

**THE INFLUENCE OF JOINT ATTENTION AND PARTNER TRUSTWORTHINESS  
ON CROSS-MODAL SENSORY CUEING**

**Maartje C. de Jong\* and H. Chris Dijkerman**

**Published in Cortex (2019), 119: 1-11.**

Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, the Netherlands

\*Corresponding author.

Spinoza Centre for Neuroimaging,

Meibergdreef 75,

1105 BK, Amsterdam, the Netherlands.

E-mail address: [dejongmc@gmail.com](mailto:dejongmc@gmail.com) (M.C. de Jong)

**Keywords**

Joint attention, alpha oscillations, visuo-tactile cueing, gaze following, trustworthiness

## **Abstract**

Joint attention refers to the coordinated attention between social partners to an object of shared interest, usually involving shared gaze toward the object. In the laboratory, however, joint attention is often investigated using computerized gaze cueing tasks that do not allow shared gaze. Instead, these computerized tasks require the participant to maintain fixation on the virtual partner's face, while the partner gazes to the left or right. Here we designed a modified gaze cueing task that better simulates a natural joint attention episode by allowing shared gaze, while still maintaining tight experimental control. In our computerized task the participant's gaze and the gaze of a virtual partner were manipulated independently, resulting in shared or unshared gaze. Following each gaze shift of the virtual partner a touch stimulus was delivered on one of the cheeks of the participant. We analyzed behavioral and neural (electro-encephalography) responses to the touch. Faster reaction-times and stronger lateralization of alpha power were observed when the touched cheek was in a jointly attended hemispace compared with a singly attended or unattended hemispace. Importantly, these effects were unique to joint attention and could not be explained as the additive effects of own gaze and gaze cue direction. Underlining its social nature, we found that the behavioral effect was absent when we repeated our experiment with nonsocial cues (arrows) instead of gaze cues. Furthermore, when we compared trustworthy with untrustworthy virtual partners (trustworthiness judgements based on facial appearance) we found the effect only for trustworthy and not for untrustworthy virtual partners. We conclude that joint attention based on shared gaze influences attentional orienting such that cross-modal sensory processing at the jointly attended location is facilitated, particularly when the partner is trustworthy. This indicates that social interactions and trustworthiness judgements affect cortical and behavioral responses to sensory information.

## Introduction

Joint attention refers to a situation in which two individuals have a common focus of attention on the same object (Reddy, 2003; Schilbach et al, 2013). The importance of joint attention as a social skill and a prerequisite for typical human social cognition is well-established. Infants engage in joint attention approximately by the end of the first year of life and their ability to do so is predictive for further development of linguistic and socio-emotional skills during childhood (Carpenter et al, 1998; Reddy, 2003). Furthermore, joint attention is thought to be associated with “Theory of Mind”, the ability to infer what someone else is thinking (Baron-Cohen, 1994; Schilbach et al, 2013), and reduced establishment of joint attention is an early predictor of autism spectrum disorders (Baron-Cohen, 1994; Dawson et al, 2004).

It is inherently difficult to investigate social phenomena like joint attention in an ecologically valid setting and still maintain tight experimental control. Therefore, many investigations in the laboratory have focused on a pre-stage of joint attention known as ‘gaze following’, which is a shift of spatial attention in the direction of someone else’s gaze. Using so-called ‘gaze cueing tasks’, it has been shown that gaze following occurs involuntarily, even when the observed gaze direction is detrimental to task performance (Emery, 2000; Frischen et al, 2007). In a typical gaze cueing task the image of a face that looks either to the left or right is presented on a computer screen and participants respond to a subsequent target stimulus appearing left or right of the face. Reaction times are faster when the face gazes in the direction of the target than when the face gazes in the opposite direction.

In a typical gaze cue task participants are not allowed to gaze at the same location as a virtual partner, as strict fixation on the virtual partner’s face is required. Thereby, there is no shared gaze between the participant and the virtual partner. Even when we consider joint attention in its most basic form, i.e., as sharing gaze on a common object without requiring knowledge or awareness of the other’s direction of attention (Kingstone et al., 2000; Mundy & Jarrold, 2010; cf. Tomasello & Hamann, 2012), this lack of shared gaze poses a cardinal difference between joint attention and the situation tested in a gaze cueing task. It therefore remains unclear whether joint attention, here narrowly defined as ‘sharing gaze’, elicits involuntary orienting by itself, or whether the involuntary part of joint

attention is restricted to the pre-stage of mere gaze following. We know that a participant's own gaze direction also elicits involuntary orienting (Bisley, 2011; Itti & Koch, 2001; Shephard et al, 1986/2009). Our question is whether the influences of own gaze and partner's gaze (gaze cueing) operate independently, or whether they interact, in which case it would make a difference whether own gaze and the partner's gaze are directed to the same location or not.

To answer this question we designed a modified gaze cueing task in which both the participant's gaze and the partner's gaze were manipulated independently and we tested whether their influences interact. If they do not, then own gaze and partner's gaze elicit orienting by themselves, irrespective of whether they are shared or not. However, if they do interact, then neural processes dedicated to own gaze and partner's gaze are integrated, providing a ground for an elemental form of joint attention based on shared gaze and a basis for the development of reciprocal social interactions (Caruana et al, 2017; Mundy & Jarrold, 2010; Pfeiffer et al, 2013; Schilbach et al, 2013). Hypothesizing that they interact, we anticipate that shared gaze elicits additional speeding of reaction times, beyond an extent that can be explained by the added effects of own gaze and partner's gaze.

We instructed participants to fixate on a stimulus left or right of a virtual partner's face and thus view the face peripherally, in the corner of their eyes. Following a gaze shift of the virtual partner a touch stimulus was delivered to one of the cheeks of the participant. Participants responded to the touch by pressing a button and reaction times and electro-encephalography (EEG) were recorded. The aim of measuring neural activity is to assess whether behavioral influences of joint attention are reflected in – and may be attributed to – modulations in sensory processing. We expected sensory processing of the touch stimulus to be reflected in modulations at fronto-central sites over the face-representation in sensory cortex in the theta and alpha frequency band (Genna et al, 2017; Hu et al, 2013). The alpha band is particularly interesting, as it has been attributed to attention and extended processing of touch (Genna et al, 2017; Klimesch, 2012).

We used a touch rather than a visual target to circumvent the misalignment of the spatiotopic field with the participant's retinotopic field (due to their averted gaze). Moreover, we considered a touch feasible and ecologically valid, considering existing knowledge about visuo-tactile cueing (Soto-

Faraco et al, 2005; Spence et al, 2004) and visuo-tactile integration near the face (Brozzoli et al, 2012; Graziano & Cooke, 2006).

To assess the social nature of possible effects of shared gaze, we repeated our experiment with a nonsocial stimulus, i.e., an arrow (Frischen et al, 2007), hypothesizing that processes related to joint attention are not activated when own gaze direction is shared with an arrow's pointing direction.

Arrows are often used as a control stimulus for gaze cues, because they are omnipresent in modern society, they do not contain eye-like features (minimizing the risk of pareidolia; Ichikawa et al., 2011; Takahashi & Watanabe, 2013) and they orient attention in a comparable way (Kuhn & Kingstone, 2009; Ristic et al., 2002; Tipples, 2002).

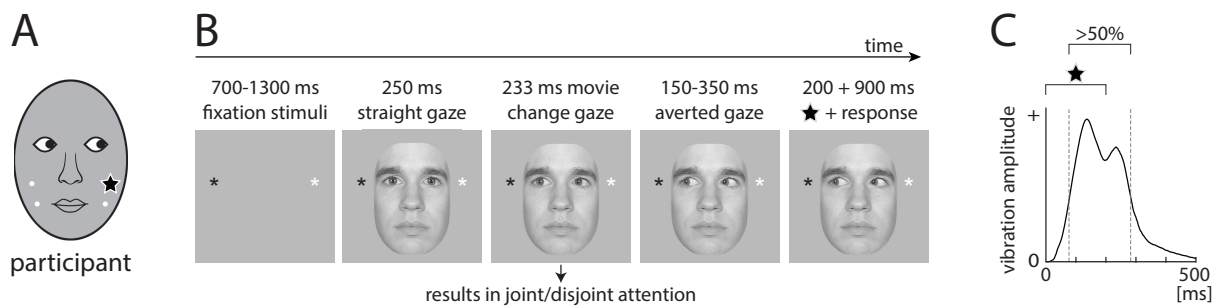
Additionally, we were interested to see if effects of shared gaze are influenced by the perceived trustworthiness of the virtual partner (trustworthiness judgements based on first impression).

Trustworthiness inferences are made within a split second (Todorov, 2008; Willis & Todorov, 2006) and could, thereby, potentially influence fast social processes like the effects of shared gaze or gaze following tested here. Consistent with earlier work showing that partners who tend to look away from a target stimulus are rated more untrustworthy (Bayliss & Tipper, 2006; Manssuer et al., 2015), we anticipated that processes related to joint attention are attenuated when a partner is untrustworthy.

## Materials and Methods

### Participants

Experiments were carried out in accordance with the Declaration of Helsinki (World Medical Association 2000) and approved by the ethical committee of the Faculty of Social and Behavioral Sciences at Utrecht University. Participants gave written informed consent prior to participation, were right-handed, had normal or corrected to normal vision, normal touch sensation and no psychiatric or neurological history. A priori power analysis indicated a recommended sample size of 24 participants (estimated effect-size: moderate to small,  $f = 0.25$ ; power = 80%;  $\alpha = 0.05$ ). We measured more participants in anticipation of technical failures or auditory transfer of the vibrotactile stimulus (i.e., using audition rather than touch to determine the touched location). All inclusion/exclusion criteria used in this study were established prior to data collection and we report all manipulations and all measures in the study. The data analysis included 24 participants (12 female, 12 male; mean age 24.1 years  $\pm 0.9$  SEM). Four additional participants were excluded because of auditory transfer of the vibrotactile stimulus, as assessed by a questionnaire (more information in 'Touch stimuli'-paragraph below; including these participants yielded similar behavioral results). Three additional participants were excluded because of technical failures. The study procedures and analyses were not pre-registered in a time-stamped, institutional registry prior to the research being conducted. The ethical committee of the Faculty of Social and Behavioral Sciences at Utrecht University does not grant permission for anonymized public archiving of the data, but data will be shared upon request with all requestors (requestors do not need to meet certain criteria; all participants gave written informed consent for data sharing). The data are stored at the Faculty of Social and Behavioural Sciences of Utrecht University and can be obtained by contacting H.C. Dijkerman. Stimuli, questionnaires and analysis code are publicly archived at the Open Science Framework (<https://osf.io/4n9yx>; DOI 10.17605/OSF.IO/4N9YX).



**Figure 1:**

**Paradigm.**

A) Schematic illustration of the averted gaze of the participant and the 4 possible touch locations on the cheek (white dots). Per trial only 1 of the touch locations vibrated (illustrated as black star). The participant indicated whether the touch was high or low on the cheek, irrespective of which cheek. Gaze direction of the participant was fixed across trials to eliminate oculomotor effects (but alternated between experimental blocks to avoid tiring of the eye muscles).

B) Fixation stimuli were presented throughout the trial. A variable delay after trial onset the virtual partner appeared with straight gaze. Then gaze direction of the virtual partner gradually changed to one of the fixation stimuli during a short movie. As a result, the participant and the virtual partner then fixated the same fixation stimulus (joint attention) or opposite fixation stimuli (disjoint attention). The tactile stimulator was activated a variable delay after that (illustrated by black star). The participants reported the vertical location of the touch (up/down) with a button press.

C) Vibration amplitude over time relative to onset of the stimulator. The stimulator was activated for 200 ms (illustrated by black star). After activation the rotating mass inside the stimulator needed some time to accelerate and produce a perceivable vibration. Then, when activation has stopped, the rotating mass needed some time to decelerate and thereby stop the vibration. The vibration was perceivable approximately 80-280 ms after onset of the stimulator, when it was >50% of its maximal amplitude (see Methods). The delay of approximately 80 ms between activation of the stimulator and delivery of the vibration does not differ between conditions, but should be taken into account when interpreting absolute reaction times and neural activity time-locked to onset of the tactile stimulator.

**Procedure and task**

Participants were seated in an electronically and acoustically shielded room, wearing a head-cap for electro-encephalography (EEG), using a chin-rest and looking at a computer screen (refresh-rate: 60 Hz; size: 302 x 378 mm). We placed the computer screen relatively close to the face of the participant (viewing distance was 50 cm), because visuo-tactile integration around the face is stronger inside

than outside peri-personal space (Brozzoli et al, 2012; Iachini et al, 2014). In addition, we ensured that this viewing distance was not further away than the inter-personal distance that our participants prefer during a social interaction. Specifically, preferred inter-personal distance when asking a stranger for directions was measured after the experiment for each participant and averaged to 73.4 cm  $\pm$ 4.2 SEM when they interacted with the experimenter and 76.7 cm  $\pm$ 4.9 SEM when they interacted with a life-sized poster of an unknown standing individual.

Participants completed 576 experimental trials (48 of which were pseudo-catch trials, as described below) and 48 catch trials, divided over 4 experimental sessions that were separated by short breaks. Four participants completed extra trials to compensate for data loss due to technical problems (more information about data loss below in 'Behavioral analysis'-section). Fixation stimuli (black and white asterisk, each 1.0°) were placed 11.9° apart (center to center) at eye-height of the to-be-presented virtual partner and alternated their position between sessions (Figure 1B). A session was subdivided in 6 blocks of 24 trials, each lasting 68 seconds. At the start of a block an instruction in text indicated which of the two fixation stimuli should be fixated. The instructed fixation position alternated between blocks to avoid tiring of the eye muscles and remained fixed within blocks to eliminate oculomotor effects (Schutz et al, 2011; Wardak et al, 2011). The first trial of each block was excluded from the analysis, because it was preceded by eye movements (related to reading the instruction in text and initiating fixation of the instructed asterisk). Details of the face stimuli and the touch stimulus are described in the following paragraphs. Importantly, the face stimuli were presented in the center of the screen (between the fixation stimuli) and there were 4 locations on the cheeks of the participant where the touch stimulus could be delivered: left up, left down, right up and right down (Figure 1).

Participants used their right thumb and index finger to operate two buttons on a button box. The button operated by the thumb was positioned to the lower left of the button operated by the index finger. On experimental trials, participants indicated perceived vertical location of the touch (thumb= low; index finger= high). The horizontal location (i.e., which cheek) was task-irrelevant to eliminate cueing effects on responding. On 48 catch trials the asterisk that the participant was fixating turned red and the participant was instructed to indicate gaze direction of the virtual partner (thumb= left; index finger= right). Correct perception (91.6% correct responses) of the gaze of the peripherally-



viewed virtual partner was validated (in line with Canadas & Lupianez, 2012; here we analyzed catch trials for which the experimental task would have required a different response than the catch task). Additionally, catch trials prevented participants from ignoring the visual modality. One participant mistakenly reported horizontal touch location instead of the partner's gaze direction on catch trials. This participant was included in the analysis, because the change in color of the asterisk that indicated a catch trial was correctly used by this participant to switch to an alternative task, indicating that the visual modality was not ignored. On 48 pseudo-catch trials the non-fixated asterisk colored red. Confirming that participants followed fixation instructions, they treated these trials as experimental trials (94.3% correct responses).

To assess whether effects of joint attention were associated with certain personality traits participants completed 3 questionnaires after they completed the experiment, namely the Autism spectrum Quotient (AQ; Baron-Cohen et al., 2001; Hoekstra et al., 2008), the State and Trait Anxiety Inventory (STAI; Spielberger et al., 1983) and the Behavioral Inhibition and Behavioral Activation Scales (BIS/BAS; Carver & White, 1994). Obtained scores were not significantly associated with effects of joint attention reported below, possibly due to very small variability in the scores of our non-clinical participant group.

### Face stimuli

A virtual partner appeared 700-1300 ms after trial onset, in the center of the screen, displaying straight gaze and subtending 9.0° horizontally. After 250 ms its gaze direction gradually changed during a short movie to either the left or the right (randomly) and was fully averted 233 ms later. A cheek stimulator was activated 150-350 ms later and the trial ended 1100 ms after that (Figure 1). The 9-frame movies of the gaze change were used in previous publications (e.g. de Jong et al., 2008) and contain face stimuli of 9 different actors with neutral expression taken from the validated stimulus set of MacArthur Foundation Research Network on Early Experience and Brain Development. Development of the MacBrain Face Stimulus Set was overseen by Nim Tottenham and supported by the John D. and Catherine T. (contact Nim Tottenham at [tott0006@tc.umn.edu](mailto:tott0006@tc.umn.edu) for more information concerning the stimulus set). Using Photoshop (Adobe) straight and averted eyes were edited, images were converted to grayscale and faces were matched on size, shape, contrast and luminance

(mean luminance equaled the mid-gray background luminance). A gamma correction was applied using Matlab (MathWorks Inc.). The movies were generated by morphing straight-gaze images to averted-gaze images of the same actor using Smartmorph (Meesoft).

### Touch stimuli

Stimulators on the cheeks looked like small button batteries, were attached using adhesive tape and produced a touch like the vibration of a mobile phone (eccentric rotating mass in cylindrical casing, 8 mm diameter, 3.4 mm thick, Precision Microdrives, Model: 308–100). Stimulators were placed high on the left cheek (below zygomatic bone), low on the left cheek (above mandible bone) and at equivalent positions on the right cheek (Figure 1A). Per trial 1 stimulator was activated (for 200 ms). Note that after activation of the stimulator the rotating mass inside the stimulator needed some time to accelerate and produce a perceivable vibration. Also, when activation of the stimulator has stopped the rotating mass needed some time to decelerate and thereby stop the vibration. The time-course of this acceleration and deceleration is provided in Figure 1C, showing that vibration amplitude was perceivable approximately 80-280 ms after stimulator onset (we here estimated an amplitude >50% of the maximum amplitude to be perceivable, based on information provided by Precision Microdrives). In the 80-280 ms interval the vibration frequency was ~100 Hz, as measured with a Bosh Sensortec BST-BMI160 inertial measuring unit. The delay of approximately 80 ms between activation of the stimulator and delivery of the vibration does not differ between conditions, but should be taken into account when interpreting absolute reaction times and neural activity time-locked to onset of the tactile stimulator.

The stimulators produced negligible sound, but facial bones/tissue or legs of eyeglasses may transmit the vibration to the ear. To minimize auditory transfer participants listened to white noise sound through ear-phones. After the experiment participants indicated whether they used audition or touch to determine the horizontal and vertical location of the vibration on a visual analogue scale that was labeled 'touch only' on the left end and 'audition only' on the right end. For analysis the positions marked on the scale were converted to a numerical scale ranging from 0 for audition only to +100 for touch only. As mentioned earlier, we excluded 4 participants who mainly used audition to determine the horizontal and/or vertical location (i.e., scores  $\leq 50\%$ ; more extreme cutoffs of 75% and 0%

yielded equivalent results, data not shown). Participants also rated the pleasantness of the tactile stimulus on a visual analogue scale, assuring it was not aversive ('very unpleasant'-label on the left end, 'very pleasant'-label on the right end; judgements converted to numerical scores between -100 and +100 for very unpleasant and very pleasant, respectively; average score:  $+22.7 \pm 6.8$  SEM).

### Trustworthiness

Before and after the experiment participants rated their first impression of the trustworthiness of each virtual partner as based on facial appearance, using a visual analogue scale (labeled 'untrustworthy' on the left end and 'trustworthy' on the right end; partners were presented with straight gaze). For analysis the judgements were converted to a numerical scale ranging from -100 for untrustworthy to +100 for trustworthy (as in Manssuer et al., 2015). One participant only completed the rating before the experiment. This participant was included in the analysis.

### Experiment with nonsocial cues

To assess the social nature of observed effects, we repeated our experiment showing arrows instead of virtual partners. Durations of the initial non-directional image (arrow-shaft only; subtending  $4.8^\circ$ , equaling center-to-center distance between the actors' pupils), the movie and the directional image (arrow with head/vanes; subtending  $6.4^\circ$ , equaling distance between outer corners of the actors' eyes; illustration in Figure 2D) were the same as in the main experiment, meaning the arrow head and vanes appeared gradually during the movie. Each of 9 images of an arrow was matched to one of the 9 actors (vertical position and size matched to actors' eyes). The arrows ranged 0-100% in luminance with mean luminance equaling background luminance. On catch trials participants were instructed to report arrow direction.

This experiment was performed by another group of participants and did not include EEG recordings. Analysis included 19 participants (14 female, 5 male; age  $29.1$  years  $\pm 2.9$  SEM). Three participants mistakenly reported horizontal touch location on catch trials. These participants were included in the analysis, because they correctly used the visual signal that indicated a catch trial to switch to an alternative task, showing that they did not ignore the visual modality. Six additional participants were excluded: 3 participants mainly used audition to determine touch location (including these participants

yielded equivalent behavioral results), 1 participant performed the experimental task on catch trials and 2 participants performed the catch task on pseudo-catch trials.

#### Electro-encephalography recording and preprocessing

Electro-encephalography (EEG) was sampled at 2048 Hz and recorded with 64 silver/silver-chloride (Ag/AgCl) flat type active electrodes (Active Two system, Biosemi) positioned at standard locations on an elastic cap (Quickcap, Neuromedical supplies of Neurosoft inc.) and referenced to an additional active electrode (Common Mode Sense) during recording. Two electrodes in the cap provided an active ground. Horizontal and vertical electro-oculograms were measured. Offline preprocessing was performed using BrainVision Analyzer 2 (Brainproducts GmbH). A 0.1-200 Hz band-pass filter was applied and signals were referenced to the average of all scalp electrodes. Artifacts due to eye movements were removed (algorithm by Gratton et al. 1983) and epochs time-locked to the onset of the tactile stimulator were extracted from the continuous data. Artifact rejection was performed on individual channels and included removal of epochs with excessive amplitude changes ( $>100 \mu\text{V}$  within 200 ms). We also checked for recording failure ( $<1 \mu\text{V}$  differences within 200 ms and steps per sample point  $>50 \mu\text{V}$ ), but these seldom occurred.

#### Further analysis of electro-encephalography

Time-frequency transformation was performed over a frequency range of 5-13 Hz (5-60 Hz in Supplementary Materials, Figure S2) using in-house-developed Matlab code (Mathworks Inc.) and the Matlab toolbox EEGLab (Delorme & Makeig, 2004). We used a 3-18 cycle Morlet wavelet tapered with a Hanning window, taking frequency steps of 1 Hz and time steps of 10 ms. To isolate induced activity, the event-related potential per participant, per condition, was subtracted from the single trial data prior to the transformation. Normalization was performed per epoch, per frequency (normalized power = power / mean power). Mean normalized power over a 100-ms pre-stimulus baseline interval was subtracted. We analyzed a fronto-central electrode site associated with tactile processing (site FC5/FC6; Figure 4A; Genna et al, 2017; Hu et al, 2013). We analyzed contralateral and ipsilateral power as well as power lateralization according to: lateralization = (contralateral-ipsilateral) / (contralateral+ipsilateral).

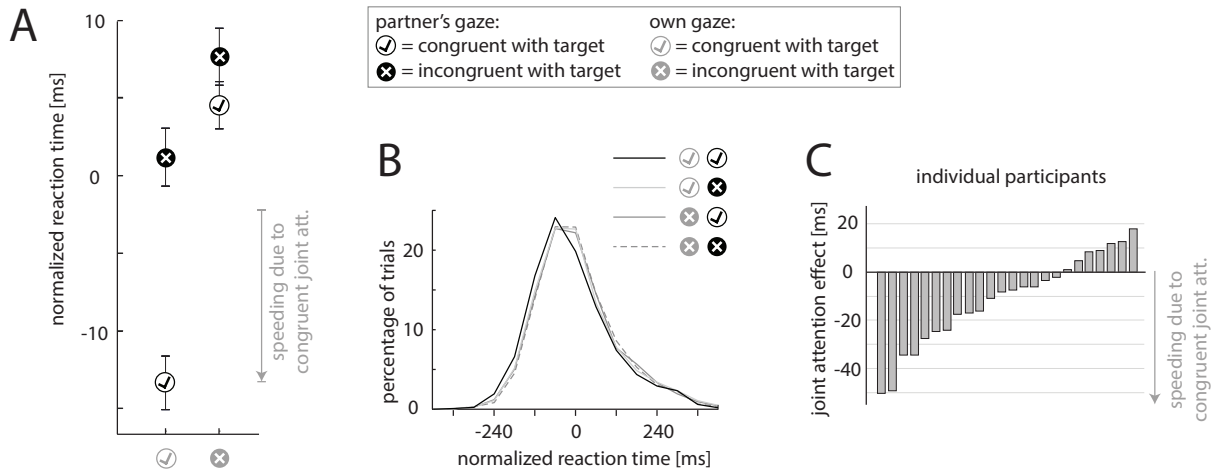
### Behavioral analysis

Experimental trials in which no button press was recorded were excluded from the analysis (7.6% of trials; data loss was mostly due to technical problems with the button box or the tactile stimulator). Using Matlab (Mathworks) the mean reaction time per participant, per touch location was subtracted from the data and outliers were removed (2.0% of experimental trials). Outliers included reaction times <100 ms and >1100 ms (before normalization) and reaction times >3 standard deviations from the participant's mean reaction time (following e.g. Hietanen et al., 2008; Manssuer et al., 2015). Trials with an incorrect response (3.4%) and trials for which no EEG was recorded (6.6%; due to technical failures) were also excluded. These exclusion criteria were also applied to the EEG analysis. The final number of analyzed trials was 459.6 on average  $\pm 10.6$  SEM (114.9 per condition).

### Statistical analysis

Unless indicated otherwise, we performed a repeated-measures analysis of variance (ANOVA) over participants, with own gaze and partner's gaze as within factors. The within factors had two levels: congruent and incongruent with target laterality.

## Results



**Figure 2:**

The influence of joint attention on reaction times.

A) Joint attention toward the touch hemispace resulted in exceptionally fast reaction times, i.e., faster than could be explained by the added effects of the participant's gaze and the virtual partner's gaze per se (referred to as *own* and *partner's* gaze, respectively). Check marks and X-marks indicate gaze to touched and untouched hemispace, respectively. Grey and black symbols refer to own and partner's gaze, respectively. Error bars indicate  $\pm$ SEM.

B) Reaction-time distributions per condition shown for illustrational purposes. The distribution for jointly attended (black line) compared with singly attended (solid light grey and dark grey lines) and unattended (dashed grey line) targets was shifted in time, but not markedly different in shape.

C) The amount of additional speeding due to joint attention, i.e., shared gaze, for the individual participants (difference in reaction time between congruent joint attention and the average of the two conditions with disjoint attention, minus the difference between this average and incongruent joint attention).

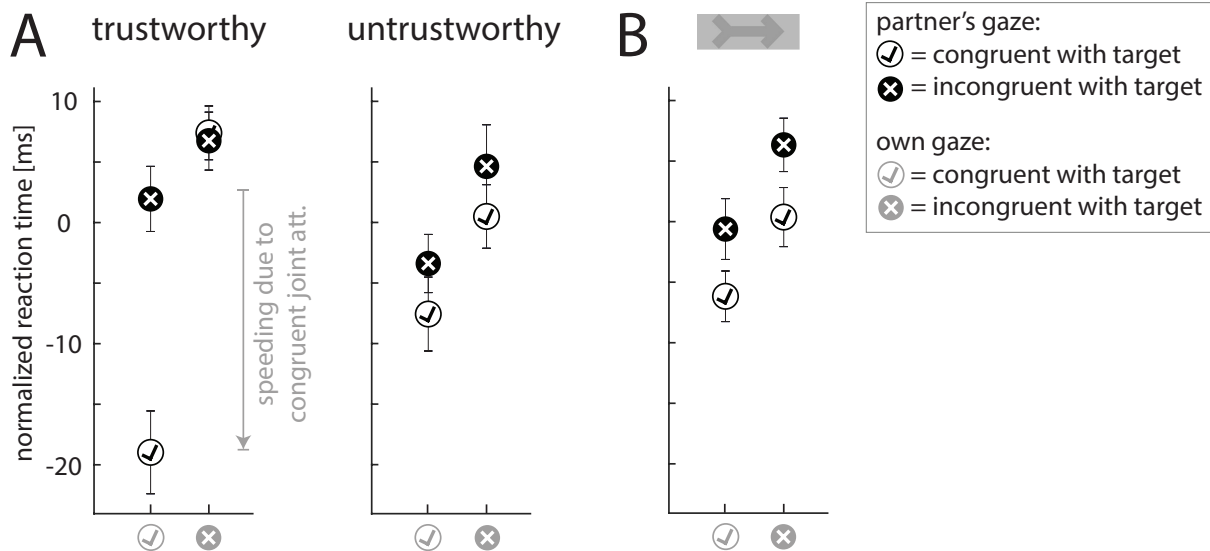
### Reaction times

Mean reaction time was 599 ms  $\pm$ 17 SEM (this reaction time includes approximately 80 ms that is needed for the tactile stimulator to accelerate and produce a perceivable vibration, see Methods and Figure 1C). Reaction times were faster when the cheek in the gazed-at hemispace was touched than when the other cheek was touched, both regarding the participant's gaze (own gaze:  $F_{1,23}= 29.2$ ,  $p= 0.000017$ ) and the partner's gaze (partner's gaze:  $F_{1,23}= 23.1$ ,  $p= 0.00008$ ; Figure 2; mean normalized reaction times: congruent joint attention:  $-13.4 \pm 1.7$  SEM; only own gaze congruent:  $1.2 \pm 1.9$  SEM;

only partner's gaze congruent:  $4.5 \pm 1.5$  SEM; incongruent joint attention:  $7.7 \pm 1.8$  SEM). Importantly, there was an interaction between these main effects (own gaze x partner's gaze:  $F_{1,23} = 8.8$ ,  $p < 0.007$ ), which revealed that, when the touched cheek was in a jointly attended hemispace, there was a speeding of reaction times that exceeded the added effects of own gaze and partner's gaze per se (i.e., joint attention to touched hemispace differed from all other conditions: all  $t_{(23)} > 5.2$ , all  $p < 0.00003$ , while the other conditions did not differ from each other: all  $t_{(23)} \leq 2.0$ , all  $p \geq 0.06$ ; Figure 2). This additional speeding was more robust for long compared with short delays between partner's gaze shift and touch ( $>250$ -ms compared with  $\leq 250$ -ms between final face and touch onset; delay as additional within factor: delay x own gaze x partner's gaze:  $F_{1,23} = 4.6$ ,  $p = 0.04$ ; own gaze x partner's gaze, long delays:  $F_{1,23} = 10.2$ ,  $p = 0.004$ ; short delays:  $F_{1,23} = 0.1$ ,  $p = 0.7$ ). The main effects of own gaze and partner's gaze were not influenced by the duration of this delay.

#### Trustworthiness of the virtual partner

Trustworthiness ratings were similar over time (repeated-measures ANOVA over participants with time and actor as within factors; time:  $F_{1,20} \approx 0.0$ ,  $p \approx 1.0$ , actor:  $F_{8,160} = 5.6$ ,  $p = 0.00003$ , time x actor:  $F_{8,160} = 1.0$ ,  $p = 0.4$ ), so we averaged ratings obtained before and after the experiment. Faces with a positive and negative averaged rating were labeled 'trustworthy' (5 faces; score:  $+13.6 \pm 2.9$  SEM) and 'untrustworthy' (4 faces; score:  $-5.6 \pm 4.6$  SEM), respectively. The difference between these sets of faces was highly significant ( $t_{(23)} = 4.8$ ,  $p = 0.00007$ ), reflecting the consistency of the ratings across participants. Trustworthy faces differed from untrustworthy faces with respect to the interaction between own gaze and partner's gaze (trustworthiness as additional within factor: trustworthiness x own gaze x partner's gaze:  $F_{1,23} = 7.3$ ,  $p = 0.01$ ; repeating the analysis with 4 faces in each group, leaving out the trustworthy face with the lowest score, yielded equivalent results: trustworthiness x own gaze x partner's gaze:  $F_{1,23} = 5.9$ ,  $p = 0.02$ ). There were no other effects of trustworthiness. The interaction between own gaze and partner's gaze was significant for trustworthy but not untrustworthy faces (Figure 3A; own gaze x partner's gaze, trustworthy faces:  $F_{1,23} = 18.7$ ,  $p = 0.0002$ ; untrustworthy faces:  $F_{1,23} \approx 0.0$ ,  $p \approx 1.0$ ). Partial analyses showed that trustworthy faces resulted in faster responses than untrustworthy faces when the touch was in a jointly attended hemispace, with no or less robust influences of trustworthiness in the other conditions.



**Figure 3:**

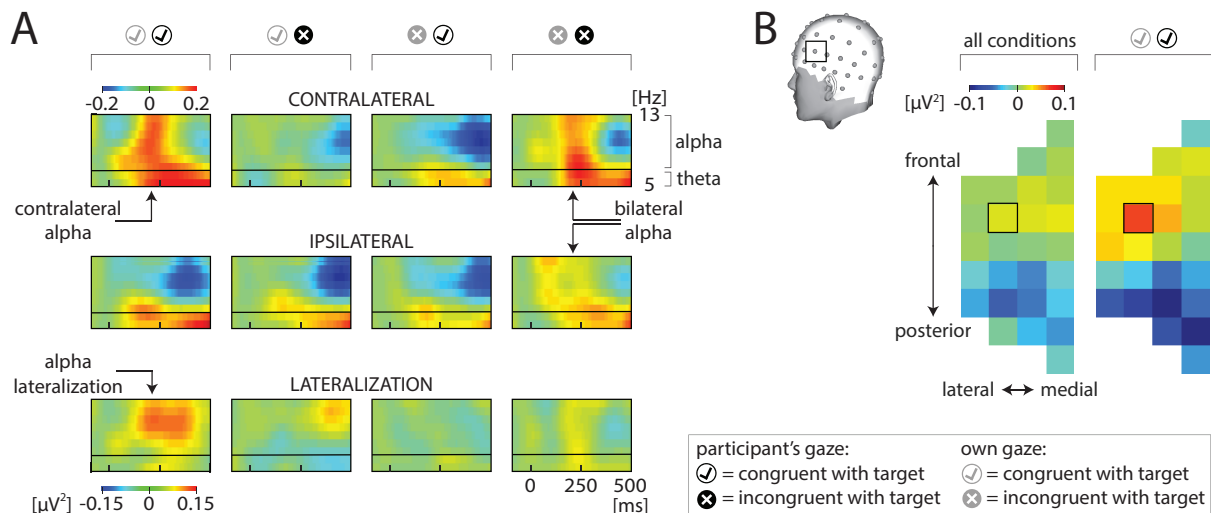
A) The same data as in Figure 2A, here plotted separately for trustworthy and untrustworthy virtual partners. The additional speeding of responses with jointly attended touch was present for trustworthy but not untrustworthy virtual partners. Reaction times for untrustworthy virtual partners showed a modulation by own gaze and partner's gaze per se, but these effects did not interact, thus revealing no additional modulation by joint attention. Conventions as in Figure 2A.

B) When we used arrows as a spatial cue instead of gaze shifts by a virtual partner there was no additional speeding of responses when the participant's gaze was shared with the arrows pointing direction toward the touch hemisphere. Conventions as in panel A.

### Social vs. nonsocial cues

Mean reaction time was 587 ms  $\pm$ 20 SEM (this reaction time includes approximately 80 ms that is needed for the tactile stimulator to accelerate and produce a perceivable vibration, see Methods and Figure 1C). The arrows were effective spatial cues ( $F_{1,19} = 5.4$ ,  $p = 0.03$ ), but contrary to gaze shifts, the influence of arrows did not interact with the influence of own gaze (Figure 3B; own gaze x arrow:  $F_{1,18} = 0.0$ ,  $p = 0.9$ ), although this difference between arrows and gaze cues was only marginally significant in an overall analysis (cue type as additional between factor: cue type x own gaze x cue direction:  $F_{1,41} = 3.2$ ,  $p = 0.08$ ).





**Figure 4:**

The influence of joint attention on touch processing

A) Modulation in theta and alpha power at the site-of-interest. Top: contralateral increase in alpha power was observed for joint but not disjunct attention, possibly reflecting enhanced processing of joint compared with disjunct attention (as seen prior to the touch, Supplementary Figure S1). Middle: ipsilateral alpha power increase was present only for joint attention toward the untouched hemispace, possibly reflecting allocation of spatial attention to this hemispace. Bottom: lateralization of alpha power ( (contralateral-ipsilateral) / (contralateral+ipsilateral) ) was observed for jointly attended, but not for singly attended or unattended touch, in line with the idea of facilitated tactile processing for jointly attended touch locations. Conventions as in Figure 2. Small head in panel B shows location of fronto-central site-of-interest.

B) Map of alpha lateralization across the scalp (frontal sites on top, lateral sites on the left) averaged for all conditions (left) and for jointly attended touch only (right), showing alpha lateralization for jointly attended touch was maximal at the site-of-interest (indicated by black square). Small head: location of fronto-central site-of-interest.

### Neural processing of touch

Two participants were excluded from these analyses because of recording artifacts at the fronto-central site-of-interest. There was a bilateral increase in theta power for all conditions. For joint but not disjunct attention contralateral alpha power was also increased (Figure 4A; 7-14 Hz, 175-250 ms time-interval, own gaze x partner's gaze:  $F_{1,21} = 15.8$ ,  $p = 0.0007$ ). This difference in contralateral alpha power was not present before onset of the target (Supplementary Materials, Figure S1), validating its post-target onset. When the touched cheek was in a jointly *unattended* hemispace there was a

concomitant ipsilateral increase in alpha power. Alpha power was thus lateralized to the contralateral side only when the touched cheek was in a jointly *attended* hemispace (Figure 4A; 175-375 ms time-interval:  $t_{(21)} = 3.9$ ,  $p = 0.0007$ ) and not in the other conditions (all  $t_{(21)} < 1.2$ , all  $p \geq 0.3$ ; own gaze x partner's gaze:  $F_{1,21} = 12.6$ ,  $p = 0.002$ ).

Across conditions more alpha lateralization was associated with faster reaction times (slope of regression line:  $t_{(86)} = -2.4$ ,  $p = 0.02$ ), while there was no relation with contralateral or ipsilateral alpha power. The above-reported alpha lateralization for touch in a jointly attended hemispace was narrowly localized to the site-of-interest for tactile stimulation of the face, corroborating it reflected facilitated touch processing (Figure 4B). There was no pre-target alpha lateralization (Supplementary Materials, Figure S1).

An overall analysis yielded no significant influences of trustworthiness on alpha lateralization. In separate analyses for trustworthy and untrustworthy faces, however, we observed a pattern in line with the behavioral results for trustworthiness (Supplementary Materials, Figure S2).

## Discussion

Considering the important role of shared gaze in joint attention, we designed a modified gaze cue task that allowed shared gaze between the participant and the virtual partner and investigated interactions between own gaze and partner's gaze direction. We found speeded responses to touch in a jointly attended compared with a singly attended or unattended hemispace. Importantly, this speeding reflected an effect unique to joint attention, as it exceeded the added effects of own gaze and partner's gaze per se (Figure 2). To our knowledge no such automatic effect of joint attention based on shared gaze has yet been reported. The additional speeding was significant for long but not short delays between the partner's gaze shift and the touch, whereas the effects of own gaze and partner's gaze per se were not influenced by this delay. This suggests joint attention involves more extended neural processing than mere gaze following.

Underlining its social nature, the additional speeding was not observed when we presented arrows instead of gaze cues (Figure 3B; interpretation in line with: Becchio et al., 2008; Greene et al., 2009; Ristic et al., 2002), indicating that sharing own gaze direction with the pointing direction of an arrow does not elicit additional attentional orienting. Our results are in line with the idea that merely observing someone else's gaze does not elicit socio-cognitive processes associated with reciprocal social interactions, while interacting through gaze shifts does (Caruana et al, 2017; Mundy & Jarrold, 2010; Schilbach et al, 2013; Shephard et al, 2009). Hence, they suggest that typical gaze cueing tasks are well-suited to investigate gaze following as a pre-stage of joint attention (Frishen et al., 2007), but are less suited to investigate joint attention as a social interaction, i.e., a more advanced social skill associated with the synergy between own gaze and partner's gaze direction (Caruana et al., 2017; Mundy & Jarrold, 2010; Pfeiffer et al., 2013; Schilbach et al., 2013).

### Neural processing of touch

Electro-encephalography (EEG) revealed that alpha power was lateralized to the contralateral site only when the touched cheek was in the jointly attended hemispace and not in the other conditions (Figure 4A), suggesting our behavioral results regarding response times were associated with enhanced sensory processing of the touch. Indeed, this alpha lateralization was found specifically at

the site-of-interest for tactile stimulation of the face, corroborating its association with tactile processing. It occurred roughly 175-375 ms after touch onset, a latency at which bottom-up activation of primary and secondary somatosensory cortex has already occurred and extended tactile processing is ongoing (Nguyen et al., 2004/2005; Suzuki et al., 2004). It could thus reflect an enhancement of extended rather than initial touch processing, in line with the modulation of reaction times by joint attention, which existed for long but not short delays between the partner's gaze shift and touch. Reaction times correlated with alpha lateralization but not with contralateral or ipsilateral alpha power, corroborating that more alpha lateralization was associated with faster responding.

Regarding contralateral alpha, we observed a power increase for joint but not disjoint attention (Figure 4A). This effect resembled previously reported contralateral alpha with nociceptive tactile stimulation, particularly when stimuli were attended (Hu et al., 2013; Ohara et al., 2004). It might reflect a general enhancement/prioritization of sensory processing when gaze direction is shared. In line with this interpretation, we found enhanced neural modulations prior to the touch that we attribute to increased hedonic value and/or priority of joint attention compared with disjoint attention (Supplementary Materials, Figure S1B). A concomitant ipsilateral alpha increase was found only when the touched cheek was in the jointly *unattended* hemispace (Figure 4A), possibly related to spatial attention being allocated to the untouched cheek.

#### Trustworthiness of the virtual partner

Even though all virtual partners were non-predictive and had a validated neutral expression, the modulation of reaction times by joint attention was present for virtual partners who were judged to be trustworthy but not for those who were judged to be untrustworthy (Figure 3A). Trustworthiness inferences are made reliably and automatically within a split second, even by young children (Cogsdill et al., 2014; Willis & Todorov, 2006), possibly reflecting an evolutionary advantage of detecting trustworthiness from facial appearance (Todorov, 2008; Willis & Todorov, 2006). Our results show that trustworthy partners elicited greater speeding of responses to jointly attended touch than untrustworthy partners, suggesting our first impression of someone influences the way we process and respond to jointly attended stimuli.

The above-reported alpha lateralization that was unique for jointly attended touch was not significantly modulated by trustworthiness, although there was a trend suggesting it was more robust for trustworthy than for untrustworthy faces (Supplementary Materials, Figure S2F).

We cannot exclude the possibility that the observed effects of trustworthiness reflect differences in visual characteristics between the trustworthy and untrustworthy faces rather than the trait trustworthiness. However, if so, it would be expected that gaze following per se is also affected by trustworthiness, an effect we did not observe, as only the interaction between own gaze and partner's gaze direction was influenced by trustworthiness. Further supporting the idea that our results relate to trait trustworthiness, the visual characteristics of the faces were well-matched and we found no systematic differences between the features of trustworthy and untrustworthy faces.

### Limitations

We here tested our modified gaze cue task for the first time, showing there is an interaction between own gaze and partner's gaze regarding behavioral as well as neural responses. Future investigations with larger sample sizes could extend our knowledge of individual differences. For example, they could investigate whether individual differences or gender differences in gaze following (Frishen et al., 2007) are correlated with differences in the interaction between own gaze and partner's gaze direction. Also, clinical populations with Autism Spectrum Disorders often perform normal in typical gaze cueing tasks, but may show abnormalities on our task that involves joint attention based on shared gaze (de Jong et al., 2008).

As we recorded neural activity from the scalp, we do not know precisely which brain regions the observed neural modulations originated from. Importantly, note that it is unlikely that our neural findings reflect motor preparation, because 1) sensorimotor activity associated with the hands is expected at sites located more medial, 2) response-related effects were averaged out when averaging across touch locations and 3) participants responded to the vertical location of the touch while gaze was manipulated in the horizontal dimension.

Given that one cannot see his/her own cheek, the present paradigm took advantage of the spatially unfocused nature of visuo-tactile cueing (visual cue and tactile target do not need to be at the exact same location; Graziano & Cooke, 2006; Spence et al., 2004; Wiese et al., 2014). The spatial precision of joint attention is, however, unknown. Would our results hold when the participant and the partner fixated two different stimuli on the same side, or when fixation stimuli were omitted and merely the fixation *direction* was instructed? Also, are our findings specific for the defensive zone around the face (Iachini et al., 2014; Brozzoli et al., 2012; Graziano & Cooke, 2006) or do they extend to other body parts (for example the hands; Soto-Faraco et al., 2005) or to further distances (away from the body and outside the range of normal social interactions)? These questions need further investigation. Another direction for future investigations could be to test whether results are the same for intentional rather than incidental joint attention, for responding to rather than initiating a joint attention episode and for real-life partners rather than virtual partners.

Although the arrow cues and gaze cues both gradually changed into a directional cue during a short movie, we cannot rule out that differences in motion, or other low-level visual features, contributed to the observed differences between gaze and arrow cues reported here. Notably, it has been shown that gaze cueing does not rely on motion (no difference between sudden onset gaze cue and implied motion gaze cue; Canadas et al., 2012; Kuhn & Kingstone, 2009).

### Conclusion

We found that effects of gaze following are modified by the gaze direction of the participant. Specifically, when the touched cheek was in a jointly attended hemispace reaction times were faster than could be explained by the added main effects of own gaze and the virtual partner's gaze direction. These behavioral effects were accompanied by neural modulations that indicated more elaborate tactile processing of jointly attended touch. Furthermore, the behavioral effects were absent when nonsocial cues (arrows) were used and when the virtual partner was considered untrustworthy at first impression. We conclude that joint attention influences the way we process and respond to stimuli at a jointly attended location, particularly when the partner is trustworthy. Our findings indicate that social interactions and trustworthiness judgements modify our behavioral and cortical responses to cross-modal sensory stimulation.

## Acknowledgements

This study was supported by a NWO VICI grant to HCD (453-10-003). We thank Chantal Kemner for providing excellent lab facilities and Joshua Snell for his assistance. We also thank Ivan Toni and Chantal Kemner for helpful suggestions regarding the data. The authors have no competing interests to declare.

## References

- Baron-Cohen S. 1994. How to Build a Baby That Can Read Minds - Cognitive Mechanisms in Mindreading. *Cahiers De Psychologie Cognitive-Current Psychology of Cognition*. 13(5):513-552.
- 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. *J Autism Dev Disord*. 31(1):5-17.
- Bayliss AP, Tipper SP. (2006). Predictive gaze cues and personality judgments: Should eye trust you? *Psychol Sci*, 17(6), 514-520.
- Becchio C, Bertone C, Castiello U. 2008. How the gaze of others influences object processing. *Trends Cogn Sci*. 12(7):254-258.
- Bisley, J. W. 2011. The neural basis of visual attention. *J Physiol*. 589(Pt 1):49-57.
- Brozzoli C, Makin TR, Cardinali L, Holmes NP, Farne A. 2012. Peripersonal Space: A Multisensory Interface for Body-Object Interactions. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. Boca Raton (FL).
- Canadas E, Lupianez J. 2012. Spatial interference between gaze direction and gaze location: a study on the eye contact effect. *Q J Exp Psychol (Hove)*, 65(8), 1586-1598.
- Carpenter M, Nagell K, Tomasello M. 1998. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr Soc Res Child Dev*. 63(4):1-143.
- Caruana N, McArthur G, Woolgar A, Brock J. 2017. Simulating social interactions for the experimental investigation of joint attention. *Neuroscience & Biobehavioral Reviews*. 74:115-125.

- Carver CS, White TL. 1994. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319-333.
- Cogsdill EJ, Todorov AT, Spelke ES, Banaji MR. 2014. Inferring character from faces: a developmental study. *Psychol Sci*. 25(5):1132-1139.
- Dawson G, Toth K, Abbott R, Osterling J, Munson J, Estes A, Liaw J. 2004. Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Dev Psychol*. 40(2):271-283.
- Delorme A, Makeig S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. 134(1):9-21.
- de Jong MC, van Engeland H, Kemner C. 2008. The attentional effect of gaze shifts is influenced by emotion and spatial frequency, but not in autism. *J Am Acad Child Adolesc Psychiatry*. 47:443-454.
- Emery NJ. 2000. The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*. 24(6):581-604.
- Frischen A, Bayliss AP, Tipper S. P. 2007. Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychol Bull*. 133(4):694-724.
- Genna C, Oddo CM, Fanciullacci C, Chisari C, Jorntell H, Artoni F, Micera S. 2017. Spatiotemporal Dynamics of the Cortical Responses Induced by a Prolonged Tactile Stimulation of the Human Fingertips. *Brain Topogr*. 30(4):473-485.
- Gratton G, Coles MG, Donchin E. 1983. A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol*. 55:468-484.
- Graziano MS, Cooke DF. 2006. Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia*. 44(13):2621-2635.
- Greene DJ, Mooshagian E, Kaplan JT, Zaidel E, Iacoboni M. 2009. The neural correlates of social attention: automatic orienting to social and nonsocial cues. *Psychol Res*, 73(4), 499-511
- Hietanen JK, Leppanen JM, Nummenmaa L, Astikainen P. 2008. Visuospatial attention shifts by gaze and arrow cues: an ERP study. *Brain Res*, 1215, 123-136.



- Hoekstra RA, Bartels M, Cath DC, Boomsma DI. 2008. Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. *J Autism Dev Disord*, 38(8), 1555-1566.
- Hu L, Peng W, Valentini E, Zhang Z, Hu Y. 2013. Functional features of nociceptive-induced suppression of alpha band electroencephalographic oscillations. *J Pain*. 14(1):89-99.
- Iachini T, Coello Y, Frassinetti F, Ruggiero G. 2014. Body space in social interactions: a comparison of reaching and comfort distance in immersive virtual reality. *PLoS One*. 9(11):e111511.
- Ichikawa H, Kanazawa S, Yamaguchi MK. 2011. Finding a face in a face-like object. *Perception*, 40(4), 500-502.
- Itti L, Koch C. 2001. Computational modelling of visual attention. *Nat Rev Neurosci*. 2(3):194-203.
- Kingstone A, Friesen CK, Gazzaniga MS. 2000. Reflexive joint attention depends on lateralized cortical connections. *Psychol Sci*, 11(2), 159-166.
- Klimesch W. 2012. Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn Sci*, 16(12), 606-617.
- Kuhn G, Kingstone A. 2009. Look away! Eyes and arrows engage oculomotor responses automatically. *Atten Percept Psychophys*, 71(2), 314-327.
- Manssuer LR, Roberts MV, Tipper SP. 2015. The late positive potential indexes a role for emotion during learning of trust from eye-gaze cues. *Social Neuroscience*, 10(6), 635-650.
- Mundy P, Jarrold W. 2010. Infant joint attention, neural networks and social cognition. *Neural Networks*. 23(8-9):985-997.
- Nguyen BT, Tran TD, Hoshiyama M, Inui K, Kakigi R. 2004. Face representation in the human primary somatosensory cortex. *Neurosci Res*. 50(2):227-232.
- Nguyen BT, Inui K, Hoshiyama M, Nakata H, Kakigi R. 2005. Face representation in the human secondary somatosensory cortex. *Clin Neurophysiol*. 116(6):1247-1253.
- Ohara S, Crone NE, Weiss N, Lenz FA. 2004. Attention to a painful cutaneous laser stimulus modulates electrocorticographic event-related desynchronization in humans. *Clin Neurophysiol*. 115(7):1641-1652.
- Pfeiffer UJ, Vogeley K, Schilbach L. 2013. From gaze cueing to dual eye-tracking: Novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience and Biobehavioral Reviews*. 37(10):2516-2528.

- Reddy V. 2003. On being the object of attention: implications for self-other consciousness. *Trends Cogn Sci.* 7(9):397-402.
- Ristic J, Friesen CK, Kingstone A. 2002. Are eyes special? It depends on how you look at it. *Psychon Bull Rev*, 9(3), 507-513.
- Schilbach L, Timmermans B, Reddy V, Costall A, Bente G, Schlicht T, Vogeley K. 2013. Toward a second-person neuroscience. *Behav Brain Sci.* 36(4):393-414.
- Schutz AC, Braun DI, Gegenfurtner KR. 2011. Eye movements and perception: a selective review. *J Vis.* 11(5):9.
- Shepherd M, Findlay JM, Hockey RJ. 1986. The relationship between eye movements and spatial attention. *Q J Exp Psychol A.* 38(3):475-491.
- Shepherd SV, Klein JT, Deaner RO, Platt M. L. 2009. Mirroring of attention by neurons in macaque parietal cortex. *Proc Natl Acad Sci U S A.* 106(23):9489-9494.
- Soto-Faraco S, Sinnott S, Alsius A, Kingstone A. 2005. Spatial orienting of tactile attention induced by social cues. *Psychon Bull Rev.* 12(6):1024-1031.
- Spence C, Pavani F, Maravita A, Holmes N. 2004. Multisensory contributions to the 3-D representation of visuotactile peripersonal space in humans: evidence from the crossmodal congruency task. *J Physiol Paris.* 98(1-3):171-189.
- Spielberger CD, Gorsuch RL, Lushene PR, Vagg PR, Jacobs GA. 1983. *Manual for the State-Trait Anxiety Inventory.* Consulting Psychologists Press.
- Suzuki T, Shibukawa Y, Kumai T, Shintani M. 2004. Face area representation of primary somatosensory cortex in humans identified by whole-head magnetoencephalography. *Jpn J Physiol.* 54(2):161-169.
- Takahashi K, Watanabe K. 2013. Gaze cueing by pareidolia faces. *Iperception*, 4(8), 490-492.
- Tipples J. 2002. Eye gaze is not unique: automatic orienting in response to uninformative arrows. *Psychon Bull Rev*, 9(2), 314-318.
- Todorov A. 2008. Evaluating faces on trustworthiness: an extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Ann N Y Acad Sci.* 1124:208-224.
- Tomasello M, Hamann K. 2012. Collaboration in young children. *Q J Exp Psychol (Hove)*, 65(1), 1-12.
- Wardak C, Olivier E, Duhamel JR. 2011. The relationship between spatial attention and saccades in the frontoparietal network of the monkey. *Eur J Neurosci.* 33(11):1973-1981.

Wiese E, Wykowska A, Muller HJ. 2014. What we observe is biased by what other people tell us: beliefs about the reliability of gaze behavior modulate attentional orienting to gaze cues. *PLoS One*, 9(4), e94529.

Willis J, Todorov A. 2006. First impressions: making up your mind after a 100-ms exposure to a face. *Psychol Sci*. 17(7):592-598.

## SUPPLEMENT

The influence of joint attention and partner trustworthiness on cross-modal sensory cueing

M.C. de Jong, H.C. Dijkerman

### Supplementary Introduction

Below we describe additional analyses of neural activity that are aimed to facilitate comparison between our findings and existing literature. In addition to the time-frequency analysis described in the main text that focused on low frequencies, we here provide an analysis over a broad range of frequencies (5-60 Hz) and an analysis of the event-related potential associated with target processing. We discuss the latter in relation with previously observed event-related potentials in gaze cueing paradigms (de Jong et al., 2008; Hietanen et al., 2008; Schuller & Rossion, 2004).

Furthermore, in this supplement we provide a detailed analysis of neural activity occurring during the pre-target time-interval. First and foremost, we assessed whether differences in alpha power and alpha lateralization were present in the pre-target time-interval and could have influenced the baseline-period used for analysis of post-target effects. We found no such pre-target modulations, validating the post-target onset of alpha modulations reported in the main text. Furthermore, we were interested to see whether the Late Positive Potential (LPP), which is believed to reflect prioritization in face processing (Langeslag et al., 2007; Manssuer et al., 2015; Schupp et al., 2004), is modified by joint attention. Finally, we tested whether the gaze cue elicited a larger low-frequency power decrease contralateral compared with ipsilateral to the observed gaze direction, as would be expected based on previous literature on visual cueing (Figure S1C; Haegens et al, 2011; Thut et al, 2006; van Ede et al, 2011).

## Supplementary Methods

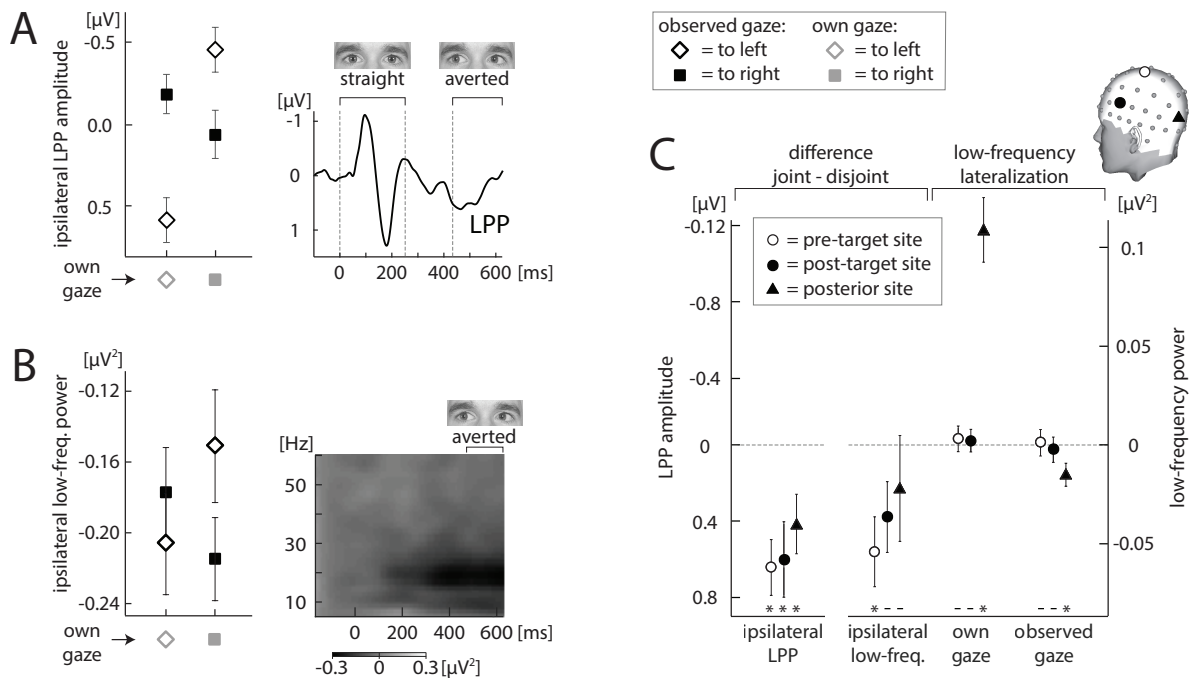
### Neural activity prior to the touch

We analyzed event-related potentials and the time-frequency decomposition of neural activity time-locked to the onset of the virtual partner. In the analysis of event-related potentials we applied a 0.1-30 Hz band-pass filter and a 50 Hz Notch filter (instead of a 0.1-200 Hz band-pass filter as in the time-frequency analysis). Epochs time-locked to the onset of the virtual partner were extracted from the continuous data and for the analysis of event-related potentials they were baseline corrected over a 100-ms pre-stimulus interval and then averaged. Regarding the time-frequency analysis, we used an earlier baseline interval (-170 to -70 ms), because post-stimulus activity was smeared into the pre-stimulus period due to the time-frequency transformation. We did not observe such smearing for the touch stimulus, possibly because of the rise-time of the vibration. Time-frequency transformation was performed over a frequency range of 5-60 Hz.

We focused on activity occurring just prior to the touch (mean over 483-633 ms time-interval). In line with existing literature there was a Late Positive Potential (LPP) in this time-interval that is believed to reflect prioritization in face processing (Langeslag et al., 2007; Schupp et al., 2004) and a marked decrease in low-frequency power (Genna et al., 2017; Haegens et al., 2011, van Ede et al., 2011). These effects are typically found at central sites. We focused our analysis on a central site (C1/C2; Figure 3C) ipsilateral to own gaze direction, because the ipsilateral hemisphere predominated regarding visual processing (visual stimuli were in the visual hemifield contralateral to own gaze direction).

### Neural activity after onset of the touch

We analyzed the time-frequency decomposition over a frequency range of 5-60 Hz and event-related potentials, focusing on the peak of the first negative deflection (N1) in the event-related potential contralateral to the touch (Genna et al., 2017; Katus et al., 2012; minimum voltage within a 130-230 ms time-interval per participant, per condition).



**Figure S1:**

The influence of joint attention neural modulations prior to the touch

A) Amplitude of the Late Positive Potential (LPP) was larger when own gaze and observed gaze were both to the left or both to the right (joint attention) than when one was to the left and the other to the right (disjoint attention), possibly reflecting prioritization of processing joint over disjoint attention. LPP amplitude was averaged over the pre-touch period in which the virtual partner displayed averted gaze. The central site-of-interest (white dot in head shown in panel C) was ipsilateral to own LPP gaze direction, because the visual stimuli were in the contralateral visual hemifield (due to averted own gaze). Diamonds and squares indicate gaze to left and right, respectively. Grey and black symbols refer to own and observed gaze, respectively. Negativity is plotted upward. Error bars indicate  $\pm$ SEM. Inset on the right shows the event-related potential averaged across conditions.

B) Decrease in low-frequency power (5-20 Hz; including theta, alpha and low beta) was larger in amplitude for joint than for disjoint attention, possibly reflecting enhanced processing of joint compared with disjoint attention. Conventions as in panel A. Power was averaged over the same period as the LPP (panel A). Inset on the right shows the time-frequency spectrum averaged across conditions.

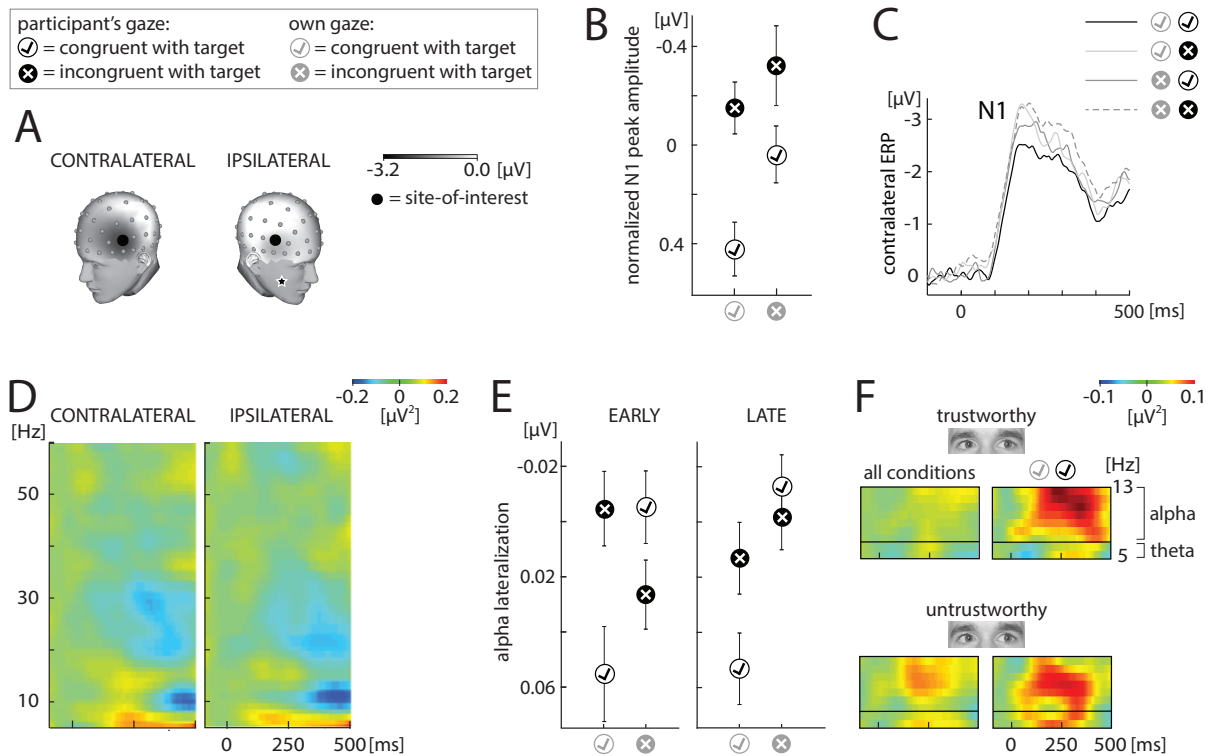
C) Comparison between sites-of-interest for pre-touch / extended visual processing (ipsilateral central site; white dots), touch processing (fronto-central site; black dots) and early visual processing (posterior site; black triangles). The differences between joint and disjoint attention were most profound on the central site. Only at the posterior site did the low-frequency power decrease differ between the hemispheres: it was larger ipsilateral to own gaze and contralateral to observed gaze direction. Small head illustrates scalp locations of the analyzed sites. Analyzed period and frequency range as in panel B. Error bars indicate  $\pm$ SEM. Asterisks indicate differences from zero ( $p < 0.05$ ).

## Supplementary Results

### Neural activity prior to the touch

At the central site-of-interest ipsilateral to own gaze direction the Late Positive Potential (LPP) was larger for joint than for disjoint attention (Figure S1A; 0.64  $\mu\text{V}$  on average; own gaze x observed gaze:  $F_{1,23}= 19.2$ ,  $p= 0.0002$ ). Contralaterally, this effect was small and nonsignificant ( $F_{1,23}= 2.9$ ,  $p= 0.1$ ). Low-frequency power decrease at the ipsilateral site was also larger for joint than for disjoint attention (Figure S1B; 5-20 Hz frequency range, i.e. theta, alpha and low beta; own gaze x observed gaze:  $F_{1,23}= 9.3$ ,  $p= 0.006$ ). This effect was also present contralaterally ( $F_{1,23}= 4.3$ ,  $p= 0.049$ ).

To assess differences between early (posterior) and extended (central) visual processing and to assess pre-target activity and the fronto-central site-of-interest for touch processing we compared activity at these 3 sites (Figure S1C). Differences between joint and disjoint attention were most robust at - but not limited to - the central site. Only at the posterior site did the low-frequency power decrease differ between the hemispheres: it was larger contralateral to *observed* gaze direction (in line with Thut et al., 2006;  $t_{(23)}= -2.5$ ,  $p= 0.02$ ) and smaller contralateral to *own* gaze direction ( $t_{(23)}= 6.7$ ,  $p= 0.0000007$ ; Figure S1C).



**Figure S2:**

The influence of joint attention on touch processing

A) Scalp distribution of the negative deflection associated with tactile processing (N1; averaged across conditions) and fronto-central site-of-interest (black dot). Black star illustrates touch location.

B) contralateral N1 peak amplitude (normalized by subtracting mean per participant) was smaller, i.e. less negative, with own gaze and/or observed gaze toward the touch hemispace, but there was no significant interaction and thus no modulation by joint attention.

C) Event-related potential time-locked to onset of stimulator. Negativity is plotted upward. Conventions as in Figure 2.

D) Time-frequency spectrum contralateral (left) and ipsilateral (right) relative to the touch stimulus at the fronto-central site-of-interest, showing bilateral increase in theta power and subsequent decrease in alpha power.

E) The pattern of alpha lateralization at the site-of-interest resembled the modulation of reaction times more in a late (250-450 ms) than in an early (175-250 ms) time-interval. Negativity is plotted upward.

F) The difference in alpha lateralization at the site-of-interest between jointly attended and other touch locations was evident for trustworthy (top) but not untrustworthy (bottom) virtual partners.



### Neural activity after onset of the touch

The N1 was localized at the contralateral fronto-central site-of-interest (Figure S2A). The peak was earlier and smaller (less negative) when the participant and/or the partner gazed toward compared with away from the touched hemifield (Figure S2B; observed gaze influenced amplitude only), but there was no interaction (own gaze x observed gaze: both  $F_{1,21} < 0.4$ , both  $p > 0.5$ ; Figure S2B-C).

The 5-60 Hz time-frequency decomposition was marked by a bilateral theta power increase and a subsequent bilateral alpha power decrease, in line with previous reports (Genna et al, 2017; Hu et al, 2013). There were no other power changes, apart from a modest bilateral beta power decrease (Figure S2D). The pattern of alpha lateralization reported in the main text (Figure 4) developed over time to resemble the modulation of reaction times more in a late (250-450 ms) than in an early (175-250 ms) time-interval (Figure S2E).

An overall analysis yielded no significant influences of trustworthiness. However, when analyzed separately, trustworthy faces showed the above-described pattern of alpha lateralization (own gaze x observed gaze:  $F_{1,21} = 8.7$ ,  $p = 0.008$ ), while untrustworthy faces did not (own gaze x observed gaze:  $F_{1,21} = 1.5$ ,  $p = 0.2$ ; Figure S2F). Differences between joint and disjoint attention that occurred irrespective of the direction of joint attention were present for trustworthy as well as untrustworthy faces, indicating trustworthiness did not modulate general/non-spatial influences of joint attention.

## Supplementary Discussion

### Neural activity prior to the touch

Neural activations prior to the touch were enhanced for joint compared with disjoint attention. An increase in the amplitude of the Late Positive Potential (LPP) indicated enhanced/prioritized processing of joint compared with disjoint attention (Figure S1A; Langeslag et al, 2007; Schupp et al, 2004; Striano et al, 2006). In addition, a decrease in low-frequency power was larger for joint compared with disjoint attention, suggesting more elaborate mental processing or a larger hedonic value of joint attention (Figure S1B; Pfeiffer et al, 2013; Redcay et al, 2010/2012; Schilbach et al, 2010). A suppression of mu-alpha rhythm may have contributed to this finding, possibly demonstrating involvement of the mirror neuron system in joint attention (Lachat et al., 2012; Pfeiffer et al., 2013; Saito et al., 2010). The modulation of low-frequency power by joint attention was similar in both hemispheres, congruent with a non-spatial interpretation of this effect. These findings show that sharing gaze with a virtual partner is associated with enhanced processing relative to gazing in opposite directions, irrespective of the direction of joint attention.

In line with previous literature on spatial cueing, we observed stronger low-frequency power decrease contralateral compared with ipsilateral to the observed gaze direction (Figure S1C; Haegens et al., 2011; Thut et al., 2006; van Ede et al., 2011).

### Neural activity after onset of the touch

In line with the behavioral results reported in the main text, there was a trend suggesting that alpha lateralization unique for jointly attended touch was present for trustworthy but not untrustworthy partners (Figure S2F), while neural differences reflecting gaze following per se or reflecting non-spatial differences between joint and disjoint attention were present for both trustworthy and untrustworthy faces.

The negative event-related deflection associated with tactile processing (N1) was smaller (less negative) when participant and/or virtual partner gazed toward the touch hemisphere (Figure S2B-C; there was no modulation by joint attention). Contrarily, previous reports found a larger instead of

smaller amplitude for gazed-at stimulus locations (de Jong et al., 2008; Hietanen et al., 2008; Schuller & Rossion, 2004; see also Katus et al., 2012). It could be that the N1 modulation emerged on an earlier positive deflection (P1), indicating a larger (more positive) amplitude of this peak. Unfortunately, the P1 was too small and noisy to be analyzed.

### **Supplementary References**

- Haegens S, Handel BF, Jensen O. 2011. Top-down controlled alpha band activity in somatosensory areas determines behavioral performance in a discrimination task. *J Neurosci.* 31(14):5197-5204.
- Katus T, Andersen SK, Muller MM. 2012. Maintenance of tactile short-term memory for locations is mediated by spatial attention. *Biol Psychol*, 89(1), 39-46.
- Lachat F, Hugueville L, Lemarechal JD, Conty L, George N. 2012. Oscillatory brain correlates of live joint attention: a dual-EEG study. *Frontiers in Human Neuroscience.* 6(156):1-12.
- Langeslag SJ, Jansma BM, Franken IH, Van Strien JW. 2007. Event-related potential responses to love-related facial stimuli. *Biol Psychol.* 76(1-2):109-115.
- Redcay E, Dodell-Feder D, Pearrow MJ, Mavros PL, Kleiner M, Gabrieli JD, Saxe R. 2010. Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. *Neuroimage.* 50(4):1639-1647.
- Redcay E, Kleiner M, Saxe R. 2012. Look at this: the neural correlates of initiating and responding to bids for joint attention. *Front Hum Neurosci.* 6(169):1-14.
- Saito DN, Tanabe HC, Izuma K, Hayashi MJ, Morito Y, Komeda H, Uchiyama H, Kosaka H, Okazawa H, Fujibayashi Y, Sadato N. 2010. "Stay tuned": inter-individual neural synchronization during mutual gaze and joint attention. *Front Integr Neurosci.* 4(127):1-12.
- Schilbach L, Wilms M, Eickhoff SB, Romanzetti S, Tepest R, Bente G, Shah NJ, Fink GR, Vogeley K. 2010. Minds Made for Sharing: Initiating Joint Attention Recruits Reward-related Neurocircuitry. *Journal of Cognitive Neuroscience.* 22(12):2702-2715.
- Schuller AM, Rossion B. 2004. Perception of static eye gaze direction facilitates subsequent early visual processing. *Clin Neurophysiol.* 115(5):1161-1168.

- Schupp HT, Ohman A, Junghofer M, Weike AI, Stockburger J, Hamm AO. 2004. The facilitated processing of threatening faces: an ERP analysis. *Emotion*. 4(2):189-200.
- Striano T, Reid VM, Hoehl S. 2006. Neural mechanisms of joint attention in infancy. *Eur J Neurosci*. 23(10):2819-2823.
- Thut G, Nietzel A, Brandt S. A, Pascual-Leone A. 2006. Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci*. 26(37):9494-9502.
- van Ede F, de Lange F, Jensen O, Maris E. 2011. Orienting attention to an upcoming tactile event involves a spatially and temporally specific modulation of sensorimotor alpha- and beta-band oscillations. *J Neurosci*. 31(6):2016-2024.