agents, again, cross-pairing only within valence, resulting in a total of 102 social interaction images. The final set of social interaction stimuli included 28 negatively valenced, 30 neutral valenced and 44 positively valenced images. Altogether, our stimuli database included 149 (47 control and 102 social interaction) images depicting two agents performing an activity and/or embedded in a range of contextual settings (e.g. supermarket, school and restaurant). See Figure 2 for one example image of each category (control and social interaction). Photoshop was used to adjust and cross-pair the images (Adobe Photoshop CS5).

Norming task

The survey was set up using LimeSurvey (LimeSurvey Project Team, 2012). To collect norms, we asked participants to rate all images on four rating scales querying about the valence, intensity, social engagement and complexity of the image. First, participants judged the emotional valence experienced by the persons in the image (from 1 = strongly negative to 7 = strongly positive). Second, participants judged the intensity of feelings experienced by the persons in the image (i.e. perceived arousal; from 1 = extremely low to 7 = extremely high). Third, participants judged the degree of social interaction or engagement depicted in the image (from 1 = completely no interaction or engagement with the other person to 7 = extremely high degree of interaction or engagement with the other person). Fourth, participants rated the complexity of the image (from 1 = extremely low to 7 = extremely high). Additionally, we gave participants the option to comment on the image in case they noticed something unclear/strange/unusual in the image. All participants were presented with all 149 images in random order, with the order of the questions fixed.

The survey started with demographic questions (age, sex, years of education and ethnicity), the autism-spectrum quotient (AQ)-10 (Allison et al., 2012), and three questions from each of two scales: the free will inventory (Nadelhoffer et al., 2014) and the mentalization scale (Dimitrijević et al., 2018). The questionnaires were included for the purpose of another study and are therefore not further discussed here.

Results

Analyses were conducted following the analysis procedure of Teh et al. (2018). To detect outliers, for every scale and every participant, we correlated the participant's image ratings with the mean image ratings across all other participants. From the set of obtained correlation coefficients, we then identified the outliers per scale across participants (i.e. outside 1.5 times the interquartile range above/below the upper/lower quartile). We removed nine outlying participants for the valence scale, three for the intensity scale and eight for the social scale. The complexity scale had no outliers, but one participant was excluded as the participants' standard deviation for this scale was zero. The mean correlation over images between each participant's rating and the overall mean rating of all participants minus the participant themselves for the respective scales

were as follows: $M_{\text{correlation valence}} = 0.89$, $M_{\text{correlation intensity}} = 0.68$, $M_{\text{correlation social}} = 0.87$, $M_{\text{correlation complexity}} = 0.22$. The low correlation for the complexity scale likely reflects low validity due to the question being ambiguous. Therefore, we excluded this scale from further analyses.

To measure the inter-rater agreement, intra-class correlations (ICCs) were conducted for each scale using a multiple-raters, consistency, two-way random-effects model. The ICCs were excellent for all three scales: $ICC_{\text{valence}} = 1$, $ICC_{\text{intensity}} = 0.98$ and $ICC_{\text{social}} = 0.99$, indicating a high degree of agreement among raters on the valence, intensity and social engagement depicted on the images. The distributions of the image ratings for valence ($M = 4.25$, $s.d. = 1.46$), intensity ($M = 4.43$, $s.d. = 1.04$) and social ($M = 4.05$, $s.d. = 1.78$) are shown in Figure 1.

See Supplementary Material for a comparison of the mean ratings of the current study and the original study by Teh et al. (2018). The Supplementary Material also includes additional analyses on the relationships between emotional valence, intensity and social interaction similar to the original study by Teh et al. (2018).

Frequency tagging study

Using the validated stimuli from study 1, in study 2 we investigated whether EEG frequency tagging can be used to measure the process of inferring social interactions from context by measuring neural responses to rich social scenes depicting either social interaction or not. We had the following hypothesis: the neural response to scenes depicting interaction will be stronger than the neural response to scenes without interaction. We did not have a hypothesis concerning the topographical distribution of the effect.

In addition to this confirmatory research question, we also explored the relationship between ASD symptomatology and social interaction processing. To explain the social interaction difficulties associated with ASD (American Psychiatric Association, 2013), much ASD research has been conducted on the neural correlates of social stimulus (e.g. face) processing. However, less is known about the neural correlates of social interaction processing. In line with a behavioural study that found a negative correlation between the ability to recognize meaningful human interactions and ASD symptomatology (Van Boxtel et al., 2017), we expected social interaction recognition to correlate negatively with ASD symptomatology.

Methods

Participants

This study had the following inclusion criteria: normal or corrected-to-normal vision, no psychiatric or neurological condition and sufficient knowledge of the Dutch language. To determine the sample size, we conducted an a priori power analysis with a significance level of 0.05, a power level of 95% and an effect size of $d_z = 0.80$, chosen based on the fact that previous frequency tagging studies have reported large effects due to the high SNR (e.g. Norcia et al., 2015). This power analysis indicated that we needed at least 23 participants to achieve our goal. However, we decided to test five more participants to ensure a large enough sample even after the possible exclusion of bad-quality data. Eventually, no data had to be excluded based on data quality. Therefore, our final sample contained 28 participants. Note that of those 28 participants, one participant was excluded and replaced with another participant because it was revealed after the test session that the participant had an attention deficit hyperactive disorder (ADHD) diagnosis, not meeting the inclusion

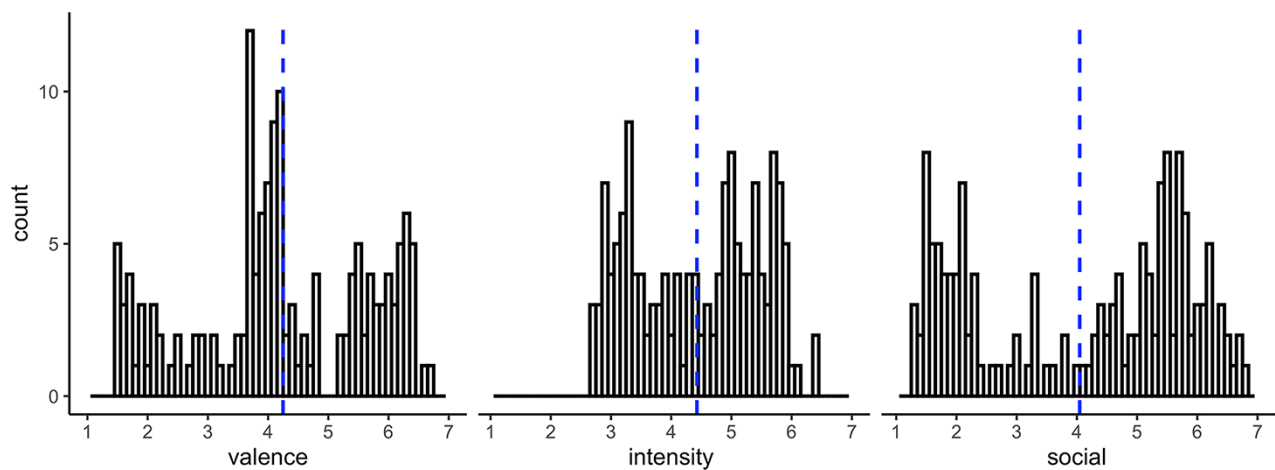


Fig. 1. Distribution of picture ratings with the mean dashed blue line for valence, intensity and social interaction.

Table 1. Participant characteristics

	M (s.d.)	Range
Age	22.43 (1.77)	20–29
Gender (% female)	50%	n/a
Years of education	15.89 (1.83)	10–19
Ethnicity (% White)	92.9%	n/a

Note. $N = 28$; ethnicity: one participant was Asian, and one participant was half White, half South-African.

criteria. See Table 1 for participant characteristics. Participants were reimbursed for their time. This study was approved by the Research Ethics Committee of the Faculty of Psychology and Educational Sciences (2020/112).

Frequency tagging task and procedure

Participants were seated in a Faraday cage of ~80–100 cm from a 24-inch computer monitor with a refresh rate 60 Hz. Before the start of the experiment, all participants signed informed consent and completed a questionnaire.

Subsequently, participants completed the frequency tagging task programmed in PsychoPy3 (Peirce et al., 2019). Stimuli were presented by sinusoidal contrast modulation at a frequency of 1.66 Hz (600 ms per stimulus). The task included four types of stimuli: images that depicted social interaction, images that depicted no social interaction and the scrambled versions of both image types. Scrambled stimuli were included to control for potential low-level differences between the two image types. The four stimulus types were presented in separate blocks, and each block was presented four times resulting in 16 blocks in total. A block consisted of 110 stimuli, and a single block lasted 66 s. A block started and ended with a 3 s fade-in (0–100%) and 3 s fade-out (100–0%) period to avoid abrupt eye movements and blinks due to the sudden (dis)appearance of stimuli.

The task included no practice block, but participants were told that they would see images of social interaction, no social interaction and scrambled images. They also saw two example images from the social interaction and no social interaction categories. These example images were not included as stimuli in the actual task. We decided to be open about the included stimulus categories because with block designs, participants often become aware of the categories during the task regardless.

By giving participants this information upfront, we reduce variability regarding the time point at which participants recognized the stimulus categories. Participants were instructed to fixate on a black cross positioned in the centre of the stimuli while also continuously monitoring the presented images. To ensure a constant level of attention throughout the task and across conditions, participants had to press the spacebar as fast and accurately as possible when they detected (400 ms) a colour change of the fixation cross (which randomly occurred three to six times per block). Detection was high overall (97% on average), with no difference between the interaction and no interaction conditions when presented normal or scrambled ($P = 0.841$, $P = 0.481$). The participants were further informed that they would be presented with eight unscrambled images after the task and would have to indicate which four images had appeared during the task. On average, participants recognized the same number of images for both categories, namely 1.27 of two images. The sole aim of these instructions (i.e. identification of fixation cross colour change and identification of the shown images) was to encourage participants to pay attention to the presented stimuli. The total task duration was ~18 min, after which they completed two additional frequency tagging tasks, addressing a different research question, not further described here. The whole experiment, including the two additional frequency tagging tasks, lasted around 1 h.

After the experiment, participants filled in a questionnaire, namely, the Dutch version of the AQ (Baron-Cohen et al., 2001; Dutch version: Hoekstra et al., 2010). The AQ was administered to explore the relationship between ASD symptomatology and social interaction processing, in addition to our confirmatory research question. The AQ contains 50 items. Items are rated on a 4-point Likert scale ranging from 1 ('definitely agree') to 4 ('definitely disagree'). Reverse items are included in such a way that agreement suggests the presence of an autistic-like trait in approximately half of the items. Each of the items was later recoded to either 0 or 1 (total score ranging from 0 to 50; Baron-Cohen et al., 2001), with higher scores indicating more autistic-like traits. The internal consistency of the AQ in the current sample was acceptable, $\alpha = 0.71$.

Stimuli

Stimuli were selected from the *adjusted* PiSCES database (validation study) and included 36 images with a mean rating of three or lower on the social interaction scale for the no interaction

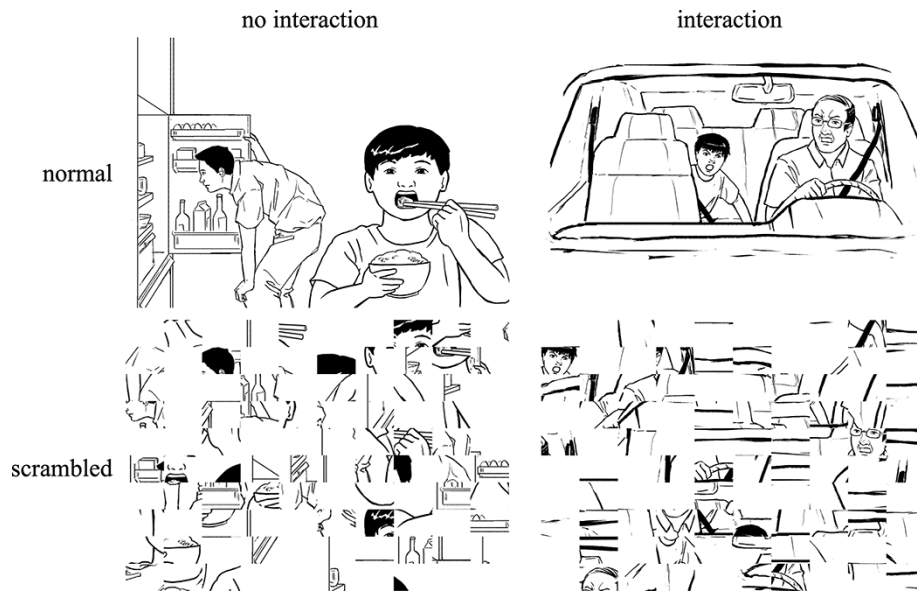


Fig. 2. One example image for each stimuli type (i.e. social interaction normal, no social interaction normal and the corresponding scrambled versions).

stimulus category ($M = 1.89$, $s.d. = 0.36$; range = 1.32–2.66) and 36 interaction images with a mean rating of five or higher for the interaction stimulus category ($M = 5.74$, $s.d. = 0.37$; range = 5.05–6.62). The stimuli from these two categories were balanced for valence, $t(69.71) = -1.01$, $P = 0.315$, and intensity, $t(63.47) = -1.57$, $P = 0.121$. The corresponding control stimuli were created by scrambling the images into a 10×10 grid. Stimuli were presented at the centre of the screen and were drawn randomly from their respective categories, never repeating the same stimulus back-to-back. See Figure 2 for an example stimulus of each category. The stimuli can be found on the Open Science Framework (https://osf.io/4r7yp/?view_only=44ee621078ea46f3893e46f9e13412c5).

EEG recording and pre-processing

EEG was continuously recorded from 64 scalp sites using an ActiCHamp amplifier and BrainVisionRecorder software (version 1.21.0402, Brain Products, Gilching, Germany). Ag/AgCl (active) electrodes were mounted in an elastic cap (ActiCAP, Munich, Germany). Electrode positions were based on the 10% system with the exception of two electrodes (TP9 and TP10) that were placed at OI1h and OI2h according to the 5% system, as we mainly expected posterior occipital activation. During EEG recording, all channels were referenced to Fz, and the sampling rate was 1000 Hz. Horizontal electro-oculogram (EOG) was recorded with FT9 and FT10 electrodes embedded in the cap. Vertical EOG was recorded with additional bipolar Ag/AgCl sintered ring electrodes placed above and below the left eye. Off-line processing of the EEG signal was done using Letswave 6 (www.letswave.org). First, a fourth-order Butterworth band-pass filter (0.1–100 Hz) was applied, after which we segmented the data to obtain epochs extending from 2 s before to 68 s after the stimulus onset. Next, ocular artefacts were removed with an independent component analysis (ICA) on the merged segmented data using the Runica algorithm and a square matrix. For each participant, the first 10 independent components (ICs) were inspected, and the IC related to eye blinks was removed manually. After ICA, we interpolated noisy or faulty electrodes. More specifically, we interpolated one electrode for seven participants using data from three (or two in case of OI1/2h electrodes)

neighbouring electrodes. We then re-referenced the signal to an average reference, before cropping the segments into 60 s epochs (3–63 s), cutting out the fade in and fade out periods. Finally, trials per condition were averaged, and subsequently, a fast Fourier transform was applied to the data of each electrode to normalized (divided by $N/2$) amplitudes (μV) in the frequency domain.

Statistical analysis approach

The statistical data-analysis approach of this study was pre-registered (<https://aspredicted.org/si3s4.pdf>), while data collection was ongoing but before pre-processing and analyses of the data. The data and analyses can be found on the Open Science Framework (https://osf.io/4r7yp/?view_only=44ee621078ea46f3893e46f9e13412c5). For the data analyses, we computed the signal to noise-subtracted amplitudes (SNS) at each frequency bin by subtracting the average voltage amplitude of the 20 neighbouring bins (10 on each side, excluding the immediately adjacent bin; Norcia *et al.*, 2015) from the amplitude of the frequencies of interest. Based on a visual inspection of pilot data (see Supplementary Figure 2 for a visualization of the SNR pilot data across all conditions and over electrodes of interest) not included in this study, and as preregistered, the SNS was calculated as the sum of the first eight harmonics (i.e. the frequencies of interest: 1.66, 3.33, 5.00, 6.66, 8.33, 10.00, 11.66 and 13.33 Hz; Retter *et al.*, 2021). To ensure an unbiased (independent of condition effects or hypotheses) selection of electrodes, regions of interest (ROIs) were chosen based on visual inspection of the scalp topography across participants and conditions (collapsed localizer approach, Luck and Gaspelin, 2017, see Supplementary Figure 3 for the collapsed scalp topography), which revealed lateral posterior activity with a maximum at PO8. The right posterior cluster included PO8 and two neighbouring electrodes (PO4 and O2), and the left posterior cluster included the corresponding electrodes on the left hemisphere (PO7, PO3 and O1). See Figure 3 for a visualization of the SNR of the included harmonics across all conditions and over electrodes of interest. The SNR computation was identical to the SNS computation except that division was used instead of subtraction. On the SNS data,

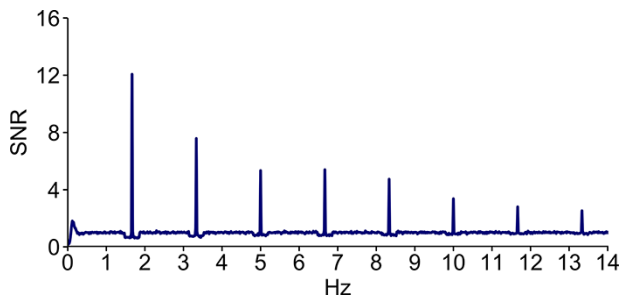


Fig. 3. SNR across all conditions and over electrodes of interest. See Supplementary Figure 4 for SNR plots per condition.

we conducted a $2 \times 2 \times 2$ repeated measures analysis of variance (ANOVA) with laterality (left and right), stimulus type (scrambled and normal) and interaction type (no interaction and interaction) as within-subject factors. Although not specified in our preregistration, we included laterality as an additional factor, given that visual inspection of our collapsed localizer revealed activity to be lateralized. Significant interactions were followed-up by two-tailed paired-sample *t*-tests. For the *t*-test, *P*-values as well as Bayes factors (BFs) are reported. BFs were calculated with a non-informative Jeffreys prior and a Cauchy prior on the standardized effect size (Rouder et al., 2012). As we expected large effect sizes, we used a wide prior ($r = 1$).

As preregistered, we also explored the relationship between ASD symptomatology and social interaction processing. For this, we performed a Spearman rho correlation between AQ total scores and the Stimulus Type \times Interaction Type interaction effect [(normal interaction – scrambled interaction) – (normal no interaction – scrambled no interaction)]. Here too, we reported *P*-values as well as BFs, but for these correlations, we used default priors.

Results

The repeated measures ANOVA revealed significant main effects of all three factors. The main effect of Laterality showed a stronger response in the right cluster ($M = 4.71$, $s.d. = 1.77$) than in the left cluster ($M = 4.08$, $s.d. = 1.97$), $F(1, 27) = 7.08$, $P = 0.013$, $\eta_p^2 = 0.21$. The main effect of Stimulus Type was due to a stronger response for the normal stimuli ($M = 4.78$, $s.d. = 2.12$) than for the scrambled stimuli ($M = 4.01$, $s.d. = 1.47$), $F(1, 27) = 19.73$, $P < 0.001$, $\eta_p^2 = 0.42$. Finally, the main effect of Interaction Type indicated a stronger response for the interaction stimuli ($M = 4.49$, $s.d. = 1.80$) than for the no-interaction stimuli ($M = 4.30$, $s.d. = 1.75$), $F(1, 27) = 11.45$, $P = 0.002$, $\eta_p^2 = 0.30$. The repeated measures ANOVA further revealed a significant Laterality \times Interaction Type interaction effect, $F(1, 27) = 10.27$, $P = 0.003$, $\eta_p^2 = 0.28$, and in line with our hypothesis, a Stimulus Type \times Interaction Type interaction effect, $F(1, 27) = 27.67$, $P < 0.001$, $\eta_p^2 = 0.51$, as well as a Laterality \times Stimulus Type \times Interaction Type interaction effect, $F(1, 27) = 13.04$, $P = 0.001$, $\eta_p^2 = 0.33$. There was no interaction between Laterality and Stimulus Type, $P = 219$.

To follow-up on the Laterality \times Stimulus Type \times Interaction Type interaction effect, we looked at the Stimulus Type \times Interaction Type interaction effect, our effect of interest, separately for the left and right cluster. The left cluster Stimulus Type \times Interaction Type interaction, $F(1, 27) = 9.91$, $P = 0.004$, $\eta_p^2 = 0.27$, revealed a stronger response for interacting than for non-interacting stimuli when they were presented normally, $t(27) = 2.62$, $P = 0.014$, $BF_{10} = 2.87$, $d_z = 0.49$, but not when they

were scrambled, $t(27) = -1.26$, $P = 0.22$, $BF_{10} = 0.31$, $d_z = 0.24$. The right cluster Stimulus Type \times Interaction Type interaction, $F(1, 27) = 29.77$, $P < 0.001$, $\eta_p^2 = 0.52$, revealed a similar effect, with a stronger response for interacting than for non-interacting stimuli when they were presented normally, $t(27) = 6.19$, $P < 0.001$, $BF_{10} = 15\,321.37$, $d_z = 1.17$, but not when they were scrambled, $t(27) = -1.02$, $P = 0.318$, $BF_{10} = 0.24$, $d_z = 0.19$. However, the Interaction Type effect for normal stimuli was stronger in the right cluster than in the left cluster, $t(27) = 4.17$, $P < 0.001$, $BF_{10} = 99.03$, $d_z = 0.40$. See Figure 4 for a bar plot of the SNS per condition and for the two regions separately, Figure 5 for the topographies of all stimulus conditions and Supplementary Table 1 for means and standard deviations of all stimulus conditions for the two regions separately.

In addition to these confirmatory analyses, we also ran an explorative analysis investigating whether social interaction processing, quantified as the Stimulus Type \times Interaction Type interaction effect [(normal interaction – scrambled interaction) – (normal no interaction – scrambled no interaction)], correlates with ASD symptomatology, quantified as the AQ total score. This revealed anecdotal evidence for a positive correlation between the right cluster Stimulus Type \times Interaction Type interaction effect and AQ total scores (left cluster: $r_s = 0.17$, $P = 0.37$, $BF_{10} = 0.67$; right cluster: $r_s = 0.47$, $P = 0.011$, $BF_{10} = 1.57$).

Discussion

Previous neuroscience studies have already provided important insights into the neural processing of third-party social interactions (Decety and Cacioppo, 2012; Cowell and Decety, 2015; Arioli and Canessa, 2019; Isik et al., 2020). However, these studies all used methods that are highly susceptible to noise (e.g. fMRI, MEG and ERP) and therefore require large samples and long experiments. This is not always feasible and poses an important challenge for researchers studying infant, children or clinical samples. A possible solution to this problem could be to use EEG frequency tagging, as this technique is known to have a very high SNR (Norcia et al., 2015), but EEG frequency tagging has so far only been used to study relatively basic social processes such as face perception (Alonso-Prieto et al., 2013) and distinguishing facing from non-facing dyads (Adibpour et al., 2021). Therefore, in the current study, we investigated whether it can also be used to investigate the more complex social process of inferring social interaction from contextual information (Isik et al., 2017).

To this end, we first created and validated a database of stimuli that depict two agents either in interaction or not. As we sought to capture social interaction recognition from contextual information and not from a single social cue such as facing or not facing (Adibpour et al., 2021), we used a wide variety of social stimuli that differed greatly from each other in terms of agent configuration (e.g. facing or not facing), agent activity (e.g. talking or paying) and/or contextual background (e.g. supermarket and school). As such, the only systematic difference between the interaction and no-interaction stimuli was the absence or presence of interaction. Additionally, the wide variety of stimuli increased ecological validity because the rich social scenes resemble the complexity of what we encounter in daily life. This validated database was subsequently used in an EEG frequency tagging experiment. As hypothesized, and in line with previous findings that highlight the saliency of third-party social interaction (Su et al., 2016), we found enhanced neural responses to social scenes with social interaction compared to social scenes without social interaction.

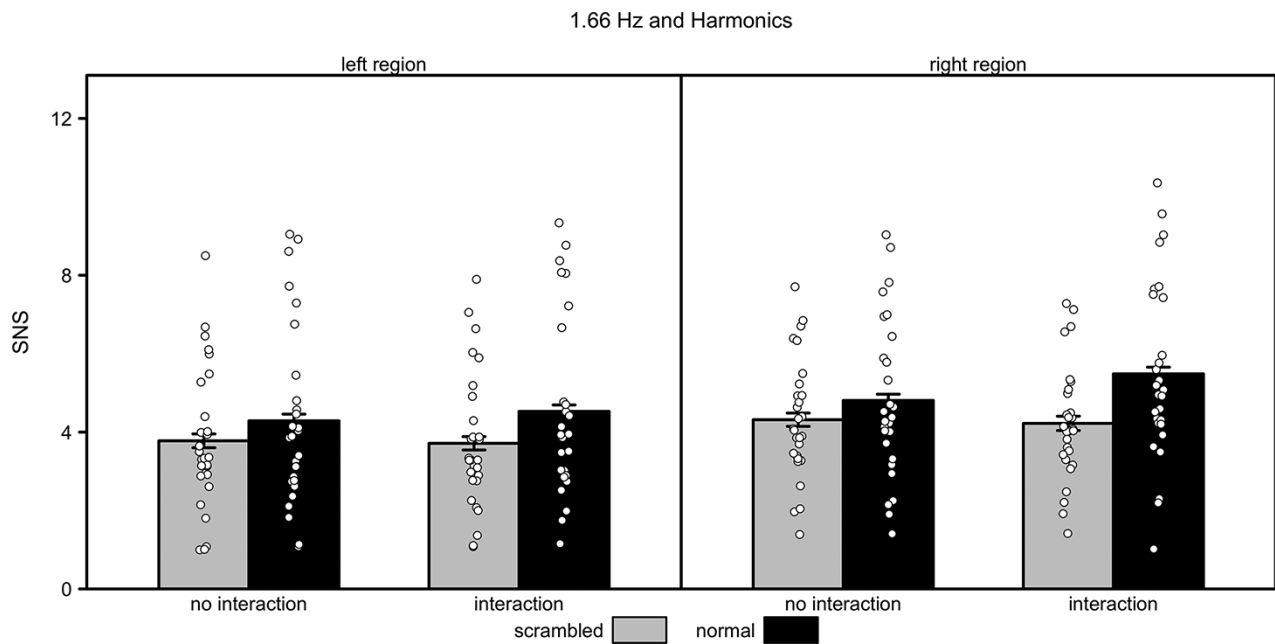


Fig. 4. SNS per condition and for the two regions separately. Error bars represent SEMs corrected for within-subject designs (Morey, 2008).

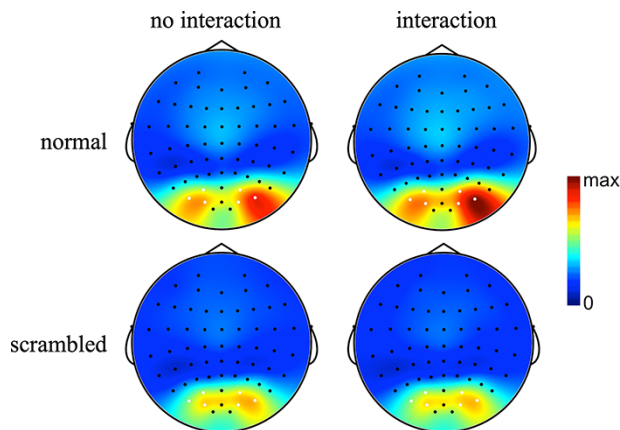


Fig. 5. Topographies per condition. Topographies are scaled from 0 to the maximum amplitude across the four conditions (i.e. $6.53 \mu\text{V}$). Included electrodes are indicated in white.

This effect appeared laterally at occipitoparietal electrodes and strongest over the right hemisphere.

Our findings imply that EEG frequency tagging can be used in future investigations on social interaction recognition, a complex cognitive process thought to require iterative top-down computations (Isik et al., 2020). This is positive because EEG frequency tagging has the key advantage of being largely resistant to noise and therefore of providing a very high SNR. This advantage is primarily driven by the fact that frequency tagging elicits a narrow response in a predefined frequency band, determined by the stimulus. As such, frequency tagging not only reduces noise but also makes it possible to objectively define the response.

Although these strengths are interesting in general for future research on social interaction recognition, they should prove particularly useful for research into different developmental stages (e.g. infants and young children) and clinical populations. Ideally, tasks for infants and children should be kept as short

as possible to accommodate their short attention span, and for certain populations that experience a diminished ability to concentrate (e.g. depression and ADHD; American Psychiatric Association, 2013). By increasing the SNR, EEG frequency tagging allows researchers to achieve this goal. Furthermore, developmental research requires a technique largely resistant to noise, as infants and young children hardly remain still (Raschle et al., 2012; Maguire et al., 2014; Azhari et al., 2020). Finally, clinical population studies often compare groups (e.g. clinical vs control group), and a high SNR is therefore needed to have enough sensitivity to be able to detect group differences. In the current study, we exploratively looked at the relationship between ASD symptomatology and social interaction processing and found anecdotal evidence for a correlation in the opposite direction as expected. This would be an interesting relationship to explore further in an autism vs neurotypical group study, given that ASD is characterized by social interaction difficulties (American Psychiatric Association, 2013). Besides the possibilities this frequency tagging method has for future clinical and developmental research into social interaction recognition, future studies may also use it to investigate more fundamental questions. The current study just takes the first step of investigating social interaction recognition with frequency tagging, but plenty of questions could still be investigated using this method, such as how social interaction processes are modulated by factors like number or proximity of agents, valence or type of interaction (e.g. joint attention, joined actions, mutual eye contact and touch).

Our findings further tentatively suggest that frequency tagging could also be used to study other complex cognitive processes. However, it is important to note that our design might have played a role in this. We used a block design instead of the nowadays more commonly used oddball design (Heinrich et al., 2009). In oddball studies, standard stimuli are presented together with oddball stimuli at fixed intervals. In these tasks, responses to the oddball stimuli represent a differential response to the oddball stimuli in contrast to the standard stimuli. For instance, presenting no social interaction stimuli at a frequency of 2.5 Hz, with social

interaction stimuli embedded every fifth stimuli (i.e. $2.5/5 = 0.5$ Hz oddball rate), should result in not only a peak at exactly 2.5 Hz, but also at 0.5 Hz, if social interaction stimuli are discriminated from no social interaction stimuli. Therefore, an advantage of oddball tasks is that there is no need to subtract two blocks (Norcia et al., 2015). However, pilot data from an oddball task acquired in preparation of this study showed no oddball response to our stimuli. Although speculative at this point, this lack of oddball response might be due to the complexity of our stimuli, which may take longer to process. Therefore, these stimuli may require a lower frequency rate than 2.5 Hz, like the one used in the current study (1.66 Hz), in order for them to be properly processed. This is, however, not ideal for oddball tasks, as a lower standard frequency rate would also imply an even lower oddball frequency rate that is likely too low, considering that experimental noise is present particularly at the lowest frequencies (Norcia et al., 2015). Thus, slower paradigms like the block design used here are perhaps more suitable to study complex cognitive processes.

Block designs do have the limitation of expectation and anticipation influences, and therewith attention influences, about the nature of the stimuli presented. However, the idea that participants paid more attention to either one of the stimulus categories is not supported by our attention check data. The memory task showed that, on average, participants recognized an identical number of images for both categories. There was also no difference in performance on the detection task between the interaction and no interaction blocks. Furthermore, this limitation is by no means unique to the method (EEG frequency tagging) or design (blockwise) used here. Other techniques such as ERP, fMRI and MEG are susceptible to attentional differences between conditions as well. Similarly, using another design, like a frequency tagging oddball design, for example, could have given rise to an explanation in terms of bottom-up attention. That is, one could then explain oddball responses as attentional capture of 'interesting stimuli', unspecific to interaction processing.

Furthermore, the current study results are limited to how we provided task instructions. More specifically, in this study, participants were informed in the instructions that they would see stimuli of social interaction and no social interaction. Although previous studies did not always explicitly include the stimulus categories of interest in the instructions, the categories are usually blatantly obvious: different vs identical faces (Alonso-Prieto et al., 2013), animals vs non-animals, birds vs non-birds, natural vs non-natural: (Stothart et al., 2017) and facing vs non-facing people (Adibpour et al., 2021); hence, awareness of the stimulus categories is not unique to this study. Furthermore, similar to other frequency tagging studies, participants were not informed about the research question, and the categories of interest did not connect to the tasks they had to perform (i.e. discriminating the two categories was not part of the tasks). However, an interesting question that remains unanswered, and that future research should address, is whether prior awareness of the stimulus categories modulates the brain response. In a similar study, Isik et al. (2020) found that social interaction recognition did not occur earlier when subjects performed an explicit compared with an implicit social interaction detection task. This gives some reason to believe that steady-state visual evoked potentials magnitude would also not depend on stimulus awareness. However, only a direct test can provide a definite answer.

Lastly, it should also be mentioned that although frequency tagging has some benefits over other neuroscientific methods, it does have its own limitations. Therefore, frequency tagging is not necessarily better than other neuroscience techniques.

For example, although it has the advantage of a high SNR, this comes at the cost of reduced temporal sensitivity. Similarly, like all measures based on EEG, spatial localization is relatively poor. Hence, which technique to use depends first and foremost on the research question and population. Given its high sensitivity, frequency tagging is a particularly useful technique for studies with clinical populations or infants, where large samples or long experiments are often not feasible. Our findings further indicate that frequency tagging can be used in those samples not only to study lower-level visual processes but also to study higher-level social processes. That said, frequency tagging alone can only reveal so much of the underlying process, and to obtain a complete picture of a process, it should therefore be studied with several complementary techniques to account for the inherent limitations of each technique.

To conclude, the current EEG frequency tagging study found evidence for enhanced neural responses to scenes depicting social interaction relative to scenes without social interaction. Our results therefore indicate that EEG frequency tagging can be used in future studies on social interaction recognition. The strengths of EEG frequency tagging will in particular open doors for expanding this research line to different developmental stages and clinical populations.

Acknowledgements

The authors would like to thank all participants for their contribution as well as Joyce Scheirlinckx for her assistance with data collection and Sara Cnudde for her help with participant recruitment.

Funding

D.O. was supported by the Special Research Fund of Ghent University (BOF18/DOC/348). E.C. was supported by two post-doctoral fellowships awarded by the Research Foundation Flanders (FWO18/PDO/049 and 12U0322N). M.B. was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) under Germany's Excellence Strategy—EXC 2002/1 'Science of Intelligence'—project number 390523135 and supported by an Einstein Strategic Professorship (Einstein Foundation Berlin). The funding bodies had no role in the design of the study, collection, analysis, interpretation of data or writing the manuscript.

Conflict of interest

The authors declared that they had no conflict of interest with respect to their authorship or the publication of this article.

Supplementary data

Supplementary data are available at SCAN online.

References

- Adibpour, P., Hochmann, J.-R., Papeo, L. (2021). Spatial relations trigger visual binding of people. *Journal of Cognitive Neuroscience*, **33**(7), 1343–53.
- Allison, C., Auyeung, B., Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: the short Autism Spectrum Quotient and the short Quantitative Checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child and Adolescent Psychiatry*, **51**(2), 202–12.e7.
- Alonso-Prieto, E., Belle, G., Van Liu-Shuang, J., Norcia, A.M., Rossion, B. (2013). The 6Hz fundamental stimulation frequency

- rate for individual face discrimination in the right occipito-temporal cortex. *Neuropsychologia*, **51**(13), 2863–75.
- American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders (5th ed.). [10.1176/appi.books.9780890425596](https://doi.org/10.1176/appi.books.9780890425596).
- Arioli, M., Canessa, N. (2019). Neural processing of social interaction: coordinate-based meta-analytic evidence from human neuroimaging studies. *Human Brain Mapping*, **40**(13), hbm.24627.
- Azhari, A., Truzzi, A., Neoh, M.J.Y., et al. (2020). A decade of infant neuroimaging research: what have we learned and where are we going? *Infant Behavior & Development*, **58**, 101389.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, **31**(1), 5–17.
- Beck, A.A., Rossion, B., Samson, D. (2018). An objective neural signature of rapid perspective taking. *Social Cognitive and Affective Neuroscience*, **13**(1), 72–9.
- Christ, O., Schmid, K., Lollot, S., et al. (2014). Contextual effect of positive intergroup contact on outgroup prejudice. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(11), 3996–4000.
- Costanzo, M., Archer, D. (1989). Interperceiving the expressive behavior of others: the Interpersonal Perception Task. *Journal of Nonverbal Behavior*, **13**(4), 225–45.
- Cowell, J.M., Decety, J. (2015). The neuroscience of implicit moral evaluation and its relation to generosity in early childhood. *Current Biology*, **25**(1), 93–7.
- Decety, J., Cacioppo, S. (2012). The speed of morality: a high-density electrical neuroimaging study. *Journal of Neurophysiology*, **108**(11), 3068–72.
- Dimitrijević, A., Hanak, N., Altaras Dimitrijević, A., Jolić Marjanović, Z. (2018). The Mentalization Scale (Ments): a self-report measure for the assessment of mentalizing capacity. *Journal of Personality Assessment*, **100**(3), 268–80.
- Efran, M.G., Cheyne, J.A. (1973). Shared space: the co-operative control of spatial areas by two interacting individuals. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences du Comportement*, **5**(3), 201–10.
- Hamlin, J.K. (2013). Moral judgment and action in preverbal infants and toddlers: evidence for an innate moral core. *Current Directions in Psychological Science*, **22**(3), 186–93.
- Heinrich, S.P., Mell, D., Bach, M. (2009). Frequency-domain analysis of fast oddball responses to visual stimuli: a feasibility study. *International Journal of Psychophysiology*, **73**(3), 287–93.
- Hoekstra, R.A., Vinkhuyzen, A.A.E., Wheelwright, S., et al. (2010). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-Short). *Journal of Autism and Developmental Disorders* 2010, **41**(5), 589–96.
- Isik, L., Koldewyn, K., Beeler, D., Kanwisher, N. (2017). Perceiving social interactions in the posterior superior temporal sulcus. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(43), E9145–52.
- Isik, L., Mynick, A., Pantazis, D., Kanwisher, N. (2020). The speed of human social interaction perception. *NeuroImage*, **215**, 116844.
- Kamp, A., Sem-Jacobsen, C.W., van Leeuwen, W.S., van der Tweel, L.H. (1960). Cortical responses to modulated light in the human subject. *Acta Physiologica Scandinavica*, **48**(1), 1–12.
- Knowles, E.S. (1973). Boundaries around group interaction: the effect of group size and member status and boundary permeability. *Journal of Personality and Social Psychology*, **26**(3), 327–31.
- Knowles, E.S. (2015). Spatial behavior of individuals and groups. In: Paulus, P.B., editors. *Psychology of Group Influence*, 2nd edn, New York, NY: Psychology Press, 53–86.
- LimeSurvey Project Team. (2012). LimeSurvey: An Open Source survey tool/LimeSurvey GmbH. Hamburg, Germany. Available: <http://www.limesurvey.org>.
- Luck, S.J., Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, **54**(1), 146.
- Maguire, M.J., Magnon, G., Fitzhugh, A.E. (2014). Improving data retention in EEG research with children using child-centered eye tracking. *Journal of Neuroscience Methods*, **238**, 78.
- Mast, M.S., Hall, J.A. (2004). Who is the boss and who is not? Accuracy of judging status. *Journal of Nonverbal Behavior*, **28**(3), 145–65.
- Morey, R.D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, **4**(2), 61–4.
- Nadelhoffer, T., Shepard, J., Nahmias, E., Sripada, C., Ross, L.T. (2014). The free will inventory: measuring beliefs about agency and responsibility. *Consciousness and Cognition*, **25**(1), 27–41.
- Norcia, A.M., Gregory Appelbaum, L., Ales, J.M., Cottareau, B.R., Rossion, B. (2015). The steady-state visual evoked potential in vision research: a review. *Journal of Vision*, **15**(6), 1–46.
- Peirce, J., Gray, J.R., Simpson, S., et al. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, **51**(1), 195–203.
- Quadflieg, S., Penton-Voak, I.S. (2017). The emerging science of people-watching: forming impressions from third-party encounters. *Current Directions in Psychological Science*, **26**(4), 383–9.
- Raschle, N., Zuk, J., Ortiz-Mantilla, S., et al. (2012). Pediatric neuroimaging in early childhood and infancy: challenges and practical guidelines. *Annals of the New York Academy of Sciences*, **1252**(1), 43.
- Regan, D. (1966). Some characteristics of average steady-state and transient responses evoked by modulated light. *Electroencephalography and Clinical Neurophysiology*, **20**(3), 238–48.
- Retter, T.L., Rossion, B., Schiltz, C. (2021). Harmonic amplitude summation for frequency-tagging analysis. *Journal of Cognitive Neuroscience*, **33**(11), 2372–93.
- Retter, T.L., Rossion, B. (2016). Uncovering the neural magnitude and spatio-temporal dynamics of natural image categorization in a fast visual stream. *Neuropsychologia*, **91**, 9–28.
- Rouder, J.N., Morey, R.D., Speckman, P.L., Province, J.M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, **56**(5), 356–74.
- Sinke, C.B.A., Sorger, B., Goebel, R., de Gelder, B. (2010). Tease or threat? Judging social interactions from bodily expressions. *NeuroImage*, **49**(2), 1717–27.
- Sliwa, J., Freiwald, W.A. (2017). Neuroscience: a dedicated network for social interaction processing in the primate brain. *Science*, **356**(6339), 745–9.
- Stothart, G., Quadflieg, S., Milton, A. (2017). A fast and implicit measure of semantic categorisation using steady state visual evoked potentials. *Neuropsychologia*, **102**, 11–8.
- Su, J., Van Boxtel, J.J.A., Lu, H. (2016). Social interactions receive priority to conscious perception. *PLoS One*, **11**, 8.
- Teh, E.J., Yap, M.J., Liow, S.J.R. (2018). PiSCES: pictures with social context and emotional scenes with norms for emotional valence,

- intensity, and social engagement. *Behavior Research Methods*, **50**(5), 1793–805.
- Van Boxtel, J.J., Peng, Y., Su, J., Lu, H. (2017). Individual differences in high-level biological motion tasks correlate with autistic traits. *Vision Research*, **141**, 136–44.
- Vestner, T., Tipper, S.P., Hartley, T., Over, H., Rueschemeyer, S.A. (2019). Bound together: social binding leads to faster processing, spatial distortion, and enhanced memory of interacting partners. *Journal of Experimental Psychology: General*, **148**(7), 1251–68.
- Vestner, T., Gray, K.L.H., Cook, R. (2020). Why are social interactions found quickly in visual search tasks? *Cognition*, **200**, 104270.
- Vestner, T., Over, H., Gray, K.L.H., Cook, R. (2022). Objects that direct visuospatial attention produce the search advantage for facing dyads. *Journal of Experimental Psychology: General*, **151**(1), 161–71.