

# **Optical imaging during toddlerhood: brain responses during naturalistic social interactions.**

**Yoko Hakuno<sup>1, 2, \*</sup>, Laura Pirazzoli<sup>3</sup>, Anna Blasi<sup>3</sup>, Mark H. Johnson<sup>3</sup>, and Sarah Lloyd-Fox<sup>3</sup>**

1. Department of Psychology, Graduate School of Human Relations, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan

2. Research Fellow of Japan Society for the Promotion of Science, 5-3-1 Kojimachi, Chiyoda-ku, Tokyo 102-0083, Japan

3. Centre for Brain and Cognitive Development, Birkbeck, University of London, Malet Street, London WC1E 7HX, United Kingdom

---

\* Corresponding author

Department of Psychology, Graduate School of Human Relations, Keio University

Address: Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan.

Tel: +81-(0)3-5427-1873, Fax: +81-(0)3-5427-1578

E-mail: yokohakuno@gmail.com (Y. Hakuno).

## **Abstract**

Despite the importance of our ability to interact and communicate with others, the early development of the social brain network remains poorly understood. We examined brain activity in 12- to 14-month-old infants while they were interacting live with an adult in two different naturalistic social scenarios (i.e., reading a picture book versus singing nursery rhymes with gestures), as compared to baseline (i.e., showing infants a toy without eye contact or speech). We used functional near-infrared spectroscopy (fNIRS) recorded over the right temporal lobe of infants to assess the role of the superior temporal sulcus - temporoparietal junction (STS-TPJ) region during naturalistic social interactions. We observed increased cortical activation in the STS-TPJ region to live social stimuli in both socially engaging conditions compared to baseline during real life interaction, with greater activation evident for the joint attention (reading book) condition relative to the social nursery rhymes. These results supported the view that the STS-TPJ region, engaged in the cortical social brain network, is already specialized in infants for processing social signals and is sensitive to communicative situations. This study also highlighted the potential of fNIRS for studying brain function in infants entering toddlerhood during live social interaction.

**Keywords:** social interactions, social signals, fNIRS, joint attention, superior temporal sulcus, temporoparietal junction

## **1. Introduction**

Social and linguistic input from others during infancy and childhood plays an important role in typical human development. Early social deprivation (e.g., low-resource orphanage care-giving) can result in specific cognitive and behavioral dysfunctions [1-4]. Typically during mother-infant interactions, infants receive a rich variety of communicative signals from caregivers, such as eye contact, infant-directed speech, touch, and contingent responsiveness [5-8], and they seem to be sensitive to those signals soon after birth [9-11]. Moreover, previous studies have found that the presence of such social signals during mother-infant interactions can enhance early language development [12-16]. Although preverbal infants might not understand the uses of social signals in the same way as older infants and toddlers do [17-19], social interactions directly impact development from an early age.

Brain regions related to processing of a social stimulus have become known as the social brain network [20]. In human adults, lesion and functional neuroimaging studies have revealed that this network includes parts of the prefrontal cortex (PFC) including the orbitofrontal cortex (OFC), amygdala, fusiform face area (FFA), temporoparietal junction (TPJ), and superior temporal sulcus (STS) [21, 22]. For example, Carrington and Bailey [23] reviewed functional neuroimaging studies examining theory of mind (ToM) skills and found that the PFC, STS, TPJ and anterior cingulate cortex (ACC) appeared to be core regions for ToM reasoning. With the aid of methodological advancements in infant neuroimaging studies, the ontogeny of the social brain network has also been investigated in recent years; we are beginning to understand how the social brain network becomes functional in early infancy [24, 25].

Electroencephalography (EEG) studies have identified early social brain responses to human

faces [26-28], emotional expressions [29], mutual gaze [9], biological motion [30-32], and joint attention [33, 34], with some evidencing immature responses compared to adults [27, 28].

Moreover, functional near infrared spectroscopy (fNIRS) studies, which allow for more precise spatial localization, have localized processing of these social signals in preverbal infants to the inferior frontal and STS-TPJ (which includes posterior regions of the superior temporal gyrus, sulcus, middle temporal gyrus, and the TPJ) regions of the cortex [31, 35-44]. For example, Otsuka, Nakato, Kanazawa, Yamaguchi, Watanabe and Kakigi [39] measured 5- to 8-month-old infants' brain responses with fNIRS and revealed that watching upright face stimuli, but not inverted face stimuli, resulted in activation in the right STS region. Lloyd-Fox, Blasi, Volein, Everdell, Elwell and Johnson [37] found greater cortical activation, at 5 months of age, in the bilateral posterior temporal cortex in response to social dynamic stimuli, such as video clips of a female actor performing hand games (e.g., "peek-a-boo"), compared to non-social dynamic stimuli, such as video clips of a moving mechanical toy. Using NIRS - magnetic resonance imaging (MRI) co-registration data [45] has recently allowed us to more confidently localize these responses within the posterior STS-TPJ region. Furthermore, recent research has found that social signals related to auditory stimuli such as mouth movements [31] and vocalizations [46-48] enhance activation in the anterior STS-TPJ region. These results suggested that young infants already have specialized areas in the temporal cortex for processing social stimuli.

In an effort to conduct strictly controlled studies, video stimuli are more commonly used than live stimuli to test social perception in infants. However, it is known that extracting information from video demonstrations can be more difficult for infants and toddlers than from

live demonstrations [49, 50]. For example, a behavioral study with 9- to 10-month-olds showed that infants learned a non-native phonetic contrast only from live exposure to a foreign language and not from televised exposure [51]. Additionally, Shimada and Hiraki [52] tested 6- to 7-month-old infants with fNIRS to see the difference between their cortical responses to live and televised actions. Though they found activation in the sensorimotor area during both live and televised action observation, the crucial contrast between activity to action and object-motion perception was only found in the live condition. Furthermore, an event-related potentials (ERPs) study revealed that infants responded to three-dimensional representations of objects more quickly than two-dimensional representations [53].

Since fNIRS is less susceptible to motion artifacts relative to functional MRI and EEG, it is an ideal imaging technique for studying infants in more ecologically valid settings. Recent fNIRS research has taken advantage of this attribute to study activity while infants are socially interacting with others [41, 42]. Recently, Lloyd-Fox, Széplaki-Köllöd, Yin and Csibra [42] studied 6-month-old infant responses to different combinations of ostensive social cues during live performances of nursery rhymes with gestures, in an infant-directed or adult-directed way. They found an additive effect of infant directed cues such that the combination of direct gaze and infant-directed speech enhanced activation in bilateral STS-TPJ regions relative to direct gaze with adult directed speech and averted gaze with infant-directed speech.

In order to explore the role of the STS-TPJ region in infants further, we extended the complexity of the interactions, and conducted a fNIRS study with an older cohort of 12- to 14-month-old infants during naturalistic play scenarios. We investigated infants' brain responses to social signals during two different social interactions with an adult experimenter.

In the interaction (I) condition, the infants viewed a female experimenter singing infant directed nursery rhymes (with accompanying hand gestures, such as “Peek-a-boo”), with direct gaze and a positive emotional face (i.e. smiling, wide eyes). In the joint attention (JA) condition, the experimenter used infant directed gaze, speech and gestures to look at a picture book. During the baseline period, the experimenter presented a toy to the infants without any social signals (with averted gaze and no speech). We measured brain activation across the infants’ right inferior frontal – posterior temporal cortex. Firstly, we hypothesized that both socially engaging experimental conditions (I and JA) would show enhanced activation relative to baseline in this region, in particular the anterior superior temporal gyrus - middle temporal gyrus (aSTG-MTG). Furthermore, given that the STS-TPJ region has been indicated to be involved in false belief tasks in older toddlers, children and adults, we hypothesized that this region may be more heavily recruited during interactions such as joint attention, which are thought to be precursors of ToM [54]. Such activity would indicate that by 12- to 14-month-old infants understand that there is an object of mutual interest requiring interpersonal interaction in the JA condition relative to the dyadic I condition.

## **2. Methods**

### *2.1 Participants*

Thirty healthy 12- to 14-month-old infants ( $M = 395.4$  days,  $S.D. = 23.5$ , 20 males) participated in the study. We excluded 4 additional infants from the analysis due to experimental error ( $n = 3$ ) or heavy fussiness ( $n = 1$ ). Of thirty infants, two infants were excluded from the behavioral analysis, but not from the fNIRS analysis, because of experimental error, and nine infants were

excluded from the fNIRS analysis, but not from the behavioral analysis, due to signal quality problems caused by movements ( $n= 4$ ), an insufficient number of valid trials because of fussiness or inattention ( $n= 3$ ), experimental error ( $n =1$ ), and failure of wearing the fNIRS headgear ( $n = 1$ ). Thus, twenty-eight infants ( $M = 396.7$  days,  $S.D. = 23.8$ , 19 males) were included in the final behavior analysis and twenty-one infants ( $M = 388.4$  days,  $S.D. = 19.7$ , 11 males) were included in the final fNIRS analysis. For these twenty-one infants, the mean head circumference, the mean semi-circumference from the left to the right ear via the top of head, and the mean semi-circumference between the pre-auricular points via the forehead were 47.3 cm ( $S.D. = 1.4$ ), 26.9 cm ( $S.D. = 1.3$ ), and 22.9 cm ( $S.D. = 1.0$ ) respectively. All participants were from a volunteer database at the Centre for Brain and Cognitive Development and gave written informed consent before they participated in the study. The study protocol was approved by the Departmental Ethics Committee, Department of Psychological Science, Birkbeck, University of London (Reference Number: 131451).

## *2.2 Stimuli*

Stimuli were presented in a live setting by a female experimenter who sat in front of the infants throughout the experiment. There were two experimental conditions: the joint attention (JA) and interaction (I) conditions (Figure 1). Within the JA condition, the experimenter interacted with the infant using a picture book. She made frequent eye contact with the infant to provoke joint attention by eye gaze, she used gestures such as pointing, as well as verbally labeling the pictures, and she displayed positive affect, with wide eyes and smiles. Within the I condition, the experimenter maintained eye contact and displayed positive affect, with wide eyes and

smiles, while either singing children's songs with hand actions or talking and performing hand games, (there were 4 different songs/games which alternated across the session, and included 'Incy Wincy Spider,' and 'peek-a-boo'). During the baseline condition, the experimenter was silent, avoided eye contact and looked towards the toy, which she held between the infant and herself, and presented a selection of moving toys, which made quiet sounds. There were four different toys to maintain the infant's attention during the baseline condition, and the experimenter picked one of them randomly for each trial. When the experimenter was not using the toys or book she placed them on a table next to her, and out of reach, but within view of the infant. A tone played every 20 s via speakers to indicate the onset and offset of each trial for Experimenter 1 and Experimenter 2, who placed event markers manually into the fNIRS recordings online. Four different researchers served as the "partner" in these real life interactions during the course of data collection for this study. Of thirty infants, fifteen infants interacted with Experimenter A, eight infants interacted with Experimenter B, four participants interacted with Experimenter C, and a further three infants interacted with Experimenter D. Prior to the study the Experimenter learnt a scripted set of instructions about how to interact during each of the conditions, and reviewed training videos generated during piloting of the live study to mimic this "standard". Instructions included the following: 1) natural transitions between trials so that when the end of the trial was signaled by the audio signal any action or sentence was completed before finishing and moving to the next trial; 2) toys and books were kept on a table next to the experimenter and returned to position on completion of each trial; 3) during the I condition they learnt 4 different songs or games (with speaking and hand actions) and alternated these across trials; 4) use infant-directed speech, eye contact, positive affect and



gestures during the experimental trials; 5) during the baseline condition watch the object and help make it move/play sounds but avoid eye contact. Videos of the experimenter interactions were reviewed to ensure that these guidelines were followed. Data from two infants (outlined as experimental error above) were excluded from the analyses as it could be seen from the videos that they interacted with the infants and used direct eye contact during the baseline trials.

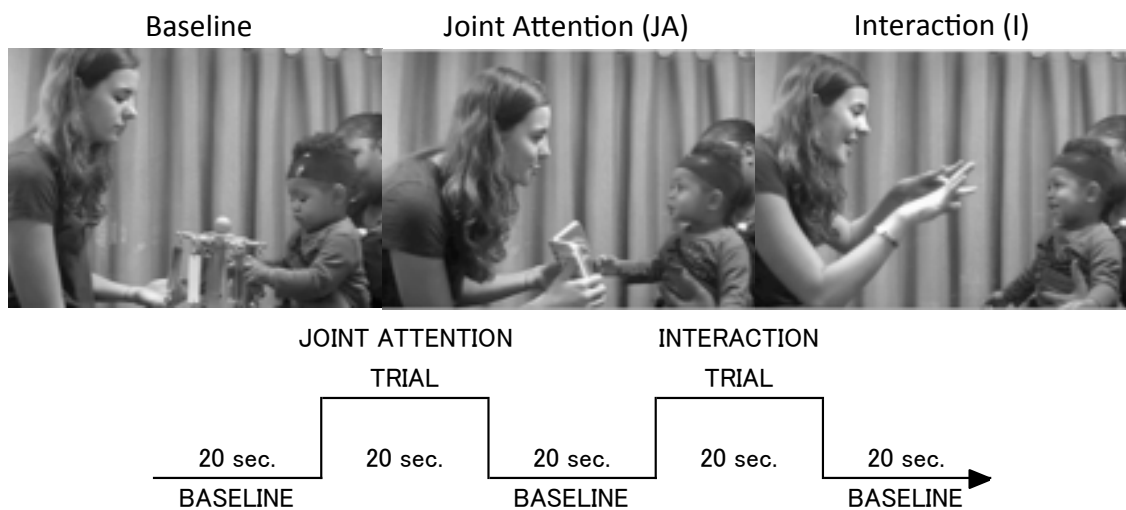


Figure 1. An infant participating in the study during the baseline and two experimental conditions (upper panel). The experimental protocol showing the order and timing of stimulus presentation for the two experimental conditions (lower panel).

### 2.3 Procedure

Within the same session each infant went through three different fNIRS experiments: two video presented experiments [48] and one live experiment. After the two video studies investigating infants' social perception and memory, we moved on to the live study, which is the focus of this study. At the beginning of the session infants sat on their parent's lap facing a female experimenter in a naturally lit room while the fNIRS headgear was quickly placed on their

heads by Experimenter 2. The parent was instructed not to interact with their infant during the stimulus presentation except when the infant became fussy or inattentive. The two types of experimental condition trials, joint attention (JA) and interaction (I), were presented in between baseline trials (Figure 1). Each trial lasted 20 s. The order of experimental condition presentation was always JA, I, I, JA, I, JA, JA, I, JA, I in a repeating loop. A camera on a tripod placed on the left side of the infant recorded both the infant and the experimenter. When the infant became bored or fussy, as judged by the experimenter, or completed 6 trials for each experimental condition, we stopped presenting stimuli and ended the experiment.

#### *2.4 Procedural Integrity*

To ensure that the live demonstration was conducted consistently by the experimenters across infants, the first author and an independent observer scored procedural integrity from videotape on approximately 25% of the samples (2 videos from each experimenter) using 20-seconds interval scoring. Note that experimenters contributed to a differing number of sessions (Experimenter A:  $N = 10$ , Experimenter B:  $N = 6$ , Experimenter C:  $N = 2$ , Experimenter D:  $N = 3$ ). Cohen's kappa [55] was 0.62, which was considered to be **substantial agreement** [56]. The percent of trials in which the following behaviors were conducted correctly were: infant-directed speech (Experimenter A: 100%, Experimenter B: 100%, Experimenter C: 100%, and **Experimenter D**: 100%), use of simple and short sentences (Experimenter A: 100%, Experimenter B: 100%, Experimenter C: 100%, and **Experimenter D**: 100%), positive expression (Experimenter A: 100%, Experimenter B: 100%, Experimenter C: 94%, and **Experimenter D**: 75%), and fluid action (Experimenter A: 100%, Experimenter B: 100%,

Experimenter C: 94%, and **Experimenter D**: 75%). In addition, the experimenters' mean proportions of fixation duration toward the infant's face were counted for approximately 25% of the samples (the same 2 videos from each experimenter), and were 40.7% (*S.D.* = 18.2) for the JA condition (Experimenter A: *M* = 48.7%, *S.D.* = 9.9; Experimenter B: *M* = 57.5%, *S.D.* = 9.3; Experimenter C: *M* = 35.9%, *S.D.* = 15.7; Experimenter D: *M* = 20.9%, *S.D.* = 12.7) and 96.2% (*S.D.* = 4.5) for the I condition (Experimenter A: *M* = 95.7%, *S.D.* = 4.6; Experimenter B: *M* = 96.8%, *S.D.* = 6.9; Experimenter C: *M* = 96.5%, *S.D.* = 3.2; Experimenter D: *M* = 95.6%, *S.D.* = 3.3). Given that Experimenters C and D were seen to be looking at infants for a lower percentage of time during the JA condition relative to A and B, we re-ran the fNIRS analyses excluding the 5 infants who observed these experimenters and found no overall change in the pattern of fNIRS activation across the condition-contrasts.

### *2.5 Data recording and processing*

A University College London (UCL)-NIRS mini topography system was used [57]. The multichannel continuous-wave mini fNIRS system uses two wavelengths at 780 and 850 nm. Six light sources (per wavelength) and four detectors were arranged within custom-built fNIRS-CBCD headgear consisting of an array over the right temporal lobe with a total of 12 2-cm source-detector channels (Figure 2). The midpoint of the lower row of channels was aligned with the pre-auricular point on the right hemisphere (scalp location T4, according to the 10–20 system) for each infant. Due to the limited number of sources and detectors of our fNIRS system, we chose to maximize coverage across one hemisphere overlaying inferior frontal to posterior temporal regions.

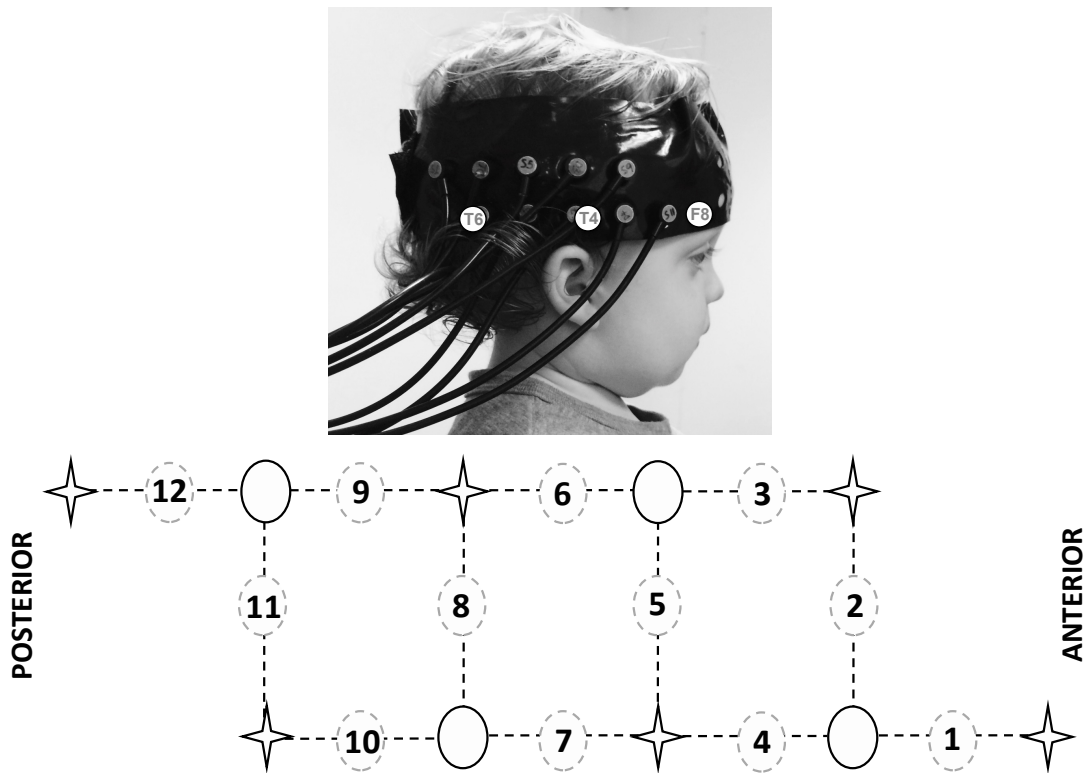


Figure 2. The fNIRS headgear and channel layout. An infant wearing the headgear with the location of the 10-20 coordinates F8, T4, and T6 overlaying the photo. T4 was located at the midpoint of the lower row of channels (the source optode between channel 4 and 7) (upper panel). The array design showing the location of the channels (dashed circle), sources (star), and detectors (full circle) (lower panel).

Prior to analyzing the behavioral data, trials were rejected from further analysis by looking time measures. Videos were coded offline by an experimenter: if an infant looked for less than <20% within a trial, the trial was considered invalid and not included in the final dataset. A minimum of two valid trials per condition was set as a threshold for inclusion within infants for the behavioral data. We couldn't obtain detailed looking time measures from two participants because we failed to video-record the session. Therefore, we excluded these two

participants from the behavioral analysis. However, we included them in the fNIRS analyses as our secondary measure of looking time (online live coding to record whether infants attended for at least 50% of each trial to guide the second experimenter's presentation of a sufficient number of trials during the session) indicated that the infants' exceeded the threshold for inclusion in analyses. Note that reanalysis of the fNIRS data excluding these infants produced the same overall pattern of results.

Following this, for the behavioral data, two kinds of behaviors were selected as dependent variables: (i) the average fixation duration toward the experimenter's face, hands (I condition) or object (JA and baseline condition), proportional to the length of the trial, and (ii) frequencies of 4 different social behaviors (gaze re-orienting, vocalizing, smiling, and pointing) per trial. For the gaze behavior, we counted the numbers of times the infant shifted their eye gaze between the experimenter's face and the hand/object, and for the other 3 behaviors we counted the number of times we observed these distinct actions for each trial. Vocalizations were counted as periods of vocalizations except negative voiced sounds such as crying and whimpering. If the separation interval of voiced sounds was more than 1 s, the vocalizations were considered to be discrete. We then averaged this number across trials for each behavior. Each dependent variable was coded by a primary coder at 100 ms intervals using behavioral coding software (GenobsX, Tokyo, Japan).

For the fNIRS data, changes in HbO<sub>2</sub> and HHb chromophore concentration ( $\mu\text{Mol}$ ) were calculated and used as hemodynamic indicators of neural activity [58]. Prior to analyzing the fNIRS data, trials, channels or participant data were rejected from further analysis by (1) looking time measures (videos were coded offline by an experimenter: <50% of looking away

within a trial for the experimental conditions and <20% of looking away for the baseline condition considered invalid) and (2) the quality of the intensity signals, using artifact detection algorithms [37, 59]. In line with previous work, channels were excluded if the coefficient of variation of the attenuation exceeded 10% or if the normalized power was larger than 50% with respect to the total power [37]. Once the attenuation data was converted into changes in concentration (see below) trials were then assessed individually for artefact. Trials were removed if concentration changes during the stimulus trials exceeded +/- 5  $\mu\text{Mol}$ . This threshold was lenient, and designed to ensure data was excluded due to abrupt changes in signal caused by motion rather than changes due to activation. For each infant, the trials and channels that survived these rejection criteria were entered into further analyses. Inclusion criteria required each channel to contain valid data in both experimental conditions. A minimum of three valid trials per experimental condition was set as a threshold for inclusion within infants, and the maximum number of rejected channels could not exceed one third of the total number of channels.

For each infant, the near-infrared intensity signal was low-pass filtered, using a cutoff frequency of 1.7 Hz. The data were then divided into blocks consisting of 4 s of the baseline trial prior to the onset of the 20 s experimental trial (the JA or I condition), plus the following 20 s baseline trial. This 44 s blocks of data were detrended with a linear fit between the first and last 4 s of the 44 s block, and converted into changes in concentration in  $\text{HbO}_2$  and  $\text{HHb}$  using the modified Beer–Lambert law with an assumption of an age-appropriate differential pathlength factor of 5.13 [60]. For each channel, valid experimental condition trials were then averaged together for each infant, and a time course of the mean change across all

valid trials in HbO<sub>2</sub> and HHb concentration was compiled for each experimental condition. Either a significant increase in HbO<sub>2</sub> concentration, or a significant decrease in HHb, is commonly accepted as an indicator of cortical activation in infant research [59]. In an initial analysis, the grand averaged hemodynamic responses ( $\mu\text{Mol}$ ) of all infants were assessed for each of the two conditions (JA and I conditions). For each channel, the maximum change (or amplitude) in HbO<sub>2</sub> (increase/decrease in chromophore concentration) and HHb (increase/decrease in chromophore concentration) in response to each experimental stimulus was compared with baseline. We followed the procedure of epoch analysis of previous research [42, 48] to assess the maximum hemodynamic change for every 5 s period from 15 s after the experimental stimulus onset to 15 s after the onset of the baseline condition (i.e. 15 – 30 s). We assessed the response from 15 s to 30 s to (a) account for the 5s delay that we observed during video coding between the bleep (and parallel event marking of the data by Experimenter 2) and the time taken for Experimenter 1 to discard the toy and begin the next trial sequence by initiating joint eye contact, and to (b) investigate the latency of the peak responses seen across participants and channels during these live naturalistic interactions. Following analysis of the experimental condition vs. baseline, channels that showed significant activation in at least one condition were entered into paired *t*-tests to compare the hemodynamic response during the specified time windows between the two conditions (JA vs. I). During statistical analyses of the group data, only significant increases in HbO<sub>2</sub> and decreases in HHb were reported, if the signals increased or decreased significantly in unison, the signal was considered inconsistent with a hemodynamic response to functional activation and not included (during analysis of the group data significant increases in HbO<sub>2</sub> and HHb were only evident in one channel (10) at one

time epoch). Channels that survived a multiple comparison correction (false discovery rate: FDR) are highlighted in Table 1 [61]. To our knowledge this was the first study to explore fNIRS responses during live naturalistic interactions with infants entering toddlerhood, therefore we included full analyses of all significant channels (pre-thresholding) in our results.

### **3. Results**

#### *3.1 Behavioral Data*

The average number of valid trials, including baseline trials, was 20.9 (*S.D.* = 4.8) with the mean duration of the recording session totaling 7.0 minutes (*S.D.* = 1.6). The average number of valid trials was 5.1 (*S.D.* = 1.3) for the JA condition and 4.9 (*S.D.* = 1.2) for the I condition. Within these valid trials, two kinds of dependent variables were analyzed, (i) proportion of fixation durations toward the experimenter's face versus to the hand/object, and (ii) frequencies of four different social behaviors (gaze shifting, vocalizing, smiling, and pointing).

Firstly, the mean proportion of looking duration to the stimuli in general (time on stimulus for both the face and the object combined) was 0.86 (*S.D.* = 0.09) for the baseline condition, 0.89 (*S.D.* = 0.08) for the JA condition, and 0.78 (*S.D.* = 0.12) for the I condition. Therefore, overall average looking time to the trials across infants was 17.2 seconds for the baseline condition, 17.8 seconds for the JA condition and 15.6 seconds for the I condition. The infants' proportion of looking time was submitted into a two-way ANOVA with the stimulus condition (Baseline / Joint attention / Interaction) and the time on stimulus (Face / Object-Hand) as within-participants factors (Figure 3). Mauchley's test of sphericity revealed a violation of sphericity for the interaction effect between the stimulus condition and the time on stimulus.



The Greenhouse–Geisser correction was therefore applied to adjust the degrees of freedom for the effect. Results revealed a significant main effect of the stimulus condition ( $F(2, 54) = 13.001, p < 0.001, \text{partial } \eta^2 = 0.325$ ) and the time on stimulus ( $F(1, 27) = 145.136, p < 0.001, \text{partial } \eta^2 = 0.843$ ). Additionally, a significant interaction between the stimulus condition and the time on stimulus ( $F(1.61, 43.57) = 643.090, p < 0.001, \text{partial } \eta^2 = 0.960$ , Greenhouse–Geisser corrected) was also found. The post hoc  $t$ -tests with Bonferroni’s correction for the simple main effect showed that looking duration to the stimuli in general was shorter in the I condition than in the baseline ( $p = 0.003$ ) or JA conditions ( $p = 0.001$ ) while no significant difference in total looking duration was found between the baseline and JA conditions ( $p = 0.640$ ). Additionally, there were significant differences between the proportions of fixation durations to the face and the object in the baseline condition ( $p < 0.001$ ), in the JA condition ( $p < 0.001$ ), and in the I condition ( $p < 0.001$ ). The infants looked at the object significantly longer than the face during the baseline and JA conditions while they looked at the face longer than the hands in the I condition (Figure 3).

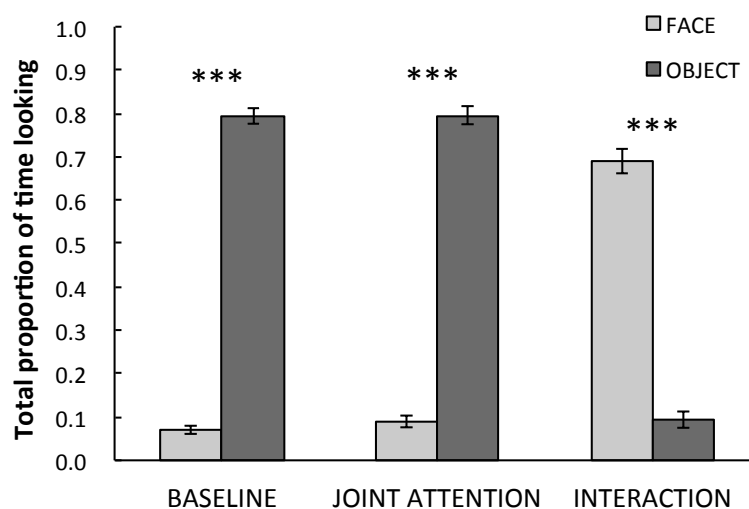


Figure 3. Mean proportions of looking times to the experimenter's face, hands (I condition) or object (JA and baseline condition) for each condition, proportional to the length of the trial. Error bars indicate 1 standard error of the mean ( $N = 28$ ). \*\*\*  $p < 0.001$ .

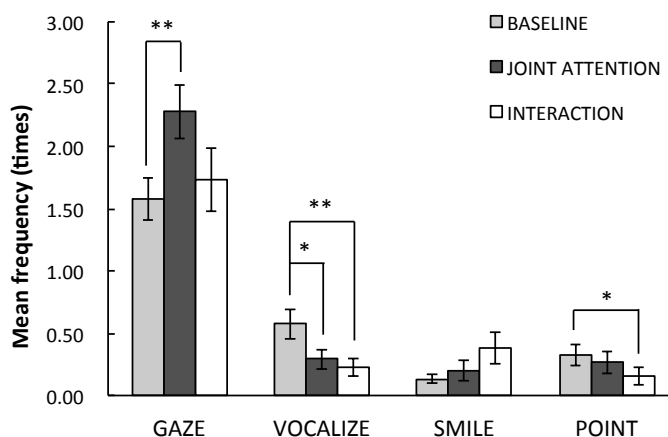


Figure 4. Mean frequencies of four different social behaviors (gaze shifting, vocalizing, smiling, and pointing) per trial for each condition. Error bars indicate 1 standard error of the mean ( $N = 28$ ). \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Secondly, the frequencies of four different social behaviors was submitted into a two-way ANOVA with the stimulus condition (Baseline / Joint attention / Interaction) and the behavior (Gaze shifting / Vocalizing / Smiling / Pointing) as within-participants factors (Figure 4). Mauchley's test of sphericity revealed a violation of sphericity for the effect of behavior and the interaction. The Greenhouse–Geisser correction was therefore applied to adjust the degrees of freedom for those effects. Results reveal no significant main effect of stimulus condition ( $F(2, 54) = 1.447, p = 0.244, \text{partial } \eta^2 = 0.051$ ), but a significant main effect of behavior ( $F(1.82, 49.13) = 66.619, p < 0.001, \text{partial } \eta^2 = 0.712, \text{Greenhouse–Geisser corrected}$ ), and a significant interaction between the stimulus condition and the behaviors ( $F(2.55, 68.90) = 5.398, p = 0.004$ ,

partial  $\eta^2 = 0.167$ , Greenhouse–Geisser corrected). The post hoc  $t$ -tests with Bonferroni's correction for the simple main effect demonstrated that the frequency of gaze re-orienting in the JA condition was higher than in the baseline condition ( $p = 0.008$ ) but not higher than in the I condition ( $p = 0.134$ ). In addition, the infants vocalized significantly longer during the baseline condition compared to the JA ( $p = 0.031$ ) or I conditions ( $p = 0.003$ ), and pointed longer during the baseline condition relative to the I condition ( $p = 0.021$ ).

### 3.2 fNIRS Data

The average number of valid trials including baseline trials was 22.9 ( $S.D. = 2.7$ ) with the mean duration of the recording session totaling 7.7 minutes ( $S.D. = 0.9$ ). The average number of valid trials were 5.7 ( $S.D. = 0.8$ ) for the JA condition and 5.4 ( $S.D. = 0.8$ ) for the I condition. The proportion of invalid channels across the 21 infants was 0.012.

Due to the exploratory nature of this study we included full analyses of all significant channels in our results section but highlight those with the strongest response (that survive FDR correction) in Table 1. A significant increase of HbO<sub>2</sub> compared to baseline for the JA condition was found over 5 channels (channels 5, 6, 7, 8, & 9), while 4 channels revealed a significant increase of HbO<sub>2</sub> for the I condition (channels 5, 6, 8, & 9) (see Figure 5 & 6). The onset and duration of the responses (see Table 1) were earliest and over a longer period in channel 5 and 8 for the JA condition (lasting from 15 – 30 s) and channel 5 for the I condition (lasting from 15 – 30 s). Interestingly, only channels in the JA condition survived the correction for multiple comparisons for this contrast to the low social baseline condition. With reference to a standardized scalp surface map of fNIRS channel coordinates and underlying anatomy for

7-month-old infants [45], and head measurements for our participants we approximated the location of the brain regions underlying these channels in this older age group. The channels with significant activation were clustered largely over the STS-TPJ regions of the cortex (channels 6, 8 & 9). A further 2 channels (5 and 7) with significant responses in the JA condition were located on the edge of this cluster, overlapping into middle temporal regions. Paired-sample channel-by-channel *t*-tests were performed (within those channels showing significant change from baseline) to directly compare the hemodynamic change observed in response to the two experimental conditions (JA versus I). This analysis found a significantly greater hemodynamic increase in HbO<sub>2</sub> to the JA condition relative to the I condition in channel 8 (Channel 8 at 20 – 25 s window:  $t(20) = 2.443$ ,  $p = 0.024$ ), indicating greater activation for the JA condition in this STS-TPJ region. No channels showed a significantly greater increase in HbO<sub>2</sub> to the I condition compared to the JA condition.

Table 1. Channels with significant activation from baseline in the JA and I conditions; statistical tests were performed within different time windows. \*  $p < 0.05$ .

	Channel	Time window (s)	<i>t</i> (20)	<i>p</i> (uncorrected)	<i>p</i> (FDR-corrected)
<b>HbO<sub>2</sub></b>					
<b>JA &gt; Baseline</b>	5	15–20	3.23	0.004	0.030 *
	5	20–25	3.92	0.001	0.014 *
	5	25–30	3.79	0.001	0.014 *
	6	25–30	2.99	0.007	0.042 *
	7	20–25	2.20	0.040	0.143
	8	15–20	2.88	0.009	0.042 *
	8	20–25	3.83	0.001	0.014 *
	8	25–30	3.58	0.002	0.017 *
	9	25–30	2.23	0.037	0.143
<b>I &gt; Baseline</b>	5	15–20	3.02	0.007	0.075
	5	20–25	2.93	0.008	0.075
	5	25–30	3.30	0.004	0.075
	6	25–30	2.33	0.031	0.197
	8	25–30	3.15	0.005	0.075
	9	25–30	2.15	0.044	0.201
<b>HHb</b>					
<b>I &gt; Baseline</b>	1	20–25	-2.38	0.028	0.276
	2	20–25	-2.32	0.031	0.276
	2	25–30	-2.72	0.013	0.237
	3	25–30	-3.22	0.005	0.163
	4	25–30	-2.22	0.038	0.276

Analyses of the HHb signal found a further 4 channels (1, 2, 3 & 4), located over inferior frontal regions, which showed significant decreases in HHb in response to the I condition. However these latter responses should be treated with caution as the decreases in HHb were accompanied by non-significant but similar decreased time courses in HbO<sub>2</sub> (see Appendix A), demonstrating a profile suggestive of contamination by motion artifact, but not to the level that our automatic algorithms would reject. For the HHb changes, there was no significant difference between the JA and I conditions within the epochs.

**(A) Joint Attention > Baseline**      **(B) Interaction > Baseline**

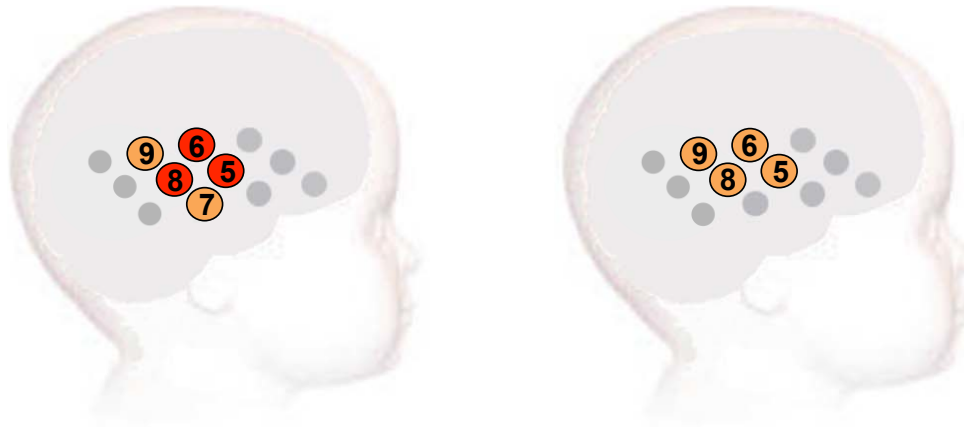


Figure 5. An overview of the significant group effect for HbO<sub>2</sub> (maximum increase in HbO<sub>2</sub> concentration) highlighted with red circles (FDR-corrected  $p < 0.05$ ) and orange circles (uncorrected  $p < 0.05$ ) for the two contrasts: joint attention condition vs. baseline (A), interaction condition vs. baseline (B).

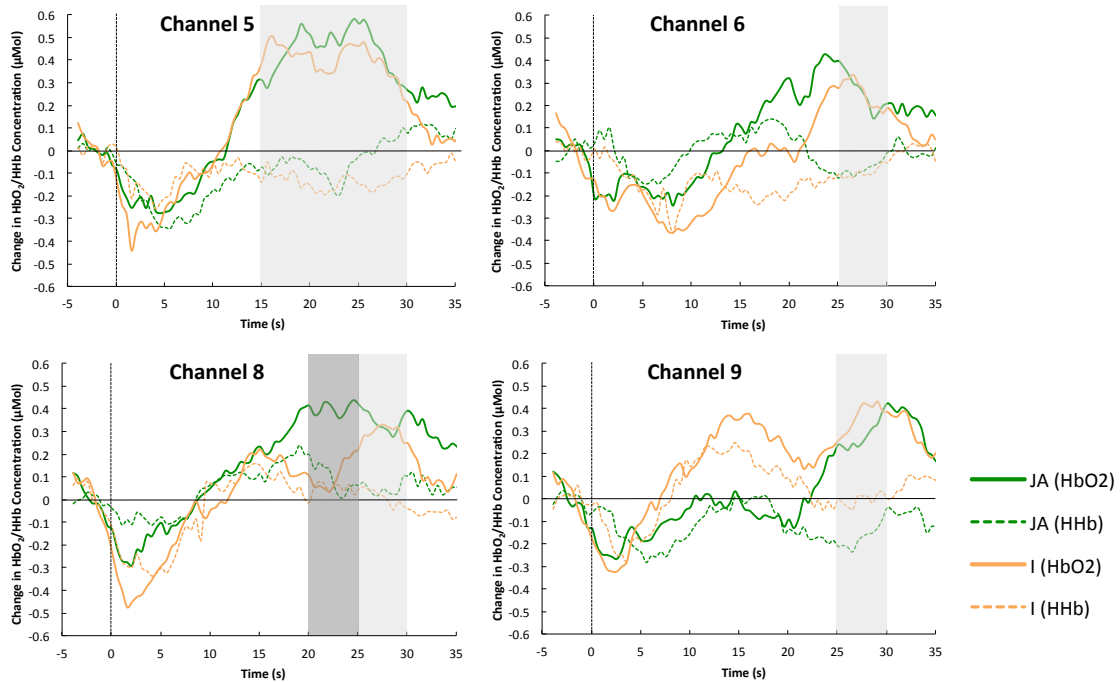


Figure 6. The grand averaged group time courses of the hemodynamic responses in HbO<sub>2</sub> and HHb to the JA and I conditions. Light greyed area indicates the time window where we found a significant increase of HbO<sub>2</sub> compared to baseline for both the JA and I conditions. The dark greyed area indicates the time window where the difference in HbO<sub>2</sub> change was significant between the JA and I conditions.

### *3.3 Relation Between Behavioral and fNIRS Data*

To further explore the relationship between the fNIRS condition specific responses and the behavioral findings we investigated associations between the activation to the I and JA conditions in channel 8 (at 20 – 30 s window) and looking time measures. We chose to assess correlations in channel 8 as it is where we found significant differences in activation between the JA and I conditions and positioned approximately over the posterior part of the STS-TPJ region. Firstly, we did not find that the average fixation duration during JA to the face (Pearson correlation;  $r(19) = 0.180, p = 0.462$ ) or to the object ( $r(19) = 0.139, p = 0.571$ ) significantly correlated with the degree of activation during JA. Therefore it is unlikely that the  $JA > I$  results can be explained by infants who attended more to the face or book during the JA condition driving an increased response during the JA condition. Interestingly, we found a significant positive correlation between the average fixation duration toward the experimenter's face ( $r(19) = 0.518, p = 0.023$ ) during the I condition and changes in HbO<sub>2</sub> in channel 8 during the JA condition (no effects were found for looks towards the object ( $r(19) = 0.071, p = 0.772$ )). This would suggest that the infants who attended to the experimenter's face for longer in the I condition showed greater brain activation in the STS-TPJ region during the JA condition (Figure 7). Also, no statistically significant correlation was found between the proportion of fixation durations toward the face during the I condition and changes in HbO<sub>2</sub> in the STS-TPJ region during the I condition ( $r(19) = 0.236, p = 0.330$ ).

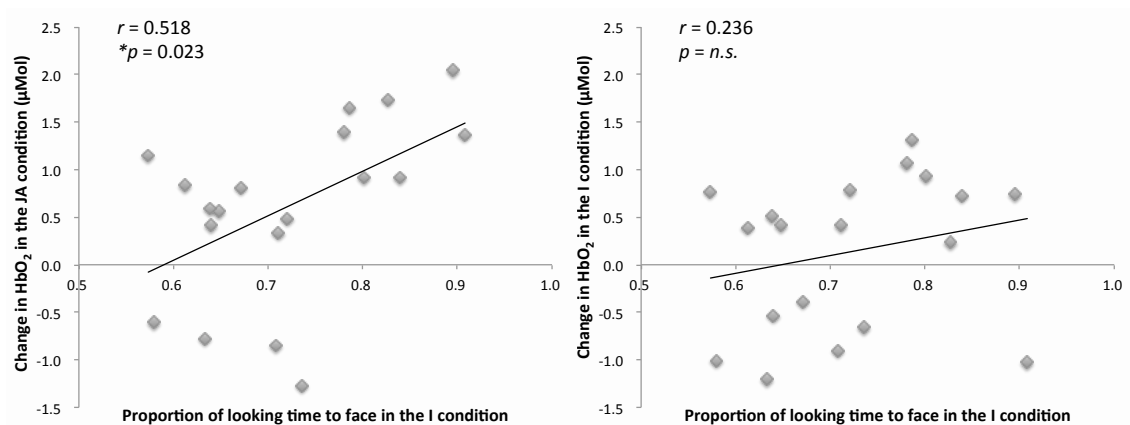


Figure 7. Individual infant responses ( $N = 19$ ) of the association between looking time to the face in the I condition and HbO<sub>2</sub> responses in channel 8 (the STS-TPJ region) in the JA condition (left panel) and I condition (right panel).

#### 4. Discussion

In the current study, we assessed 12- to 14-month-olds' cortical responses during two different types of live communicative interactions with an adult (I condition: singing a nursery rhymes with gestures; JA condition: reading a picture book) to assess the role of the STS-TPJ region in the processing of social signals such as mutual gaze, infant-directed speech, and contingent responsiveness in infants.

Our analyses of infant behavior evidenced significant differences between the JA and I conditions, thus showing that our stimulus conditions were interpreted differently by the infants. Firstly, we found (1) that the proportion of looking time toward the experimenter's face was greater than to the hands during the singing interaction (I) condition whilst the looking time to the face was lower than to the book during the book reading (JA) condition. Secondly during analyses of the frequencies of gaze shifting, smiling, vocalizing and pointing, we found that (2) that the frequency of infant gaze re-orienting (or shifts) between the face and hands/object was significantly higher in the JA condition relative to the baseline; (3) that the frequency of



vocalizations was higher in the baseline condition relative to the JA and I conditions; and (4) that the frequency of pointing gestures was higher in the baseline compared with the I condition.

Our analyses of the fNIRS data evidenced increased HbO<sub>2</sub> responses to both experimental conditions (I and JA) in a STS-TPJ region of the cortex compared to the low-social baseline condition. Furthermore, the response was significantly stronger in the JA compared with the I condition in a channel positioned over the pSTS-TPJ region. Whilst we had anticipated that looking time to the face would be lower during the JA condition relative to the I condition, given the second object of interest (book), we had not anticipated that infant's would switch attention to such a high degree (on average looking to the face for less than 10% of the trial despite the experimenter being positioned directly opposite the infant). Therefore the increased activation during the JA condition is not simply a reflection of observing more visual cues from the experimenter's face. Rather, it suggests that the STS-TPJ region is highly responsive to the communicative context of joint attention whilst the infants were listening to the experimenter describing the book and pointing at the pictures. Furthermore, we found a relationship between the amount of time spent attending to the experimenter's face during the singing (I) condition and the degree of activation during the JA condition. Whilst further work is needed to replicate these findings in additional cohorts and disentangle the relative contribution of the different cues and contexts at this age, one interpretation could be that those infants more interested in attending to adult communication (during periods of directed attention during singing) may evidence more brain activation during other forms of more complex communication such as the joint attention elicited during the sharing of a book. Our results are consistent with developmental models of social cognition where joint attention precedes the

onset of ToM and is a more advanced stage of social cognition involved with interpreting the cues before you, rather than simply observing them passively, such as during the singing condition [62-64].

We found one area of activation, over the aSTG-MTG, to show highly similar hemodynamic time courses (in channels 5 & 6) across both experimental social conditions. This likely represents voice-selective activation in the temporal cortex as the adult spoke in both experimental conditions relative to non-vocal sounds from the toys in the baseline condition. We know from previous research that adults and infants display increased responses in the anterior-middle STS region to human voice relative to non-vocal sounds [46, 48, 65-69]. Indeed, on the same day, the infants in this study were also involved in a study of social visual and auditory (vocal) cues prior to taking part in this live session, and group data from the first study showed a similar localized pattern of activation to vocal > non-vocal sounds accompanied by visual social cues [48]. However, we believe that the observed responses cannot be wholly explained by voice-selective activation. Firstly, the region of activation to the social experimental conditions in the current study extends into the posterior STS-TPJ areas, which have not previously been associated with vocal selectivity. Secondly, these findings are analogous to 6-month-old infant's cortical responses in a previous live interaction study which assessed cortical responses while watching a woman singing with gestures [42]. They found increased activation in very similar regions of the brain to infant-directed singing compared to adult-directed singing, suggesting that these brain areas selectively activated to the type and context of the social cues.

In comparison to the majority of previous infant fNIRS studies, we found that the response of the hemodynamic response function was more sustained (particularly in channels 5 & 8 with significant activation found between 15 to 30 s after stimulus onset). Whilst this response appears to be fairly late, note that this is due to the 5 s delay between the event marked beginning of the trial (marked in the acquisition software by Experimenter 2) and the onset of Experimenter 1's change in interaction as we purposefully avoided unnatural transitions by asking Experimenter 1 to finish the sentence, or action, that they were performing before changing conditions. Therefore the window is in effect covering the last 10s of the experimental trial and the first 5s of the subsequent baseline trial. Interestingly, a previous fNIRS study with similar sustained and/or late occurring activation was the previous live naturalistic interaction study with 6 month olds [42], suggesting that the complexity of the stimuli or attention required may be enhanced in live contexts. Whilst the current design attempted to mimic natural play (for example; the trials were semi-structured in length, the objects and book were set down in view, but out of reach, of the infant on a table next to the adult before beginning the next trial), whilst controlling other variables across trials (gestures were used in the I condition during singing and the JA condition to point to pictures in the book) there were certain differences that could contribute to the current findings (such as reading versus singing) which could not be addressed here. Further research with longer more naturalistic communicative interaction sequences could help clarify the relative contribution of these different factors, and how the current responses relate to different social cues during naturalistic play, singing and conversation. Furthermore, while overall measures of attention (looking time) were not predictive of the level of activation across the three conditions, we may in future work want to

more fully address other factors such as arousal by measuring skin conductance or heart rate across these longer sessions.

According to Cristia, Dupoux, Hakuno, Lloyd-Fox, Schuetze, Kivits, Bergvelt, van Gelder, Filippin, Charron and Minagawa-Kawai [70], most fNIRS studies are carried out on young infants, from neonates to 8-month-olds, and studies on infants older than 8 months are rarely performed. The authors discussed that this may reflect a limitation in the present neuroimaging techniques of testing older infants and toddlers since even fNIRS requires participants to maintain a certain degree of immobility and compliance. They also reported that the attrition rate tends to be higher in older infants (9 months or older) rather than younger infants. In the present study, we succeeded in collecting valid fNIRS data from 21 out of the 34 infants aged 12 to 14 months, resulting in an attrition rate of 38% (falling to 30% if we remove those excluded due to experimenter error). This attrition rate fell below the normal range for fNIRS studies with older infants and was even lower than the average attrition rate of 49% reported for infant ERP studies [71]. This low attrition rate supports the use of fNIRS for measuring brain activity in older infants and toddlers using study designs involving interaction with live stimuli to enhance infant attention.

There are several caveats to the strength of our current conclusions, which should be addressed in future work. Firstly, the limited number of sources and detectors of the mini-NIRS system we used in the study was only sufficient to design a single array and make measurements over inferior frontal to posterior temporal regions on one hemisphere, therefore it would be highly beneficial to continue this research with a wider coverage of the cortex. However we should note that there may be benefits to using a smaller array in live interactions

as the contact of the headgear will be stronger and less affected by artifact from participant movement. Secondly, the current findings do not allow us to fully distinguish between the contribution of the different auditory and visual social cues, and which so future research should use a more complex multifactor design to build on the current findings. Thirdly, whilst we believe from previous research [42] that this semi-structured design was suitable and that it was constructed to be similar to natural interactions with caregivers (who often rapidly switch attention between their infant and other objects in their environment during play), it would be beneficial to extend the current work into longer sessions of naturalistic play and communication to disentangle how these responses interplay across different types of play and communicative contexts. Fourthly, we need to explore the latency and shape of the hemodynamic responses within live contexts with more sensitive analyses to fully interrogate the contribution of different stimulus cues on the pattern of the responses seen.

Our study supports the view that the STS-TPJ region is already engaged in processing of communicative cues related to joint attention by the age of 12 – 14 months. Furthermore, the successful use of fNIRS in this semi-structured design opens the possibility for future research conducted during more natural conversational and play sessions with caregivers, where components of child and adult led interaction and communication could be extracted and analyzed over longer periods of time while individuals are at the lab or at home. This is one of the first fNIRS studies to investigate cortical activation during naturalistic live interactions in infants entering toddlerhood [41, 42] and highlights the potential of this technology in future basic and clinical science.

## Financial Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

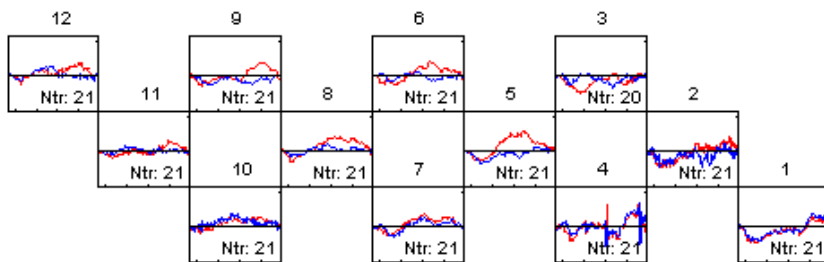
## Acknowledgements

This work was supported in part by JSPS KAKENHI Grant Number JP15J06053 to Yoko Hakuno, and the UK Medical Research Council (G0701484 ID: 85031 & MR/K021389/1) to Mark H. Johnson. We would like to thank all the infants and their parents who participated in this research at the Centre for Brain and Cognitive Development (CBCD), Birkbeck, University of London. We also appreciate the assistance offered by Mine Serttas, Sarah Hearne, Wachen Harris, and Sophie Budge.

## Appendix A

As supplementary data, time courses of hemoglobin responses to the JA and I conditions for all channels were shown in Figure 8.

### Joint Attention



### Interaction

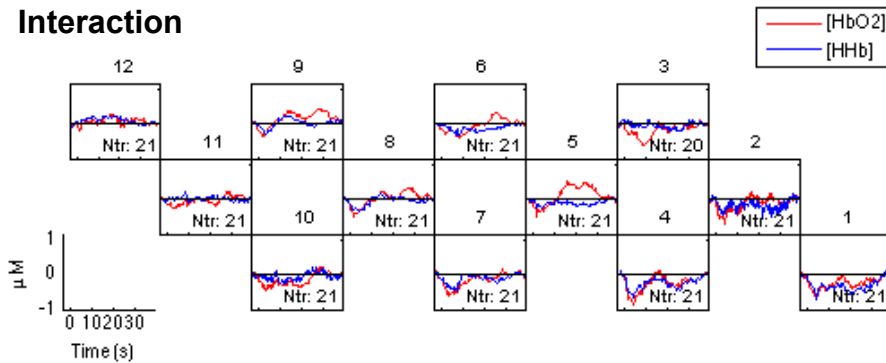


Figure 8. The grand averaged group time courses of the hemodynamic responses in HbO<sub>2</sub> and HHb to the JA and I conditions for channels 1 to 12.

### Figure Captions

Figure 1. An infant participating in the study during the baseline and two experimental conditions (upper panel). The experimental protocol showing the order and timing of stimulus presentation for the two experimental conditions (lower panel).

Figure 2. The fNIRS headgear and channel layout. An infant wearing the headgear with the location of the 10-20 coordinates F8, T4, and T6 overlaying the photo. T4 was located at the midpoint of the lower row of channels (the source optode between channel 4 and 7) (upper panel). The array design showing the location of the channels (dashed circle), sources (star), and detectors (full circle) (lower panel).

Figure 3. Mean proportions of looking times to the experimenter's face, hands (I condition) or object (JA and baseline condition) for each condition, proportional to the length of the trial. Error bars indicate 1 standard error of the mean ( $N = 28$ ). \*\*\*  $p < 0.001$ .

Figure 4. Mean frequencies of four different social behaviors (gaze shifting, vocalizing, smiling, and pointing) per trial for each condition. Error bars indicate 1 standard error of the mean ( $N = 28$ ). \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Figure 5. An overview of the significant group effect for HbO<sub>2</sub> (maximum increase in HbO<sub>2</sub> concentration) highlighted with red circles (FDR-corrected  $p < 0.05$ ) and orange circles (uncorrected  $p < 0.05$ ) for the two contrasts: joint attention condition vs. baseline (A), interaction condition vs. baseline (B).

Figure 6. The grand averaged group time courses of the hemodynamic responses in HbO<sub>2</sub> and HHb to the JA and I conditions. Light greyed area indicates the time window where we found a significant increase of HbO<sub>2</sub> compared to baseline for both the JA and I conditions. The dark

greyed area indicates the time window where the difference in HbO<sub>2</sub> change was significant between the JA and I conditions.

Figure 7. Individual infant responses ( $N = 19$ ) of the association between looking time to the face in the I condition and HbO<sub>2</sub> responses in channel 8 (the STS-TPJ region) in the JA condition (left panel) and I condition (right panel).

Figure 8. The grand averaged group time courses of the hemodynamic responses in HbO<sub>2</sub> and HHb to the JA and I conditions for channels 1 to 12.

## References

- [1] M. Rutter, "Developmental catch-up, and deficit, following adoption after severe global early privation," *Journal of Child Psychology and Psychiatry*, 39(4), 465-476 (1998).
- [2] C. A. Nelson, C. H. Zeanah, N. A. Fox *et al.*, "Cognitive recovery in socially deprived young children: The Bucharest early intervention project," *Science*, 318(5858), 1937-1940 (2007).
- [3] H. T. Chugani, M. E. Behen, O. Muzik *et al.*, "Local brain functional activity following early deprivation: A study of postinstitutionalized Romanian orphans," *Neuroimage*, 14(6), 1290-301 (2001).
- [4] T. J. Eluvathingal, H. T. Chugani, M. E. Behen *et al.*, "Abnormal brain connectivity in children after early severe socioemotional deprivation: A diffusion tensor imaging study," *Pediatrics*, 117(6), 2093-2100 (2006).
- [5] G. Csibra, "Recognizing communicative intentions in infancy," *Mind and Language*, 25, 141-168 (2010).
- [6] G. Csibra, and G. Gergely, "Natural pedagogy," *Trends in Cognitive Sciences*, 13, 148-153 (2009).
- [7] H. Papoušek, and M. Papoušek, "Intuitive parenting: Aspects related to educational psychology," *European Journal of Psychology of Education*, 4(2), 201-210 (1989).
- [8] M. H. Bornstein, C. S. Tamis-LeMonda, H. Chun-Shin *et al.*,



- “Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct,” *Developmental Psychology*, 44(3), 867-874 (2008).
- [9] T. Farroni, G. Csibra, F. Simion *et al.*, “Eye contact detection in humans from birth,” *Proceedings of the National Academy of Sciences of the United States of America*, 99, 9602-5 (2002).
- [10] J. Nadel, I. Carchon, C. Kervella *et al.*, “Expectancies for social contingency in 2-month-olds,” *Developmental Science*, 2(2), 164 (1999).
- [11] R. P. Cooper, and R. N. Aslin, “Preference for infant-directed speech in the first month after birth,” *Child Development*, 61(5), 1584-1595 (1990).
- [12] N. Altvater-Mackensen, and T. Grossmann, “Learning to match auditory and visual speech cues: Social influences on acquisition of phonological categories,” *Child Development*, 86(2), 362-378 (2015).
- [13] H.-M. Liu, P. K. Kuhl, and F.-M. Tsao, “An association between mothers’ speech clarity and infants’ speech discrimination skills,” *Developmental Science*, 6(3), 1-10 (2003).
- [14] J. Gros-Louis, M. J. West, and A. P. King, “Maternal responsiveness and the development of directed vocalizing in social interactions,” *Infancy*, 19(4), 385-408 (2014).
- [15] R. Wu, K. S. Tummeltshammer, T. Gliga *et al.*, “Ostensive signals support learning from novel attention cues during infancy,” *Frontiers in Psychology*, 5, 251 (2014).
- [16] A. Seidl, R. Tincoff, C. Baker *et al.*, “Why the body comes first: Effects of experimenter touch on infants’ word finding,” *Developmental Science*, 18(1), 155-164 (2015).
- [17] R. Barr, P. Muentener, and A. Garcia, “Age-related changes in deferred imitation from television by 6- to 18-month-olds,” *Developmental Science*, 10(6), 910-921 (2007).
- [18] S. M. Pruden, K. Hirsh-Pasek, R. M. Golinkoff *et al.*, “The birth of words: Ten-month-olds learn words through perceptual salience,”

- Child Development, 77(2), 266-280 (2006).
- [19] M. Krcmar, "Can social meaningfulness and repeat exposure help infants and toddlers overcome the video deficit?," *Media Psychology*, 13(1), 31-53 (2010).
- [20] L. Brothers, "The social brain: A project for integrating primate behavior and neurophysiology in a new domain," *Concepts in Neuroscience*, 1, 27-51 (1990).
- [21] R. Adolphs, "Cognitive neuroscience of human social behaviour," *Nature Reviews Neuroscience*, 4(3), 165-178 (2003).
- [22] R. Adolphs, "The social brain: Neural basis of social knowledge," *Annual Review of Psychology*, 60, 693-716 (2009).
- [23] S. J. Carrington, and A. J. Bailey, "Are there theory of mind regions in the brain? A review of the neuroimaging literature," *Human Brain Mapping*, 30(8), 2313-2335 (2009).
- [24] T. Grossmann, and M. H. Johnson, "The development of the social brain in human infancy.," *European Journal of Neuroscience*, 25, 909-919 (2007).
- [25] M. H. Johnson, R. Griffin, G. Csibra *et al.*, "The emergence of the social brain network: Evidence from typical and atypical development.," *Development and Psychopathology*, 17, 509-619 (2005).
- [26] M. de Haan, and C. A. Nelson, "Recognition of the mother's face by six-month-old infants: A neurobehavioral study," *Child Development*, 68(2), 187-210 (1997).
- [27] M. de Haan, O. Pascalis, and M. H. Johnson, "Specialization of neural mechanisms underlying face recognition in human infants," *Journal of Cognitive Neuroscience*, 14(2), 199-209 (2002).
- [28] H. Halit, M. de Haan, and M. H. Johnson, "Cortical specialisation for face processing: Face-sensitive event-related potential components in 3- and 12-month-old infants," *Neuroimage*, 19(3), 1180-93 (2003).
- [29] D. Stahl, E. Parise, S. Hoehl *et al.*, "Eye contact and emotional face processing in 6-month-old infants: Advanced statistical methods applied to event-related potentials," *Brain & Development*, 32(4),

- 305-317 (2010).
- [30] M. Hirai, and K. Hiraki, "An event-related potentials study of biological motion perception in human infants," *Cognitive Brain Research*, 22(2), 301-304 (2005).
  - [31] S. Lloyd-Fox, A. Blasi, N. Everdell *et al.*, "Selective cortical mapping of biological motion processing in young infants.," *Journal of Cognitive Neuroscience*, 23, 2521-2532 (2011).
  - [32] V. M. Reid, S. Hoehl, and T. Striano, "The perception of biological motion by infants: An event-related potential study," *Neuroscience Letters*, 395(3), 211-214 (2006).
  - [33] S. Hoehl, V. Reid, J. Mooney *et al.*, "What are you looking at? Infants' neural processing of an adult's object-directed eye gaze.," *Developmental Science*, 11, 10-16 (2008).
  - [34] T. Striano, V. M. Reid, and S. Hoehl, "Neural mechanisms of joint attention in infancy.," *European Journal of Neuroscience*, 23, 2819-2823 (2006).
  - [35] T. Grossmann, E. Parise, and A. D. Friederici, "The detection of communicative signals directed at the self in infant prefrontal cortex," *Frontiers in Human Neuroscience*, 4, (2010).
  - [36] M. Imafuku, Y. Hakuno, M. Uchida-Ota *et al.*, "'Mom called me!'" Behavioral and prefrontal responses of infants to self-names spoken by their mothers.," *NeuroImage*, 103, 476-484 (2014).
  - [37] S. Lloyd-Fox, A. Blasi, A. Volein *et al.*, "Social perception in infancy: A near infrared spectroscopy study," *Child Development*, 80, 986-999 (2009).
  - [38] Y. Minagawa-Kawai, S. Matsuoka, I. Dan *et al.*, "Prefrontal activation associated with social attachment: Facial-emotion recognition in mothers and infants.," *Cerebral Cortex*, 19, 284-292 (2009).
  - [39] Y. Otsuka, E. Nakato, S. Kanazawa *et al.*, "Neural activation to upright and inverted faces in infants measured by near infrared spectroscopy.," *NeuroImage*, 34, (2007).
  - [40] Y. Saito, T. Kondo, S. Aoyama *et al.*, "The function of the frontal lobe in neonates for response to a prosodic voice," *Early Human Development*,

- 83(4), 225-230 (2007).
- [41] S. Urakawa, K. Takamoto, A. Ishikawa *et al.*, “Selective medial prefrontal cortex responses during live mutual gaze interactions in human infants: An fNIRS study,” *Brain Topography*, 1-11 (2014).
  - [42] S. Lloyd-Fox, B. Széplaki-Köllöd, J. Yin *et al.*, “Are you talking to me? Neural activations in 6-month-old infants in response to being addressed during natural interactions,” *Cortex*, 70, 35-48 (2015).
  - [43] M. L. Filippetti, S. Lloyd-Fox, M. R. Longo *et al.*, “Neural mechanisms of body awareness in infants.,” *Cerebral Cortex*, (2014).
  - [44] T. Grossmann, and M. H. Johnson, “Selective prefrontal cortex responses to joint attention in early infancy,” *Biology Letters*, 6, 540-543 (2010).
  - [45] S. Lloyd-Fox, J. E. Richards, A. Blasi *et al.*, “Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants,” *Neurophotonics*, 1(2), 025006 (2014).
  - [46] T. Grossmann, R. Oberecker, S. P. Koch *et al.*, “The developmental origins of voice processing in the human brain,” *Neuron*, 65(6), 852-858 (2010).
  - [47] Y. Minagawa-Kawai, H. van der Lely, F. Ramus *et al.*, “Optical brain imaging reveals general auditory and language-specific processing in early infant development,” *Cerebral Cortex*, 21(2), 254-61 (2011).
  - [48] S. Lloyd-Fox, K. Begus, D. Halliday *et al.*, “Cortical specialisation to social stimuli from the first days to the second year of life: A rural Gambian cohort,” *Developmental Cognitive Neuroscience*, (in press).
  - [49] D. R. Anderson, and T. A. Pempek, “Television and very young children,” *American Behavioral Scientist*, 48(5), 505-522 (2005).
  - [50] R. Barr, “Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice,” *Developmental Review*, 30(2), 128-154 (2010).
  - [51] P. K. Kuhl, F. M. Tsao, and H. M. Liu, “Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning,” *Proceedings of the National Academy of Sciences of the United States of America*, 100, 9096-9101 (2003).

- [52] S. Shimada, and K. Hiraki, "Infant's brain responses to live and televised action," *NeuroImage*, 32, 930-939 (2006).
- [53] L. J. Carver, A. N. Meltzoff, and G. Dawson, "Event-related potential (ERP) indices of infants' recognition of familiar and unfamiliar objects in two and three dimensions.," *Developmental Science*, 9, 51-62 (2006).
- [54] T. Charman, S. Baron-Cohen, J. Swettenham *et al.*, "Testing joint attention, imitation, and play as infancy precursors to language and theory of mind," *Cognitive Development*, 15(4), 481-498 (2000).
- [55] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, 20, 37-46 (1960).
- [56] J. R. Landis, and G. G. Koch, "An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers," *Biometrics*, 33(2), 363-374 (1977).
- [57] N. L. Everdell, A. P. Gibson, I. D. C. Tullis *et al.*, "A frequency multiplexed near-infrared topography system for imaging functional activation in the brain.," *Review of Scientific Instruments*, 76, 093705 (2005).
- [58] H. Obrig, and A. Villringer, "Beyond the visible—imaging the human brain with light," *Journal of Cerebral Blood Flow and Metabolism*, 23(1), 1-18 (2003).
- [59] S. Lloyd-Fox, A. Blasi, and C. E. Elwell, "Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy," *Neuroscience and Biobehavioral Reviews*, 34, 269-84 (2010).
- [60] A. Duncan, J. H. Meek, M. Clemence *et al.*, "Optical pathlength measurements on adult head, calf and forearm and the head of the newborn infant using phase resolved optical spectroscopy.," *Physics in Medicine and Biology*, 40, (1995).
- [61] Y. Benjamini, "Discovering the false discovery rate," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(4), 405-416 (2010).
- [62] P. Mundy, and L. Newell, "Attention, joint attention, and social

- cognition.,” *Current Directions in Psychological Science*, 16, 269-274 (2007).
- [63] M. Tomasello, M. Carpenter, J. Call *et al.*, “Understanding and sharing intentions: The origins of cultural cognition,” *Behavioral and Brain Sciences*, 28(5), 675-735 (2005).
- [64] G. Butterworth, and N. Jarrett, “What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy,” *British Journal of Developmental Psychology*, 9(1), 55-72 (1991).
- [65] P. Belin, and M.-H. Grosbras, “Before speech: Cerebral voice processing in infants,” *Neuron*, 65(6), 733-735 (2010).
- [66] P. Belin, R. J. Zatorre, P. Lafaille *et al.*, “Voice-selective areas in human auditory cortex,” *Nature*, 403(6767), 309-312 (2000).
- [67] A. Bethmann, and A. Brechmann, “On the definition and interpretation of voice selective activation in the temporal cortex,” *Frontiers in Human Neuroscience*, 8, 499 (2014).
- [68] S. Lloyd-Fox, A. Blasi, E. Mercure *et al.*, “The emergence of cerebral specialization for the human voice over the first months of life,” *Social Neuroscience*, 7(3), 317-330 (2012).
- [69] A. Blasi, E. Mercure, S. Lloyd-Fox *et al.*, “Early specialization for voice and emotion processing in the infant brain,” *Current Biology*, 21(14), 1220-1224 (2011).
- [70] A. Cristia, E. Dupoux, Y. Hakuno *et al.*, “An online database of infant functional near infrared spectroscopy studies: A community-augmented systematic review,” *PLoS ONE*, 8(3), 1-10 (2013).
- [71] M. Stets, D. Stahl, and V. M. Reid, “A meta-analysis investigating factors underlying attrition rates in infant ERP studies.,” *Developmental Neuropsychology*, 37(3), 226-252 (2012).