

Functional imaging of the developing brain with wearable high-density diffuse optical tomography: A new benchmark for infant neuroimaging outside the scanner environment



Elisabetta Maria Frijia^{a,b,*}, Addison Billing^{a,c}, Sarah Lloyd-Fox^{d,e}, Ernesto Vidal Rosas^a, Liam Collins-Jones^{a,b}, Maria Magdalena Crespo-Llado^{d,e}, Marta Perapoch Amadó^{e,f}, Topun Austin^b, Andrea Edwards^b, Luke Dunne^g, Greg Smith^g, Reuben Nixon-Hill^g, Samuel Powell^g, Nicholas L. Everdell^g, Robert J. Cooper^{a,b}

^a DOT-HUB, Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom

^b neoLAB, Cambridge University NHS Foundation Trust, Cambridge CB2 0QQ, United Kingdom

^c Institute of Cognitive Neuroscience, University College London, London WC1N 3AZ, United Kingdom

^d Cambridge BabyLab, Department of Psychology, University of Cambridge, Cambridge, CB2 3ER, United Kingdom

^e Centre for Brain and Cognitive Development, Birkbeck, University of London, London WC1E 7HX, United Kingdom

^f Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom

^g Gowerlabs Ltd., Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom

ARTICLE INFO

Keywords:

High-Density Diffuse Optical Tomography
Functional Near-Infrared Spectroscopy
Optical Neuroimaging
Infant Cognitive Development
Infant Neuroimaging

ABSTRACT

Studies of cortical function in the awake infant are extremely challenging to undertake with traditional neuroimaging approaches. Partly in response to this challenge, functional near-infrared spectroscopy (fNIRS) has become increasingly common in developmental neuroscience, but has significant limitations including resolution, spatial specificity and ergonomics. In adults, high-density arrays of near-infrared sources and detectors have recently been shown to yield dramatic improvements in spatial resolution and specificity when compared to typical fNIRS approaches. However, most existing fNIRS devices only permit the acquisition of ~20–100 sparsely distributed fNIRS channels, and increasing the number of optodes presents significant mechanical challenges, particularly for infant applications. A new generation of wearable, modular, high-density diffuse optical tomography (HD-DOT) technologies has recently emerged that overcomes many of the limitations of traditional, fibre-based and low-density fNIRS measurements. Driven by the development of this new technology, we have undertaken the first study of the infant brain using wearable HD-DOT. Using a well-established social stimulus paradigm, and combining this new imaging technology with advances in cap design and spatial registration, we show that it is now possible to obtain high-quality, functional images of the infant brain with minimal constraints on either the environment or on the infant participants. Our results are consistent with prior low-density fNIRS measures based on similar paradigms, but demonstrate superior spatial localization, improved depth specificity, higher SNR and a dramatic improvement in the consistency of the responses across participants. Our data retention rates also demonstrate that this new generation of wearable technology is well tolerated by the infant population.

1. Introduction

The investigation of cortical responses to controlled external stimuli has been fundamental to our improved understanding of the functional architecture of the human brain over the last few decades (Corbetta and Shulman, 2002; Kanwisher, 2010; Allison et al., 1999). Among the many available neuroimaging techniques, it is functional magnetic resonance imaging (fMRI) that has been at the heart of these advancements (Buckner et al., 1996; Culham et al., 2003; Pelphrey et al., 2005). The

blood-oxygen-level-dependant (BOLD) signal to which fMRI is sensitive is a result of functional hyperaemia: a localized increase in cerebral blood flow that results in changes in the concentration of both oxyhaemoglobin and deoxyhaemoglobin that provide a proxy for neuronal activity.

While the advantages of fMRI are numerous, there are many applications for which fMRI is poorly suited. Functional neuroimaging of infancy is one such application. fMRI places many restrictions on both

* Corresponding author at: DOT-HUB, Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom.
E-mail address: elisabetta.frijia.18@ucl.ac.uk (E.M. Frijia).

the subject and the experimental paradigm. Because of the need to remain still, fMRI of infants is almost always conducted on sleeping or sedated subjects. This significantly limits the types of paradigm that can be undertaken and as a result, the number of fMRI studies of healthy infants and young children remains disproportionately low (Ellis and Turk-Browne, 2018). Functional near-infrared spectroscopy (fNIRS) is a non-invasive, non-ionizing, silent and low-cost neuroimaging technique that can accommodate a relatively high degree of movement from the subject (Quaresima et al., 2012) and can be performed in a typical lab environment on awake infant subjects. The standard fNIRS method uses arrays of optodes (sources and detectors of near-infrared light), sparsely distributed over the scalp to provide measures of haemodynamics in the underlying tissues. When compared with fMRI, fNIRS demonstrates higher temporal resolution, and has the advantage of being able to determine the concentration of both oxy- and deoxy-haemoglobin. However, the spatial resolution of typical fNIRS systems is on the order of 3 cm, significantly lower than fMRI (Scholkmann et al., 2014). Furthermore, the depth sensitivity of fNIRS is limited to the superficial cortex (Quaresima et al., 2012), and typical fNIRS layouts provide little (if any) depth information, meaning changes in haemoglobin concentrations in the brain can be difficult to distinguish from those in the extracerebral tissues (Gagnon et al., 2012; Funane et al., 2014).

Despite these limitations, the emergence of fNIRS has enabled the study of infant functional responses across a range of domains, including the visual (Meek et al., 1998), auditory (Sakatani et al., 1999), olfactory stimulation (Bartocci et al., 2000) and somatosensory (Verriotis et al., 2016). One particular area of interest has been the social brain network which develops rapidly in infancy (Lloyd-Fox et al., 2009; Frank et al., 2009; Minagawa-Kawai et al., 2009). Found (in part) in the superior temporal sulcus (STS), this network underpins the processing of social stimuli and social interaction. Specific patterns of activation have been identified in response to social stimulation of adults using fMRI (Pelphrey et al., 2005; Allison et al., 2000). To investigate the development of this network in the infant brain, a social stimulus paradigm was developed over several years by researchers at the Centre for Brain and Cognitive Development (CBCD, Birkbeck, University of London) (Lloyd-Fox et al., 2009; Lloyd-Fox et al., 2012). Numerous fNIRS studies across multiple centres have used the resulting paradigm to investigate the development of infant cognitive abilities and compare the results obtained between groups at different ages, in different countries, of different socio-economic status (SES), and within families with different likelihoods for developing disorders such as autism spectrum disorder (ASD) (Lloyd-Fox et al., 2017; Perdue et al., 2019; Lloyd-Fox et al., 2013).

Current fNIRS technologies can be broadly divided into two categories: benchtop and wearable devices. Benchtop devices provide a reasonable number of near-infrared sources and detectors, (typically ~20–60) (Scholkmann et al., 2014), but usually require the same number of optical fibre bundles to carry light to and from the scalp. The necessity of optical fibres presents significant mechanical challenges, particular for infant subjects. Both the weight and the physical size of the fibre bundle limits the number that can be comfortably applied to the infant scalp, and the weight of the optical fibres can render a system more vulnerable to detachments or displacements of the optodes when a subject moves, resulting in motion artifacts. The available wearable devices are much smaller in size and are rapidly improving (Scholkmann et al., 2014). However, many still include relatively large control units and/or significant cabling, and all current wearable devices remain limited in terms of channel count and/or field of view (Zhao and Cooper, 2017).

Diffuse optical tomography (DOT) is an evolution of fNIRS that permits the production of three-dimensional images of the optical properties of a target object (White and Culver, 2010). Channels with a range of source-detector separations and spatially-overlapping sensitivity distributions are essential to DOT approaches (White and Culver, 2010). This necessitates detector technology with a very high dynamic range, and denser arrangements of source and detector optodes than are possible with standard fNIRS devices.

Table 1
Age and head measurement information for all participants.

	Mean	SD	Range
Age (days)	168	30.6	131–216
Head Circumference (cm)	43.2	1.6	40–45.5
Pre-auricular points via Fpz (cm)	21.5	1.8	18–23.5
Helix to helix via apex (cm)	24.9	1.2	23–26.5
Tragus to tragus via apex (cm)	28.5	1.7	26–31.5
Nasion-to-inion (cm)	27.9	2.2	25.5–31.5

The concept of high-density diffuse optical tomography (HD-DOT) takes this one step further. In HD-DOT, an array of sources and detectors is employed that is dense enough to provide a continuous distribution of both ‘short separation’ channels at ~10 mm source-detector separation and longer channels spanning the 10–40 mm range (White and Culver, 2010). This approach has been shown to significantly improve the disentanglement of extracerebral and cerebral haemodynamics (Funane et al., 2014), and has the capacity to yield cortical activation maps that approach the resolution of fMRI (Eggebrecht et al., 2014). However, because these HD-DOT technologies require very large numbers of optodes, they compound the mechanical challenge associated with studying infant populations.

In the last few years, the miniaturization of the optoelectronics required for DOT has accelerated significantly (Zhao et al., 2019). The first functional images of the human brain obtained using a fibre-less HD-DOT system were reported in late 2016 (Chitnis et al., 2016). This system demonstrated a novel modular design architecture that provided a dense configuration of sources and detectors while still allowing the system to conform to the scalp. This modular approach has since been advanced (Zhao et al., 2019) and adopted by other research groups (Wyser et al., 2017; Zimmermann et al., 2019; Von Lüthmann et al., 2020). In this paper, we describe the application of a prototype, wearable technology that permitted the acquisition of what we believe to be the first high-density DOT data-set in infants of 4 - 7 months of age, one of the most frequently studied age groups in developmental science (Cristia et al., 2013). We used the CBCD social stimulus paradigm that has been applied extensively to infants by multiple groups using a range of devices to allow direct comparison of the resulting data. In doing so we sought to demonstrate what can now be achieved with optical neuroimaging in the developmental neurosciences.

2. Materials and methods

2.1. Participants

Twenty healthy, term-born infants between 4 and 7 months of age were recruited to this study (6 females and 14 males), which took place at the Cambridge Babylab. Written, informed parental consent was obtained for each participant prior to their participation. All aspects of the protocol were approved by the NHS London Riverside Research Ethics Committee (IRAS ID 241,042). Demographic information for the cohort is provided in Table 1.

2.2. HD-DOT system

We employed a prototype, modular, high-density DOT system known as “LUMO”, developed by UCL spin-out Gowerlabs Ltd., UK. The system consists of multiple independent modules (or “tiles”, Fig. 1) that together create a dense network of sources and detectors, while still allowing the system to conform to the scalp. Each hexagonal sensor tile is equipped with 3 dual-wavelength LED sources (at 735, 850 nm) and 4 photodiode detectors. Each module weighs ~6 g and measures 29 mm across at the widest point. Within-tile measurements are obtained with source-detector separations of approximately 10 mm and 20 mm. Cross-tile measurements are acquired for all separations within an array but

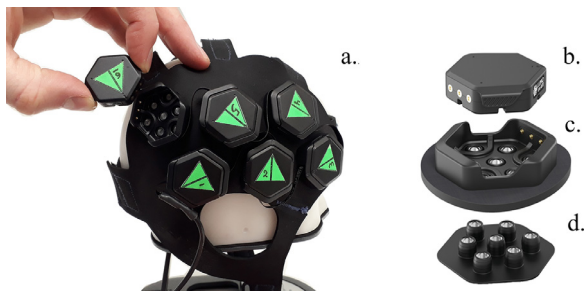


Fig. 1. The cap containing 12 LUMO tiles in a 2×6 tile array, arranged to provide sensitivity to the superior temporal lobes and surrounding areas is shown in (a). Note the custom, adjustable neoprene cap that hosts the docks (c) into which the hexagonal tiles (b) are located. The light-guide piece shown in (d) contains seven short optical fibres with length 4.5 mm. Note the green high-contrast triangular labels, which supported photogrammetric registration.

channels up to ~ 45 mm are expected to yield acceptable signal quality. The modules are located in a chain of “docks” that are clipped into a flexible neoprene cap. The docks provide power and data transfer, and can be positioned anywhere within the cap. The dock chain connects via a single flexible cable to a “hub”, which itself connects via USB to a laptop PC. A disposable, flat, rubber “light-guide” containing seven short plastic optical fibres couples light through the dock to and from the scalp.

The neoprene cap itself was designed specifically for this study to provide good optical coupling, efficient fitting (< 5 min) and to be comfortable for the infant. To obtain good stability and avoid displacements due to motion, the cap was made of neoprene with a high-friction, rubberized surface. The cap consisted of two halves connected by multiple Velcro™ straps to allow it to be rapidly adjusted to fit a range of different head sizes and shapes.

For this study we designed a layout of 12 docks; 6 per hemisphere. Each set of 6 tiles provides 18 sources, 24 detectors and a total of 432 logical channels per wavelength. Approximately 210 channels per hemisphere were expected to provide a source-detector separation below 45 mm. Optode locations and the associated channels are illustrated in Fig. 2. This array design represents the equivalent of 84 optical fibres. The whole array, including the neoprene cap, docks and tiles, weighs ~ 195 g. The frame rate of this prototype device was 4.6 Hz. The whole HD-DOT system (cap, tiles, hub, cabling and laptop) fitted in a small backpack and was carried to the site to perform each study.

2.3. Cap positioning and spatial registration

The infant’s head circumference was measured around the crown and distances between the pre-auricular points over the apex and via Fpz were recorded. After measurement, the cap was adjusted and placed on the infant’s head, such that the two arrays of tiles were located over the superior temporal lobes and the temporoparietal junction (Lloyd-Fox et al., 2009). We first aligned the array to the ear locations using specific markings on the cap, ensuring that it was sitting symmetrically relative to the face and that the front edge followed the curve of the eyebrows. The posterior part of the array covered the scalp locations T3–T5 and T4–T6, with anterior parts of the array placed toward F7 and F8 in the EEG 10–20 coordinate system. The cap design, positioning and fit are shown in Fig. 2.

As part of the development of this study, we established a new, infant-friendly method for acquiring the three-dimensional position of sensors on the scalp relative to the cranial landmarks. For each participant, we acquired a three-dimensional model of the head surface using the TrueDepth camera functionality of an X-series iPhone (Apple, Inc., CA, USA) and the app Scandy Pro (Scandy LLC, LA, USA), which together allow scaled, 3D point clouds to be acquired and exported in a readily

accessible format. As the inevitable movement of the infant makes acquiring a continuous 360-degree scan nearly impossible, we acquired several depth-resolved ‘snapshot’ scans from different angles for each participant while the cap was in place. Each snapshot scan lasted approximately 1–2 s, and was repeated if the baby moved during acquisition. The angle of each snapshot was chosen to ensure at least two LUMO tiles and one cranial landmark were in the frame in each case and together the snapshots would cover both arrays, the nasion, inion, pre-auricular points and the vertex (Cz). The resulting multiple partial point clouds (see example shown in Fig. 2(b)) were then registered to one another using the software package CloudCompare (www.danielgm.net/cc), which uses equivalent point pairs that are manually identified across different partial scans to rigidly transform those partial point clouds to produce a complete model (Fig. 2(c)). To identify the exact location of sources and detectors, each tile was tagged with a specially designed fluorescent green triangular marker (Figs. 1(a), 2(a–c)), the points of which were positioned directly above the sources of each tile. Because the tile and docks are of known dimensions, identifying the locations of the points of each triangle (Fig. 2(c) provides sufficient information to determine the location of each optode location on the scalp without approximation. The location of subject-specific cranial landmarks (nasion, inion, pre-auricular points, Cz) and the position of each source and detector on the scalp could then be determined from the completed point cloud. These positions allowed a multi-layer mesh model of the six-month-old infant head to be registered to each infant’s landmark space to create a subject-specific structural model (Fig. 2(d)).

2.4. Paradigm

The experiment required infants to be seated on their parent’s lap, and observe audio-video recordings displayed on a 46-inch plasma screen monitor positioned at a distance of approximately 70 cm. The stimulus presentation was controlled, using MATLAB®, by a second experimenter from a desktop PC out of view of the infant (see Fig. 3), which also recorded the output of a camera embedded in the monitor.

The paradigm alternated between stimulation and baseline blocks. The visual component of the stimulation blocks consisted of full-colour, life-size videos of female actors moving their eyes, silently mouthing English vowels, or performing hand gestures including ‘peek-a-boo’ and ‘incy wincy spider’. These were presented for a period of 9–12 s. The baseline condition consisted of full-colour, still images of different types of vehicle, each presented in a random order for a pseudorandom duration of 1–3 s to create blocks of 9–12 s in duration. Each stimulation block was presented in combination with one of three possible auditory conditions: silence (S), vocal (consisting of human non-speech vocal sounds, V), and non-vocal (consisting of environmental sounds, N). The paradigm ended after a maximum of 18 trials (6 trials per condition) or after the infant became bored and fussy. The full experiment lasted approximately 7 min and the order of presentation of the stimuli was the same across all infants (Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2013).

During the study, and in line with previous applications of this paradigm, the parent was instructed not to interact with the infant. In order to prevent infants from grabbing the cap, parents were told to gently hold the infant hands if needed. If infants became distracted during the stimulus presentation, alerting sounds were played to help draw the infant’s attention back to the screen (Lloyd-Fox et al., 2009).

2.5. Data pre-processing

Each experimental session was recorded by a video camera system embedded into the frame of the plasma screen monitor. The infants’ eye movements were coded manually offline to evaluate the total screen-directed looking time during each trial. For a trial to be considered valid, we required that the infant was looking at the screen for a minimum of

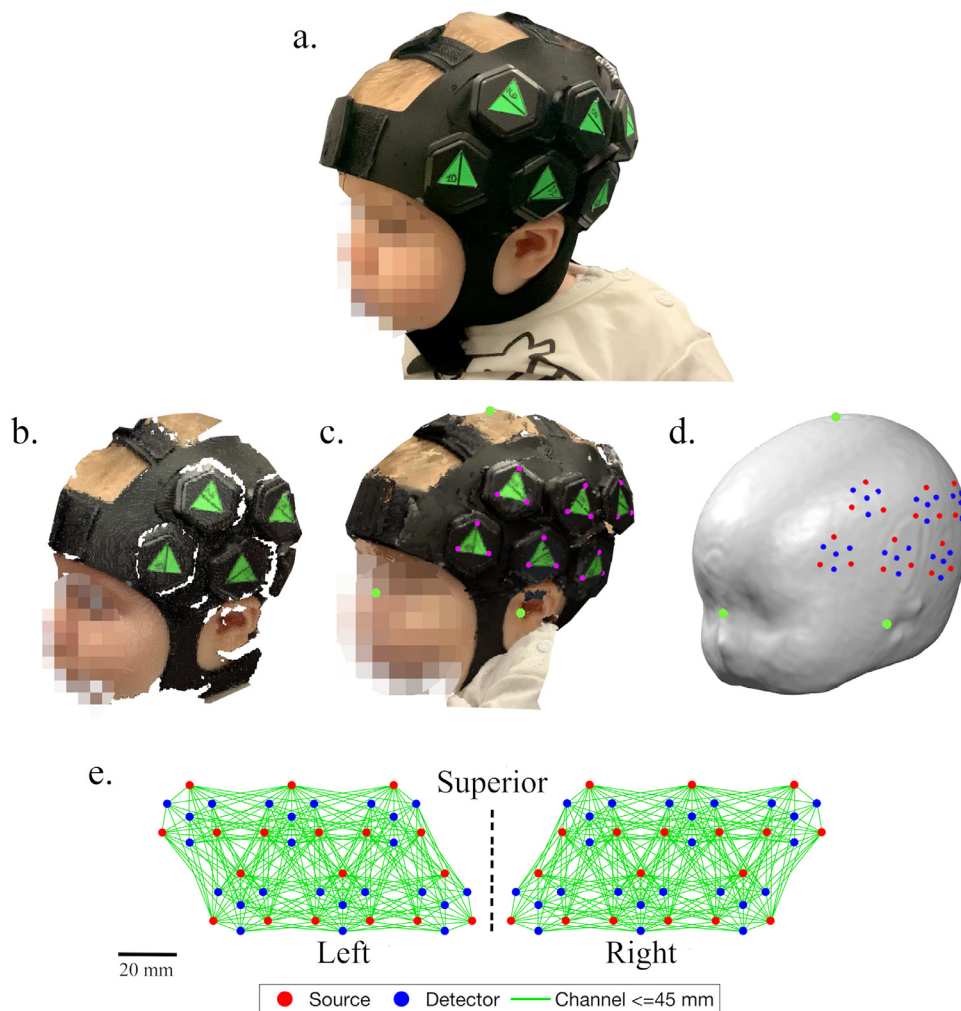


Fig. 2. (a) A side view photograph of an infant wearing the cap with tiles visible over the left hemisphere. Note the Velcro strap connections running over the midline of the head. (b) A single snapshot point cloud image of the same infant. (c) The complete model for that infant, created from multiple point cloud images with cranial landmarks (green points) and sources marker locations (magenta points). (d) The source positions (red points) and detector positions (blue points) shown on the subject-registered six-month anatomical atlas model. (e) A 2D representation of the full 12 tile array, showing channels with source-detector separation of approximately 45 mm or less.

60% of the experimental trial period (Lloyd-Fox et al., 2014). A minimum of three valid experimental trials for each condition were required to include an infant's data in our analysis (Lloyd-Fox et al., 2014).

To allow a more direct comparison with previous studies that have used this paradigm and fibre-based fNIRS devices, our HD-DOT data pre-processing followed established protocols for infant fNIRS research (Di Lorenzo et al., 2019). First, for simplicity and to reduce data size we rejected channels with a source-detector separation of > 60 mm. Channels were also rejected if their coefficient of variation exceeded 8% (Lloyd-Fox et al., 2009). Note that this rejection was performed based on an interval of data during which no motion artifacts were apparent on visual inspection. In order to identify the impact of motion artifacts, a channel-by-channel analysis was performed using functions from the Homer2 fNIRS processing package (www.homerfnirs.org, (Huppert et al., 2009)). The function *hmrMotionArtifactByChannel* was used with the parameters $tMotion = 1$, $tMask = 1$, $STDE-Vthresh = 15$, $AMPthresh = 0.4$ (Di Lorenzo et al., 2019). After motion detection, we performed a motion correction procedure that consisted of a spline (Scholkmann et al., 2010) and wavelet deconvolution method (Molavi and Dumont, 2012). The spline approach models periods of data classified as motion artifacts on a channel-by-channel basis with a cubic spline interpolation and then subtracts that model from the original signal. The function also corrects the signal baseline to ensure the continuity of the data. Informed by the value used previously in other studies we set the parameter $p = 0.99$ (Di Lorenzo et al., 2019). The wavelet method is based on a discrete wavelet transform decomposition of each channel data. Wavelet coefficients identified as outliers are assumed to

model artifact and are removed. To identify outliers, we used an interquartile range factor of 0.8, which has been shown to perform well in the recovery of the haemodynamic response function (HRF) from infant data (Di Lorenzo et al., 2019).

After motion correction, the data was band-pass filtered in range 0.01–0.5 Hz using a FIR digital filter to remove slow drifts and high frequency/cardiac noise. Channel-wise changes in haemoglobin concentration were obtained using the modified Beer-Lambert law (Delpy et al., 1988), with a differential pathlength factor (DPF = 5.1) appropriate for infants used for both wavelengths (Duncan et al., 1995). Finally, block-averaged relative changes in oxyhaemoglobin (HbO) and deoxyhaemoglobin (HbR), were computed for 25-sec-long epochs starting from 5 s before the onset of each trial for each of the 3 stimulation conditions (Di Lorenzo et al., 2019). The haemodynamic response at the group level was obtained by compiling the average time courses for each infant into an averaged response.

2.6. Head modelling and image reconstruction

We employed a four-layer mesh model of the 6-month infant head. The data used to produce this model (consisting of average T1-weighted MRI templates for head and brain, and tissue probability maps for grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF)) was derived from the NIHPD Objective-2 and MCBI-USC databases, and was obtained from the Neurodevelopmental MRI database (Sanchez et al., 2012; Richards et al., 2016; Almlí et al., 2007; Richards and Xie, 2015). A majority vote procedure was implemented using the tissue probability

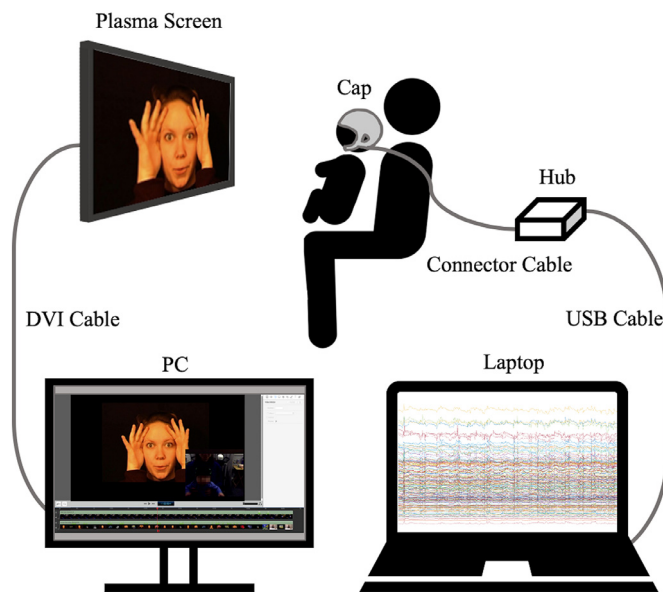


Fig. 3. A representation of the full experimental set-up showing all devices and the sitting positioning of parent and infant during the study. The experiment required infants to be seated on their parent's lap at approximately 70 cm from the screen. The desktop PC was used to control the stimuli, which were presented on a 46-inch plasma screen and to record the session using an integrated camera. The LUMO system includes a laptop for data acquisition, a hub, connected to the laptop through a USB cable and the cap, connected to the hub via a single 2.5 m cable.

maps to produce binary segmentations for brain tissues. An outer scalp boundary was segmented from the average T1-weighted image using Betsurf (Jenkinson et al., 2012). The resulting 4-layer tissue mask was used to create a finite element volume mesh, and scalp and grey matter surface meshes using the Iso2Mesh package (iso2mesh.sourceforge.net, (Fang and Boas, 2009)). Once this model was registered to the individual's cranial landmarks, we used TOAST++ (<http://toastplusplus.org>, (Schweiger and Arridge, 2014)) to model optical fluence and calculate a Jacobian matrix for each wavelength. A zeroth-order Tikhonov regularized reconstruction was performed using a regularization hyperparameter of 0.01. All images were reconstructed in the full mesh volume. For visualization purposes, image values at the level of the grey matter (GM) surface were also extracted. Note that all good short-separation channels were included in the reconstruction process, but no explicit regression of the short signals was performed. The code used to pre-process and reconstruct the HD-DOT data described here is now freely available as part of the DOT-HUB toolbox at www.github.com/DOT-HUB.

2.7. Statistical mapping

We performed a simple statistical mapping approach in the GM space to produce T-statistic maps that compared the group-level responses for all different conditions across all infants. A time window of between 10.5 and 15.5 s post-experimental stimulus onset was selected to represent the peak of the haemodynamic response to each of the stimulus blocks. For each condition and GM mesh node, a concatenated vector of all activation periods for all participants was created. These response vectors were then compared to equivalent vectors constructed for each of the other conditions (including the baseline condition) using a paired two-tailed *t*-test. The resulting values were corrected for multiple comparisons using the Bonferroni method on the basis of the number of nodes in the mesh, which represents an extremely conservative correction approach.

2.8. Low-density comparative analysis

To directly assess the impact of the increased channel density provided by our system, we repeated our data-processing and image reconstruction pipeline using a simulated low-density fNIRS array. This array was created by identifying a set of channels with source-detector separations in range 20 to 24.5 mm (Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2013) that covered a similar field-of-view as the full array, but included no overlapping measurements. Channel-space average HRFs, volume and grey-matter-surface images were created for each condition. To explicitly compare the localization and consistency of the low-density and high-density image results, a binary response map was created for each participant by thresholding each GM peak image (i.e. the mean of the period from 10.5–15.5 second post stimulus onset) at 25% of its own maximum. These binary response maps were then summed across all participants to create a response overlap map.

3. Results

3.1. Participant and trial rejection

Of the twenty infants recruited, zero were rejected from data collection due to 'fussing out' prior to the acquisition of sufficient trials. The average duration of the recording sessions was 6.40 min (range 5.65 to 6.65 min). The looking-time analysis resulted in the exclusion of one infant, who failed to attend to enough stimuli for the requisite duration to obtain the minimum required number of trials. Of the remaining infants, the median number of trials that yielded acceptable looking times was 18 (range 15–18) with a median looking time of 9.7 s (range 5.8–11.9 s). For condition types N, S, V the median (range) number of included trials were 6 (5–6), 6 (3–6) and 6 (5–6), respectively. The median (range) of looking times for accepted trials for conditions N, S, V were 9.9 s (6.2–11.9 s), 9.8 s (5.9–11.9 s), and 9.6 s (5.8–11.9 s). Only one trial was excluded in one of the participants because of motion artifacts identified by the automatic detection *after* performing motion artifact correction. Two participants were rejected due to technical problems during data acquisition, which in both cases were caused by user error. For one participant, it was not possible to collect depth-resolved registration images, and polhemus measurements taken on a 6-month head phantom were used instead. In total 17 viable datasets were obtained from a cohort of 20 recruited infants. The total set-up time (including seating the parent and baby, aligning and attaching the head cap and preparing for recording) did not exceed 15 min.

3.2. Data quality assessments

Fig. 4(a) shows the temporal mean of the intensity detected for all channels at both wavelengths as a function of source-detector separation for all participants. These measurements exhibit a decay with a characteristic log-linear fall-off. Channels that passed the coefficient-of-variation threshold and did not exhibit saturation are displayed in red. The remaining channels are shown in blue. Fig. 4(b) shows a histogram of the average available dual-wavelength channels (i.e. fNIRS channels) and the average accepted channels across all participants as a function of source-detector separation. The average number of good channels obtained per participant across the group was 430.9 (range 343–506). The maximum source-detector separation from which data was retained for each participant ranged from 41.8 to 59.2 mm in one extreme case. Note that the small number of short-separation channels lost were removed due to saturation effects that occurred in one early participant that was studied prior to the introduction of an automated source-power optimization software routine.

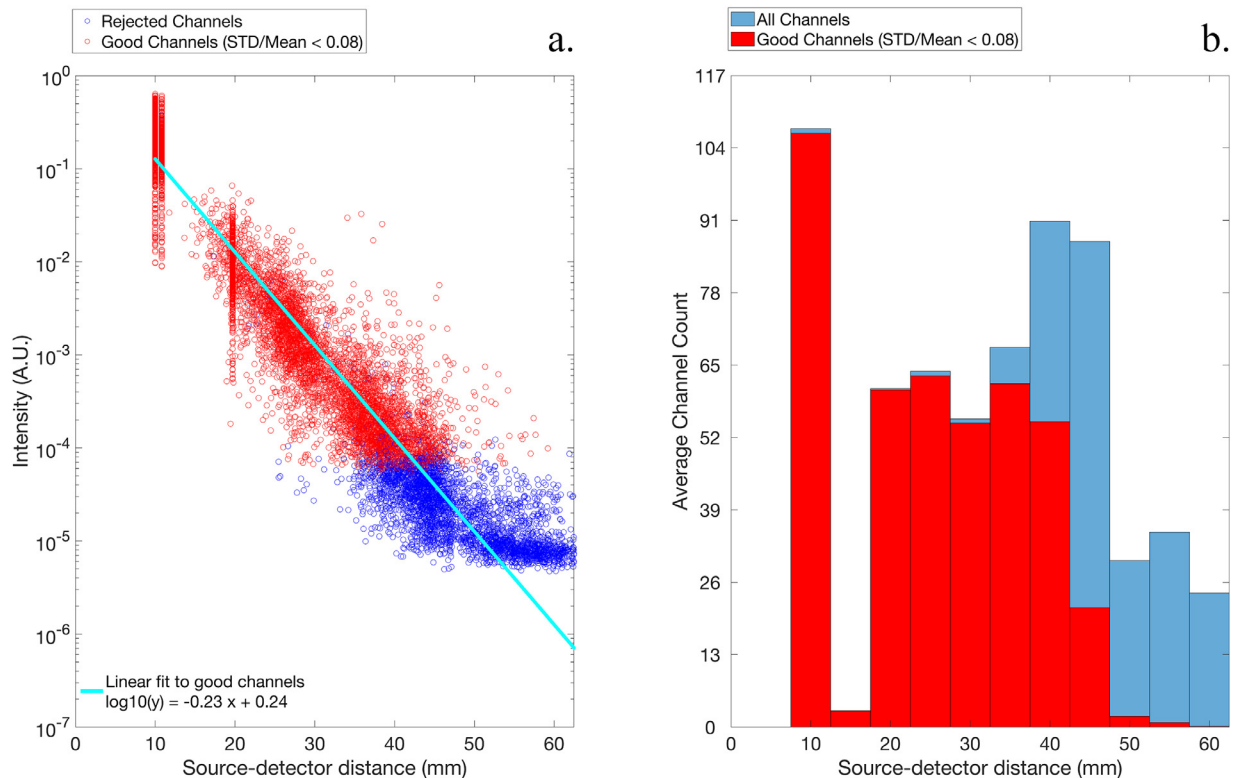


Fig. 4. (a) Mean intensity over time for every source-detector pair, wavelength and participant on a logarithmic scale, plotted against source-detector separation. (b) A histogram illustrating the average channel count across all participants as a function of source-detector distance.

3.3. Channel-wise haemodynamic responses

Fig. 5 shows the response at the group level for the vocal condition displayed on a representative 2D channel layout. Clear, well-localized, canonical haemodynamic responses consisting of an increase in HbO and a decrease in HbR are clearly evident. The largest responses occur in channels in the posterior and inferior part of the array. Activation is also apparent, but to a lesser extent, in the anterior part of the array (see inset Fig. 5(a)). The non-vocal condition yielded responses with similar gross spatial localization (Fig. S1, online supporting information). The silence condition did not result in such clear haemodynamic responses relative to the preceding 5 second baseline period in the channel space. The scale of the observed changes in response to the silence condition were noticeably smaller than those of the vocal and non-vocal conditions, and channels display a mixture of increases and decreases in HbO and HbR (see Fig. S2). The latency to peak of the maximum change in HbO in the channels that exhibit a clear functional response was approximately 12–13 s post-stimulus onset for the vocal (and non-vocal) conditions. Note the inset Figs. 5(a,b) show zoomed-in detail of two example group-level HRFs for a channel located on the left hemisphere over the anterior part of the array and a channel located on the right hemisphere over the posterior part of the array (both with a source-detector separation of ~ 27 mm). Note that these same channels are also highlighted for the non-vocal and silence conditions in supplementary Figs. S1 and S2.

Fig. 6 shows a subset of the group level haemodynamic responses obtained for the vocal condition arranged by source-detector distance. These responses were identified by finding the channel with the largest change in HbO during the peak time window, and then identifying a channel of each given source-detector separation that was centered within 10 mm of that peak channel.

3.4. Volume and grey matter image results

The high-density sampling provided by the LUMO device enabled 3D images to be reconstructed in the four-layer tissue mesh. Fig. 7 shows a coronal view, a transverse view and the grey-matter surface extraction of the reconstructed DOT images of HbO and HbR at the group level. These images were obtained by averaging the time window of between 10.5 and 15.5 s post-stimulus onset. The slices selected for the coronal and transverse views correspond to the coordinates of the node exhibiting the largest increase in HbO and decrease in HbR for each condition. For the vocal and non-vocal conditions, the peak of the HbO and HbR changes both occur in the posterior superior temporal gyrus. Note the co-incident increases in HbO and decreases in HbR apparent for both vocal and non-vocal conditions. This classical response is much less apparent in the silence v baseline map, which exhibits a more variable response. The silence condition shows a coincident increase in HbO and decrease in HbR in only the most posterior region of the field-of-view, at the boundary between the temporal and occipital lobes, and particularly in the left hemisphere. A range of other non-typical changes relative to baseline are apparent, including a decrease in HbO and increase in HbR in the anterior superior temporal gyri bordering the lateral fissures.

3.5. Low-density and high-density array comparisons

Fig. 8 provides a summary of the results of our comparison between low-density and high-density array results. Fig. 8(a) shows the simulated low-density array, while Fig. 8(b) shows the equivalent view for the full array. Figs. 8(c–f) each show a transverse view of the volume image (i) and the grey-matter surface extraction (ii) of the reconstructed HbO and HbR images for the non-vocal condition for an example participant using the low-density and the high-density arrays respectively.

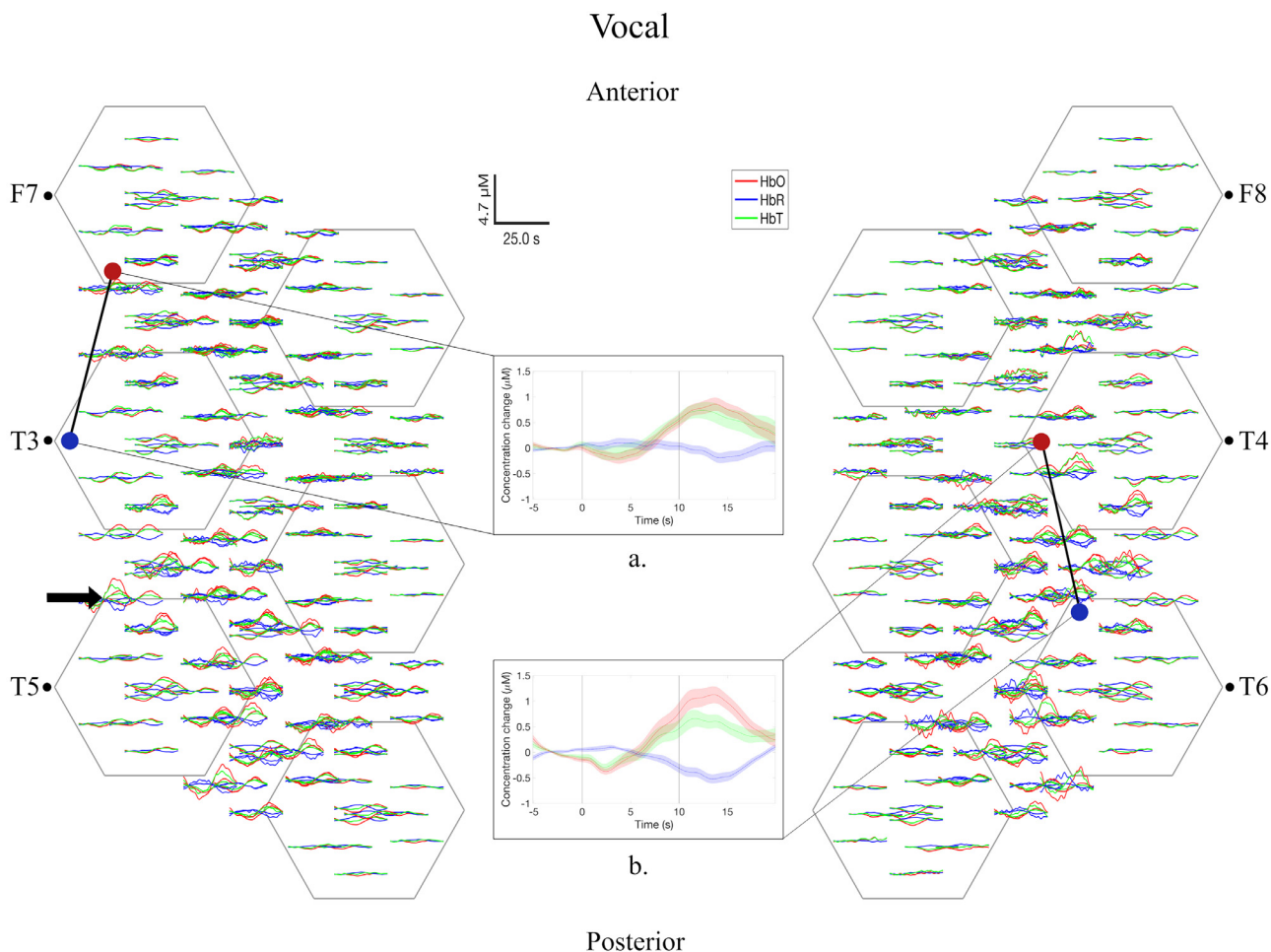


Fig. 5. Haemodynamic responses in a representative 2D array, obtained at the group level for the vocal condition. Note that a high number of channels exhibit clear haemodynamic responses. Selected HRFs for a channel located in the anterior part of the array (a) and a channel located over the posterior part of the array (b) are highlighted. The shaded areas about each line in the inset plots represent standard error across all participants. Note that the arrow shows the peak channel (that exhibiting the largest change in HbO during the peak response window).

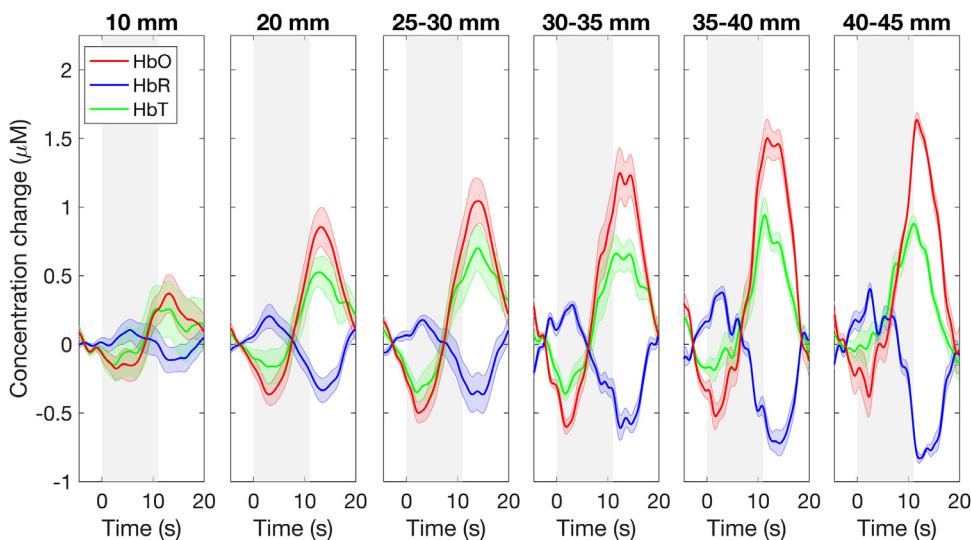


Fig. 6. A subset of the group level haemodynamic responses for the vocal condition ordered by source-detector distance. Each curve depicts the mean across subjects, with the shaded areas representing the standard error on the mean. Note the clear increase in the haemodynamic response scale as a function of source-detector separation with the largest HbO change in the 40–45 mm channel range. Note also the relatively small haemodynamic change present in the 10 mm channels.

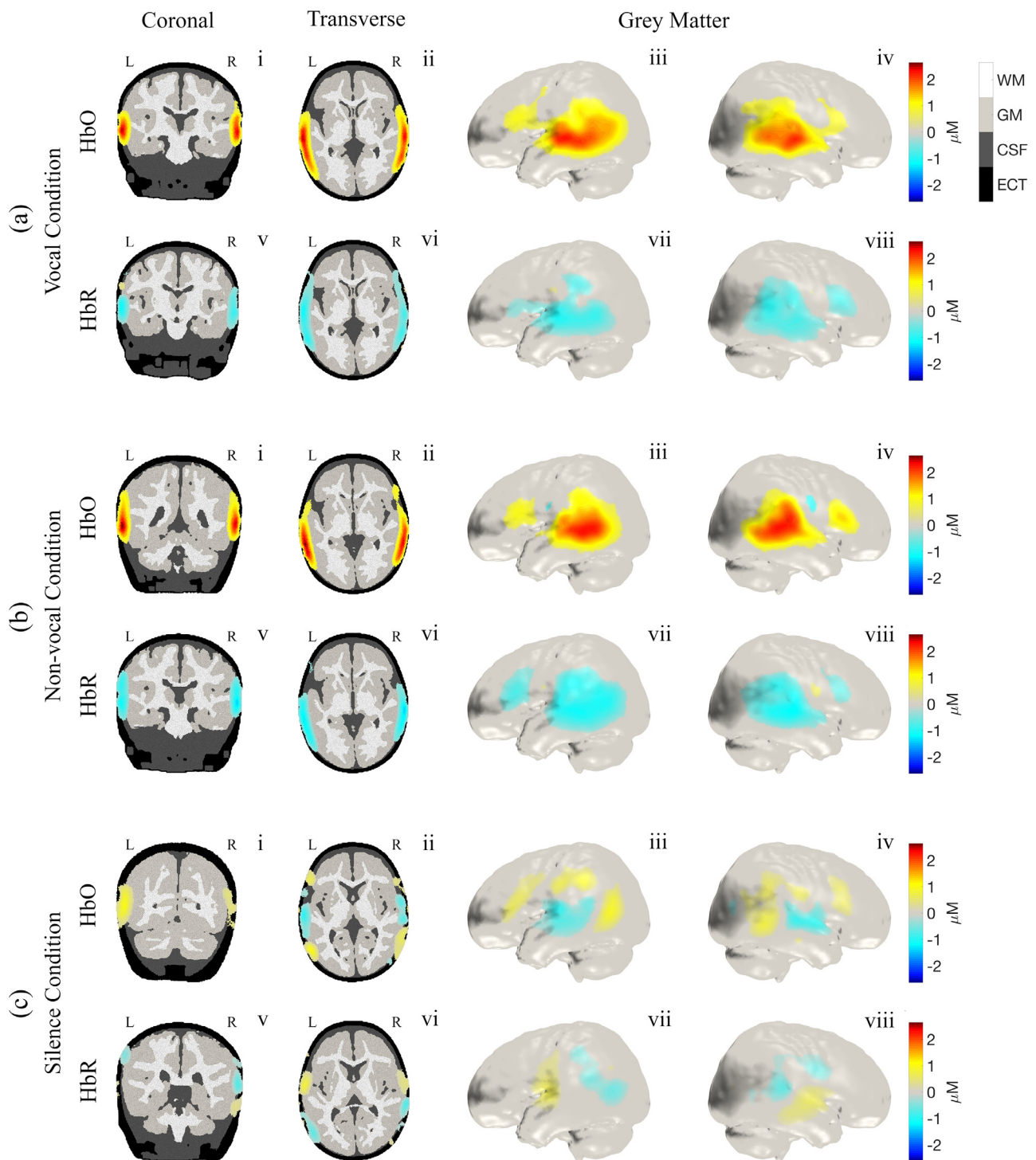


Fig. 7. Volumetric and grey-matter surface images for HbO and HbR at the group level, displayed for each of the different social conditions. Images are an average of a time window of between 10.5 and 15.5 s post-stimulus-onset, relative to mean of the 5 s pre-stimulus. Each image is thresholded at 25% of the maximum signal found in the volume. The HbO panels display the images obtained for the social conditions in coronal section (i), transverse section (ii) and an extraction of the grey-matter surface (left hemisphere (iii), right hemisphere (iv)). The subsequent panels (v-viii) show the equivalent HbR images in each case. Note that the volume images demonstrate that the response is well localized to the cortex in all three dimensions.

Figs. 8(g–j) show the equivalent for the group average. The comparison between low-density and high-density array results shows clearly that the increase in channel density yields a dramatic increase in the magnitude of the reconstructed haemodynamic response at both individual and group levels. As shown in **Figs. 8(c(i,ii))** and **8(e(i,ii))**, the changes reconstructed with the low-density array are localised to the superfi-

cial layers, rather than the cortex. The low-density results also show spatially discontinuous increases in HbO, rather than the extended, bilaterally distributed response apparent in the high-density reconstructions. Similar results were obtained for the comparison at the group level (**Figs. 8(g–j)**) where the response obtained with the low-density array is lower in amplitude and predominantly localised to the superficial tis-

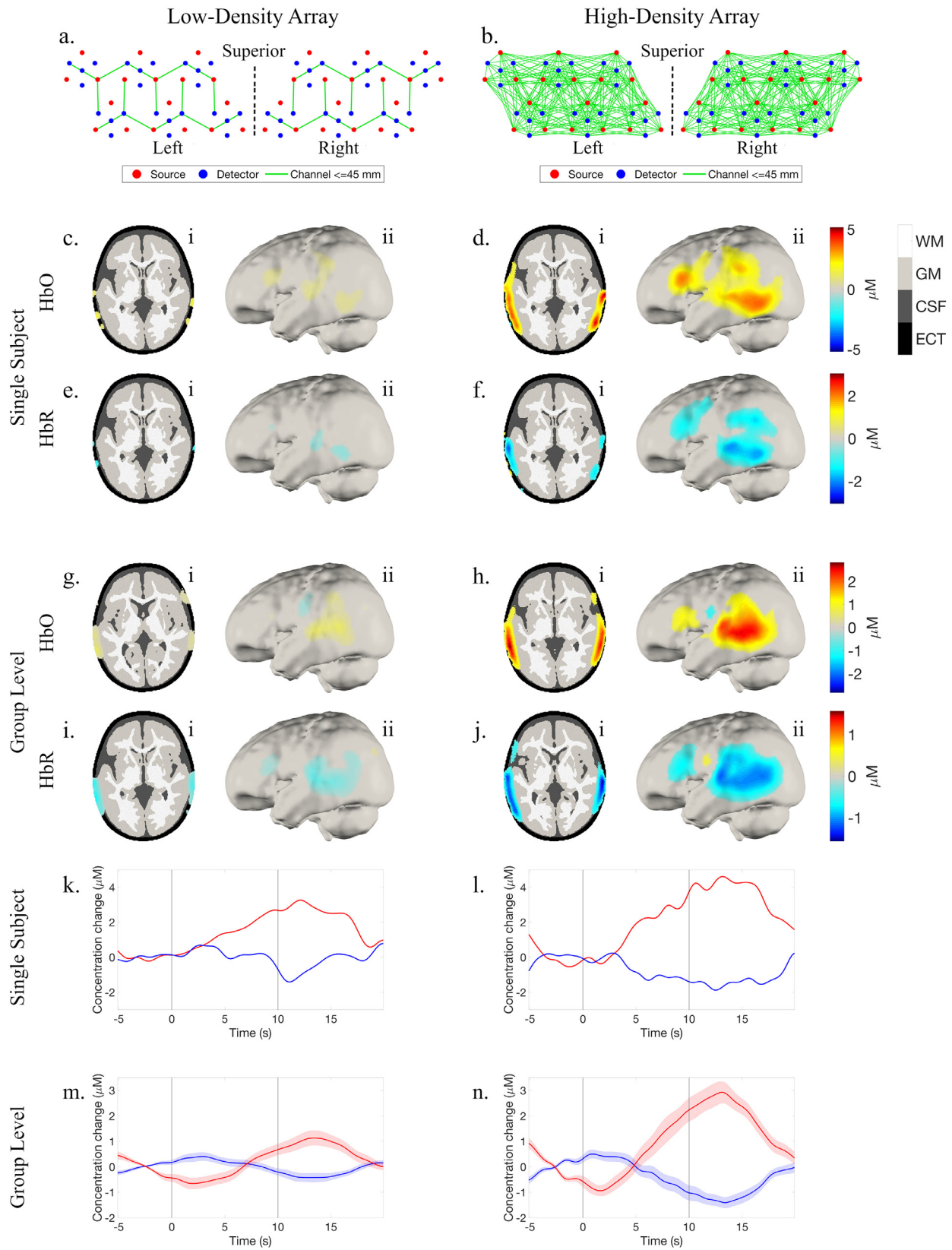


Fig. 8. Low-density and high-density results comparison. (a) The simulated low-density array with channels and source-detector distribution similar to a typical fNIRS system. The selected channels used to reconstruct low-density measurements are shown by green lines. A total of 30 channels with a source-detector separation between 20 and 24.5 mm were used. (b) The full high-density array. Volumetric (c(i) and d(i)) and grey-matter surface (c(ii) and d(ii)) HbO peak images in response to the non-vocal condition reconstructed with the low-density array and high-density array for a selected participant (c,d). Panels (e) and (f) show the equivalent for the HbR responses. Panels (g–j) show the equivalent for the group level results. Panels (k) and (l) show the peak HRFs from the reconstructed image space for the same single participant. The volume node that exhibited the highest amplitude HbO response was selected in each case. Panels (m) and (n) show the equivalent for the group-level data, but include shaded areas about each line that represent the standard error across all participants.

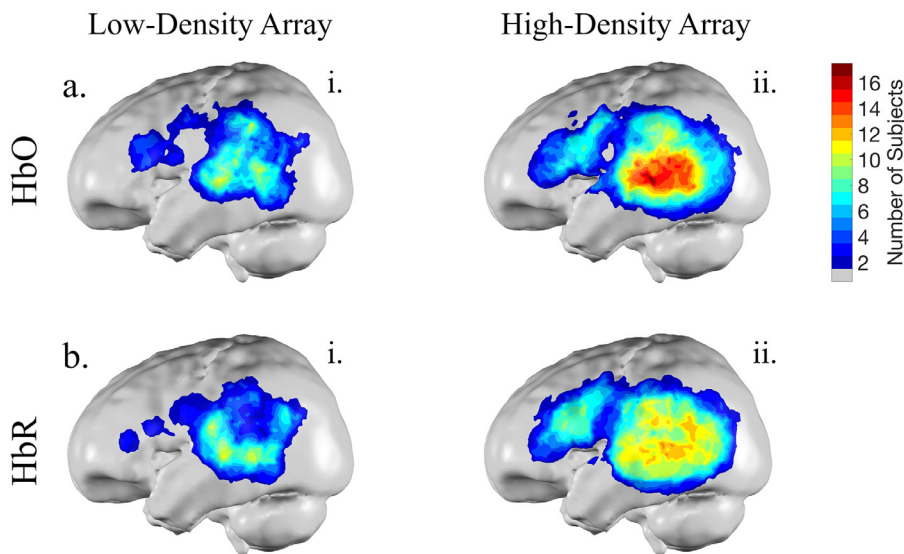


Fig. 9. The sum of the binarized, individual activation maps for HbO (a) and HbR (b) responses to the non-vocal condition, for both the low and high-density arrays.

sues. The group-level low-density response maps show a peak change in a location broadly consistent with the high-density array images, but with a smaller amplitude, a different spatial extent, and no changes apparent in the left pre-frontal cortex. Figs. 8(k,l) and 8(m,n) show the image-space haemodynamic responses extracted from the volume mesh node with the largest peak HbO signal, for the low-density and high-density arrays at the individual and group level. In this example, the signal-to-noise ratio (SNR, obtained by dividing the signal value by the standard deviation across participants) of the high-density HRF averages 216% that of the low-density equivalent for HbO and 208% for HbR. By extracting the HRF in the same fashion for the vocal condition the SNR of the high-density HRF is 180% that of the low-density equivalent for HbO and 237% for HbR.

Figs. 9(a) and 9(b) show the overlap of the binarized, individual activation maps (also for the non-vocal condition in this example) of HbO and HbR for the low and high-density arrays, respectively. The localisation of the response obtained at the individual level with the high-density array is much more consistent, with all 17 participants demonstrating a response in the superior temporal gyrus.

3.6. Statistical maps

Fig. 10 shows T-statistic maps of the response to all the social conditions compared to baseline. Note that a positive T-stat value for the contrast $A > B$ indicates that the mean of the changes in chromophore concentration for condition A within the peak window is significantly higher than that for condition B, but does not intrinsically require that either signal is associated with a canonical haemodynamic response; that information is present in Fig. 7.

For the vocal and the non-vocal conditions, the mean of the HbO response signal was significantly higher than that of baseline (Fig. 10(a) and 10(c)) and correspondingly, the HbR signal was significantly lower (Fig. 10(b) and 10(d)) compared to baseline, particularly in the posterior superior temporal gyrus stretching up to temporal-parietal junction. Note the responses were extremely significant, as demonstrated by the large T-stat scales. To analyse the responses to the visual element of the social stimulus, the silence condition was also compared to the baseline stimulus (a contrast that represents a non-social visual condition). The results shown in Fig. 10(e) and 10(f) demonstrate that HbO was significantly larger (and HbR significantly lower) for the silence condition compared to the baseline condition over the posterior temporal lobe, near the boundary with the occipital lobe, and also in the inferior frontal cortex, particularly in the left hemisphere.

Paired-sample, two-tailed t-tests were also used to compare vocal and non-vocal responses with the silence condition. As the silence condition consists of the same visual-social stimulus but without an auditory component, these comparisons constitute purely auditory contrasts. As shown in Fig. 11, the mean vocal and non-vocal auditory stimuli induced changes in HbO responses that were significantly higher than those of silence (Figs. 11(a) and 11(c)), with these changes well localised to the region of the primary auditory cortex of the superior temporal lobe. Correspondingly, the HbR signal was significantly lower (Figs. 11(b) and 11(d)) compared to that of the silence condition in similar cortical areas.

The same analysis was also performed to contrast the vocal and non-vocal conditions against one another, which identifies selective activation due solely to differences between the auditory stimuli. The results in Fig. 12 show a mixture of differing responses, particularly in the right hemisphere. The left hemisphere has several regions demonstrating significantly larger HbO values for the vocal than the non-vocal condition, but the HbR map is not consistent with a canonical haemodynamic response in all cases. Areas where there is a significantly larger HbO signal and a coincidentally smaller HbR signal include the anterior superior temporal gyrus and the superior part of the field of view in the somatomotor cortex. However, as shown by the inset HRF plots (which show the group-level HRFs obtained for a node at the indicated location, Figs. 12(e-h)), only the significant difference in the anterior superior temporal gyrus is consistent with a larger haemodynamic response; the apparent somatomotor cortex response is coincident with an inverted haemodynamic response to the non-vocal condition, as shown in Fig. 12(f) and also present below the threshold in Fig. 7(b).

4. Discussion

To our knowledge this is the first demonstration of a wearable, high-density optical imaging technology in the infant population. At present, most infant fNIRS research is undertaken using bulky optical fibre bundles or extensive cabling (Scholkmann et al., 2014), which limits experimental flexibility, restricts the subject's movements and often adds significant weight. Fibreless, wireless and/or wearable fNIRS systems have begun to overcome these issues by embedding NIR sources and detectors into wearable caps (Scholkmann et al., 2014; Zhao and Cooper, 2017), but these devices still maintain the limitations of fNIRS in terms of resolution, spatial uniformity and depth specificity. High-density diffuse optical tomography (HD-DOT) technologies exhibit increased spatial resolution and depth specificity when compared to typical fNIRS measures

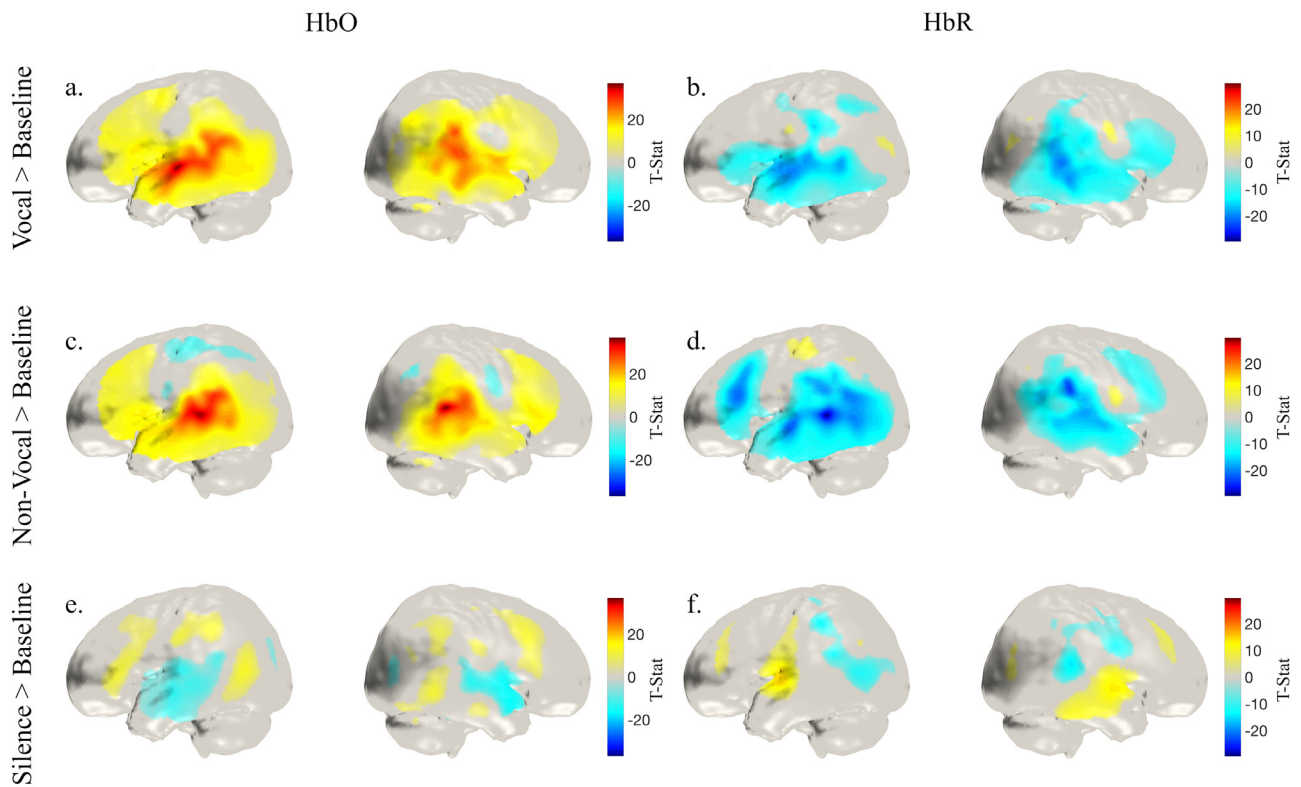


Fig. 10. T-statistic maps based on HbO (left, a, c, e) and HbR (right, b, d, f) data windowed around the peak of the response. Maps are provided for the visual social and auditory stimuli contrasted versus baseline (vocal > baseline (a,b)); (non-vocal > baseline (c,d)); and visual social stimulus versus baseline (silence > baseline (e,f)).

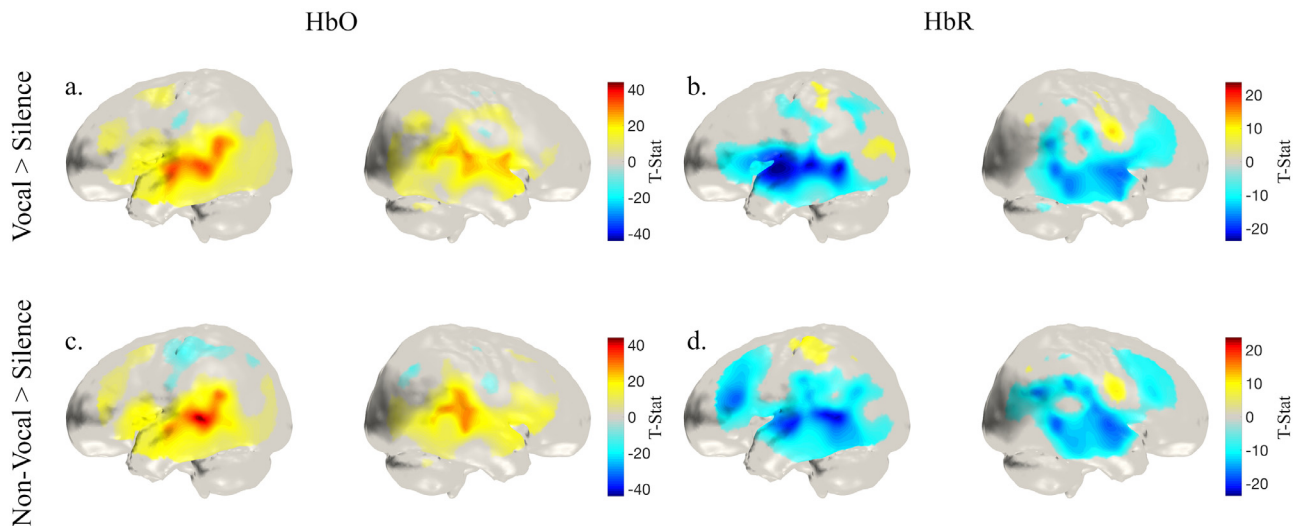


Fig. 11. T-statistic maps based on HbO (left, a, c) and HbR (right, b, d) data windowed around the peak of the response. Maps are provided that contrast vocal (a visual social and auditory condition) with silence (a visual social condition) (a,b); and non-vocal (a visual social and auditory condition) with silence (c,d). Note the differences are more confined to the region of the superior temporal cortex than those of Fig. 10.

(White and Culver, 2010; Eggebrecht et al., 2014; Joseph et al., 2006). However, the difficulty of applying HD-DOT with existing systems has limited the number of studies to date, particularly in infants and children.

The fibreless, modular and portable design of the system introduced in this study enabled the application of a large number of optodes to the head with minimal preparation or disturbance to the infant. The total time from the parent and baby entering the room to the beginning of a

recording was always between 10 and 15 min. The system also appears to be well-tolerated by the infant population. The flat light-guide pieces allowed good optical contact despite the presence of hair, which was not actively removed from underneath the optodes. The proportion of datasets used, the average duration of the recording sessions and number of included trials all reveal a high rate of data retention when compared to similar studies using fibre-based devices (Lloyd-Fox et al., 2009; Perdue et al., 2019; Lloyd-Fox et al., 2010; Grossmann et al., 2013).

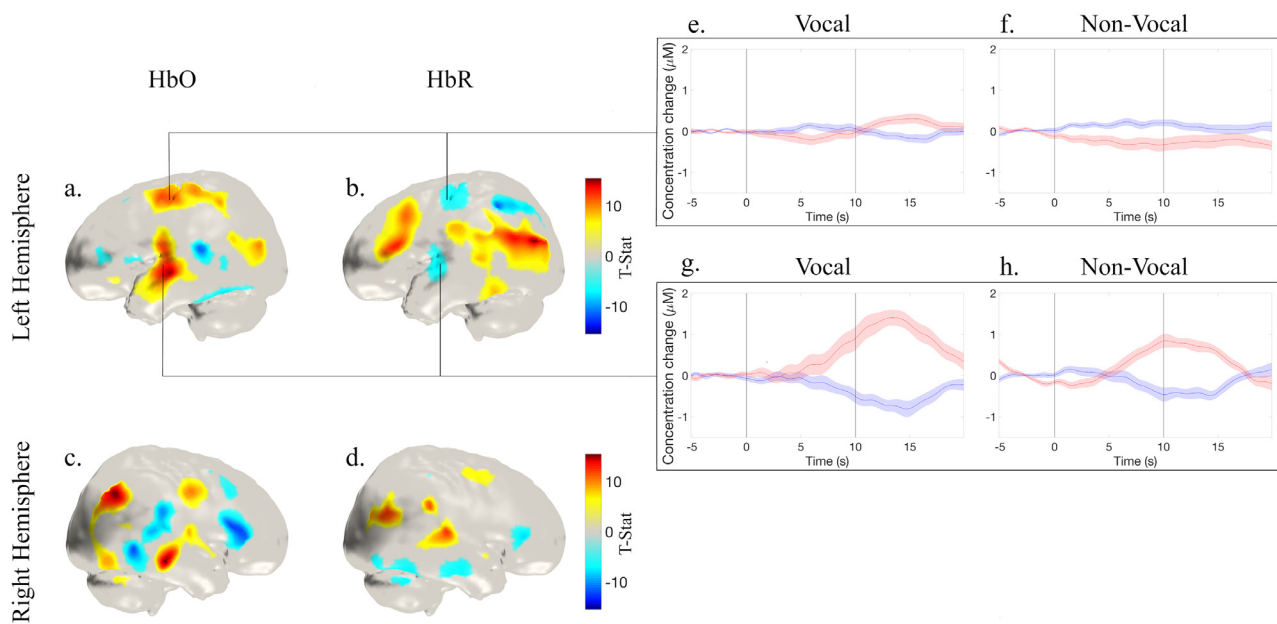


Fig. 12. T-statistic maps based on HbO (a,c) and HbR (b,d) data windowed around the peak of response that show the result of contrasting the vocal and non-vocal conditions (vocal > non-vocal). Selected HRFs for a node located in the somatomotor cortex (e,f) and a node located over the mid-portion of the superior temporal gyrus (g,h) are highlighted in the inset plots. The shaded areas about each line in the inset plots represent standard error across all participants. Note that what appears to be significantly stronger haemodynamic response to the vocal than the non-vocal condition in the somatomotor region is actually driven by a weakly inverted response to the non-vocal stimulus in this location. The only location that demonstrates both significance, and a canonical functional signal is the anterior superior temporal gyrus.

Our infant-friendly registration method enabled us to obtain the three-dimensional position of tiles relative to the cranial landmarks using the TrueDepth functionality of an X-series iPhone. Note that other devices, including the structure.io scanner (<https://structure.io>), are available to perform similar structured-illumination scanning.

The results of our pre-processing steps demonstrate impressive data quality, with measurements passing our thresholds for source-detector separations ranging from 10 to 41.8 mm in every participant (Fig. 4(a)). The group-level haemodynamic response maps (Figs. 5 and additionally S1, S2) show clear, consistent responses with high SNR. The responses are also canonical in form, with a clear and significant decrease in HbR. Many fNIRS studies of infants have reported minimal HbR contrast, or even inverted responses (Issard and Gervain, 2018), but that is demonstrably not the case in this dataset. In a manner consistent with the results of previous studies that employed the same paradigm (Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2013; Lloyd-Fox et al., 2014) our channel-wise data demonstrate functional activation in response to the visual and auditory social stimuli (vocal and non-vocal conditions) in channels located over the most posterior area of the array and (to a lesser extent) over the most anterior area of the array (Fig. 5 and Fig. S1).

One issue that is evident from the HRF curves (Figs. 5(a,b) and S1(a,b)) is that the haemodynamic response is relatively slow to begin and to recover, relative to a typical adult HRF (Plichta et al., 2007). In several cases the response has not returned to baseline in the allotted inter-stimulus interval. The block-averaging approach therefore appears to be sub-optimal for this paradigm. Note that because of this concern, we repeated our full analysis pipeline using a deconvolution model, but as the resulting T-stat maps were consistent with those obtained with the block-average approach, we chose to present the results of the more established analysis pipeline here.

The high-density sampling provided by the LUMO device allowed us to explicitly examine the effect of source-detector separation on the functional haemodynamic response curve, which is another topic of active interest in infant fNIRS research (Emberson et al., 2016).

Fig. 6 shows a haemodynamic response that clearly increases in magnitude with increasing source-detector separation; a strong indication that the signal is cerebral in origin. This result also implies that the default choice of 20–25 mm channel separations for infant fNIRS (as opposed to separations of 30 mm that are typically targeted in adult studies and for which most fNIRS devices can provide sufficient signal quality) may actually be detrimental to the SNR of the resulting functional responses, at least in some cases. Of course, the denser sampling provided by the 20–25 mm channel arrays that are typical in infant fNIRS studies does have other advantages over standard 30 mm arrays, particularly in terms of spatial specificity and ergonomics (Gervain et al., 2011).

In this paper we chose not to undertake a short separation regression approach either with the high-density data or with the simulated low-density data because of the lack of standardised approaches to do so in the infant population. When examining the HRFs associated with the shortest channels in the array (at ~10 mm), an oscillatory signal appears present, albeit one that is relatively small in amplitude (Fig. 6). This signal shares some, but not all, features of the haemodynamic response of the longer channels, suggesting that at least some component of this short-separation signal is superficial in origin. While inconclusive, this result does suggest short-separation regression strategies in infant fNIRS and HD-DOT may well be beneficial, and they have been used successfully in other work (Ferradal et al., 2016). The high-density nature of the device we present here should ensure that it is possible to undertake short separation regression methods in this (or any) population once standardised approaches are established.

Our methodology also allowed us to obtain functional images that show responses localized in all three dimensions (Fig. 7), again with co-incident increases in HbO and decreases in HbR. The reconstruction process allows us to confirm that these responses are centred in the posterior superior temporal gyrus, with activation also apparent in the inferior frontal cortex for the vocal and the non-vocal condition (Figs. 7(a) and 7(b)). The silence condition exhibit more mixed responses, with co-incident increases in HbO and decreases in HbR found only in the posterior superior temporal lobe near the boundary with the occipital

lobe, particularly in the left hemisphere (Fig. 7(c)). For all conditions, the maximum value of the reconstructed volume images is found below the superficial tissues, at a depth that is broadly coincident with the outer surface of the cerebral cortex.

The comparisons between the image results obtained using a low-density and a high-density array reveal that, while both configurations detect haemodynamic responses to the stimuli, there are clear differences in both the magnitudes and apparent locations of those responses. This is demonstrated at the single subject level in Figs. 8(c–f) and at the group level in Figs. 8(g–j). The high-density data also yields images that are localized in depth (Figs. 8(d,f,h and j)), in stark contrast to those obtained with the simulated low-density array. It is important to note, however, that the apparently superficial responses reconstructed using the low-density array (Figs. 8(c,e,g and i)) do not imply that the signal was superficial in origin. Rather, this is an artefact of the image reconstruction process when insufficient depth information is provided (Dehghani et al., 2009). Fig. 9 shows a comparison of the sum of binarized activation masks across the group for both oxyhaemoglobin and deoxyhaemoglobin with the low-density and the high-density arrays. These maps represent the number of subjects (out of 17) that showed a response in a given chromophore at a given location. An area showing a high number of subjects therefore implies a consistent response across subjects. The results of Fig. 9 demonstrate that high-density sampling yields improved spatial localization and a dramatic improvement in the consistency of the responses across participants. This effect is clearly evident in HbO (Fig. 9(a)) and although it is less dramatic (likely because of the lower contrast of HbR measurements), it is also evident in HbR (Fig. 9(b)).

In addition to improvements in spatial specificity, by comparing the reconstructed haemodynamic response curves of Fig. 8, the use of high-density sampling is shown to increase the SNR of the HRF by a factor of 210% on average across conditions and chromophores, relative to the low-density array. Assuming that the number of samples (i.e. participants) required is inversely proportional to the square of the SNR, this improvement in SNR implies that by applying HD-DOT, one could theoretically reduce the number of infants in the cohort by a factor of 4 and still achieve group-level responses comparable to those obtained with low-density fNIRS.

Extensive prior work has demonstrated that HD-DOT sampling also provides superior spatial resolution relative to low-density fNIRS measurements (White and Culver, 2010; Fishell et al., 2020). Because of the wide-spread nature of the responses to the multi-sensory stimuli we applied, a meaningful comparison of spatial resolution is unfortunately not possible here.

The statistical comparisons of the vocal and non-vocal conditions with baseline (Fig. 10) represent a contrast between a visual-social-auditory stimulus and a silent, non-social visual stimulation. This involves a wide range of cortical networks that causes a broad functional response that (while strongest within the superior temporal gyrus) is significant over much of the temporal lobe and the inferior prefrontal cortex (Figs. 10(a,b) and 10(c,d)). The statistical comparison of the silence condition (visual social with no sound) with the non-social visual baseline represents a differential visual-social contrast. The results demonstrated involvement of the posterior temporal lobe, near the boundary with the occipital lobe; the pSTS-TPJ region; and the inferior frontal cortex, particularly in the left hemisphere as shown in Figs. 10(e) and 10(f). Again, this result is broadly consistent with previous studies that used low-density fNIRS to investigate a similar-age cohort (Lloyd-Fox et al., 2009; Lloyd-Fox et al., 2018), a 12–16 months old cohort (Lloyd-Fox et al., 2017) and neonates (Farroni et al., 2013), with the exception that the neonatal study did not report inferior frontal cortex activation. These results support the hypothesis that 6 months-old infants process social dynamic stimuli within a specialized area of the temporal cortex, which seems to be operating from the first days of life (Lloyd-Fox et al., 2009; Farroni et al., 2013).

The auditory vocal and non-vocal conditions compared with the silent condition constitutes a purely auditory contrast, which results in a more focal statistical map than those found when comparing either condition to baseline (Fig. 11). Lastly, the comparison between the vocal and the non-vocal conditions (Fig. 12) identifies selective activation due to the difference between the auditory stimuli only (human voice versus environmental sounds). The only clear significant difference that was also associated with a functional response of the expected form was found in the left anterior superior temporal gyrus (Fig. 12(a)), which is preferentially activated by vocal (rather than non-vocal) sounds. Once again, this selective response in the left hemisphere is broadly consistent with the results of similar investigations in previous studies of 4–7 month-olds infants with fNIRS (Lloyd-Fox et al., 2012; Lloyd-Fox et al., 2018), though it is also important to note that other studies (using both fMRI and fNIRS and similar stimuli) have shown bilateral and right-dominant cortical activation of the superior temporal cortex (Belin et al., 2000; Blasi et al., 2011; Grossmann et al., 2010).

In conclusion, we have demonstrated the application of a new generation of wearable HD-DOT devices to the infant population. The system is well-tolerated, adaptable and easy to apply. Our data demonstrates that high-density sampling yields improved spatial specificity, better localisation and considerably higher SNR than typical fNIRS approaches. The resulting statistical response maps associated with an established social-stimulus paradigm are broadly consistent with the channel-space results of previous studies, but the improved localisation and the HD-DOT image reconstruction process permits direct association to the cortical anatomy. While there are a range of ways in which our HD-DOT data could be exploited further, we believe the results presented here represent a significant step forward in optical neuroimaging for developmental neuroscience.

Declaration of Competing Interest

This paper involves the application of a prototype technology developed by UCL spin-out company Gowerlabs Ltd. In addition to the Gowerlabs-affiliated authors, RJC holds a financial interest in Gowerlabs Ltd.

Credit authorship contribution statement

Elisabetta Maria Frijia: Conceptualization, Methodology, Data curation, Formal analysis, Writing - original draft. **Addison Billing:** Methodology, Data curation. **Sarah Lloyd-Fox:** Methodology, Data curation, Writing - review & editing. **Ernesto Vidal Rosas:** Data curation. **Maria Magdalena Crespo-Llado:** Methodology, Data curation. **Marta Perapoch Amadó:** Methodology, Data curation. **Topun Austin:** Methodology. **Andrea Edwards:** Methodology. **Luke Dunne:** Software. **Greg Smith:** Software. **Reuben Nixon-Hill:** Software. **Samuel Powell:** Software. **Nicholas L. Everdell:** Software. **Robert J. Cooper:** Conceptualization, Project administration.

Data and code availability statement

The data presented here were acquired as part of a wider research performed under ethical approval from an NHS Research Ethics Committee. Our ethical approval does not, at present, permit the public sharing of participant data. The code used in this paper has been developed and released via www.github.com/DOT-HUB.

Acknowledgments

RJC is supported by EPSRC Fellowship EP/N025946/1. SLF is supported by a UKRI Future Leaders Fellowship.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2020.117490](https://doi.org/10.1016/j.neuroimage.2020.117490).

References

- Allison, T., Puce, A., Spencer, D.D., McCarthy, G., 1999. Electrophysiological studies of human face perception. I: potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex* 9, 415–430.
- Allison, T., Puce, A., McCarthy, G., 2000. Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* doi:10.1016/S1364-6613(00)01501-1.
- Almli, C.R., Rivkin, M.J., McKinstry, R.C., 2007. The NIH MRI study of normal brain development (Objective-2): newborns, infants, toddlers, and preschoolers. *Neuroimage* 35, 308–325.
- Bartocci, M., et al., 2000. Activation of Olfactory Cortex in Newborn Infants After Odor Stimulation: A Functional Near-Infrared Spectroscopy Study. *Pediatr. Res.* 48, 18–23.
- Belin, P., Zatorre, R.J., Lafallie, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. *Nature* 403, 309–312.
- Blasi, A., et al., 2011. Early specialization for voice and emotion processing in the infant brain. *Curr. Biol.* 21, 1220–1224.
- Buckner, R.L., et al., 1996. Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging (neuroimaging single trial language prefrontal). *Neurobiology* 93.
- Chitnis, D., et al., 2016. Functional imaging of the human brain using a modular, fibre-less, high-density diffuse optical tomography system. *Biomed. Opt. Express* 7, 4275.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.
- Cristia, A., et al., 2013. An online database of infant functional near infrared spectroscopy studies: a community-augmented systematic review. *PLoS One* 8.
- Culham, J.C., et al., 2003. Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Exp. Brain Res.* 153, 180–189.
- Dehghani, H., Brian, R.W., Benjamin, W.Z., Tizzard, Andrew, Joseph, P.C., 2009. Depth sensitivity and image reconstruction analysis of dense imaging arrays for mapping brain function with diffuse optical tomography. *Appl. Opt.* 48, 137–143.
- Delpy, D.T., Cope, M., van der, P., Zee, S., Arridge, S.W., and J., W., 1988. Estimation of optical pathlength through tissue from direct time of flight measurement. *Phys. Med. Biol.* 33, 1433 To cite this.
- Di Lorenzo, R., et al., 2019. Recommendations for motion correction of infant fNIRS data applicable to multiple data sets and acquisition systems. *Neuroimage* doi:10.1016/j.neuroimage.2019.06.056.
- Duncan, A., et al., 1995. Optical pathlength measurements on adult head, calf and forearm and the head of the newborn infant using phase resolved optical spectroscopy. *Phys. Med. Biol.* 40, 295.
- EGgebrecht, A.T., et al., 2014. Mapping distributed brain function and networks with diffuse optical tomography. *Nat. Photonics* 8, 448–454.
- Ellis, C.T., Turk-Browne, N.B., 2018. Infant fMRI: a model system for cognitive neuroscience. *Trends Cogn. Sci.* 22, 375–387.
- Emerson, L.L., Crosswhite, S.L., Goodwin, J.R., Berger, A.J., Aslin, R.N., 2016. Isolating the effects of surface vasculature in infant neuroimaging using short-distance optical channels: a combination of local and global effects. *Neurophotonics* 3, 031406.
- Fang, Q., Boas, D.A., 2009. Tetrahedral mesh generation from volumetric binary and gray-scale images. In: *Proceedings of the Sixth IEEE International Conference on Symposium on Biomedical Imaging: From Nano to Macro*. IEEE Press, pp. 1142–1145.
- Farroni, T., et al., 2013. Infant cortex responds to other humans from shortly after birth. *Sci. Rep.* 3 (2851). doi:10.1038/srep02851.24092239.
- Ferradal, S.L., et al., 2016. Functional imaging of the developing brain at the bedside using diffuse optical tomography. *Cereb. Cortex* 26, 1558–1568.
- Fishell, A.K., et al., 2020. Portable, field-based neuroimaging using high-density diffuse optical tomography. *Neuroimage* 215.
- Frank, M.C., Vul, E., Johnson, S.P., 2009. Development of infants' attention to faces during the first year. *Cognition* 110, 160–170.
- Funane, T., et al., 2014. Quantitative evaluation of deep and shallow tissue layers' contribution to fNIRS signal using multi-distance optodes and independent component analysis. *Neuroimage* 85, 150–165.
- Gagnon, L., et al., 2012. Short separation channel location impacts the performance of short channel regression in NIRS. *Neuroimage* 59, 2518–2528.
- Gervain, J., et al., 2011. Near-infrared spectroscopy: a report from the McDonnell infant methodology consortium. *Dev. Cogn. Neurosci.* 1, 22–46.
- Grossmann, T., Oberecker, R., Koch, S.P., Friederici, A.D., 2010. The developmental origins of voice processing in the human brain. *Neuron* 65, 852–858.
- Grossmann, T., Cross, E.S., Ticini, L.F., Daum, M.M., 2013. Action observation in the infant brain: the role of body form and motion. *Soc. Neurosci.* 8, 22–30.
- Huppert, T.J., Diamond, S.G., Franceschini, M.A. & Boas, D.A., 2009. HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain.
- Issard, C., Gervain, J., 2018. Variability of the hemodynamic response in infants: influence of experimental design and stimulus complexity. *Dev. Cogn. Neurosci.* 33, 182–193.
- Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. *Neuroimage* 62, 782–790.
- Joseph, D.K., Huppert, T.J., Franceschini, M.A. & Boas, D.A. Diffuse optical tomography system to image brain activation with improved spatial resolution and validation with functional magnetic resonance imaging. (2006).
- Kanwisher, N., 2010. Functional specificity in the human brain: a window into the functional architecture of the mind. *Proc. Natl. Acad. Sci. U. S. A.* 107, 11163–11170.
- Lloyd-Fox, S., et al., 2009. Social perception in infancy: a near infrared spectroscopy study. *Child Dev.* 80, 986–999.
- Lloyd-Fox, S., Blasi, A., Elwell, C.E., 2010. Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci. Biobehav. Rev.* 34, 269–284.
- Lloyd-Fox, S., Blasi, A., Mercure, E., Elwell, C.E., Johnson, M.H., 2012. The emergence of cerebral specialization for the human voice over the first months of life. *Soc. Neurosci.* 7, 317–330.
- Lloyd-Fox, S., et al., 2013. Reduced neural sensitivity to social stimuli in infants at risk for autism. *Proc. R. Soc. B Biol. Sci.* 280.
- Lloyd-Fox, S., et al., 2014. Functional near infrared spectroscopy (fNIRS) to assess cognitive function in infants in rural Africa. *Sci. Rep.* 4.
- Lloyd-Fox, S., et al., 2017. Cortical specialisation to social stimuli from the first days to the second year of life: a rural Gambian cohort. *Dev. Cogn. Neurosci.* 25, 92–104.
- Lloyd-Fox, S., et al., 2018. Cortical responses before 6 months of life associate with later autism. *Eur. J. Neurosci.* 47, 736–749.
- Meek, J.H., et al., 1998. Regional hemodynamic responses to visual stimulation in awake infants. *Pediatr. Res.* 43, 840–843.
- Minagawa-Kawai, Y., et al., 2009. Prefrontal activation associated with social attachment: facial-emotion recognition in mothers and infants. *Cereb. Cortex* 19, 284–292.
- Molavi, B., Dumont, G.A., 2012. Wavelet-based motion artifact removal for functional near-infrared spectroscopy. *Physiol. Meas.* 33, 259–270.
- Pelphrey, K.A., Morris, J.P., Michelich, C.R., Allison, T., McCarthy, G., 2005. Functional anatomy of biological motion perception in posterior temporal cortex: an fMRI study of eye, mouth and hand movements. *Cereb. Cortex* 15, 1866–1876.
- Perdue, K.L., et al., 2019. Using functional near-infrared spectroscopy to assess social information processing in poor urban Bangladeshi infants and toddlers. *Dev. Sci.* 22.
- Plichta, M.M., Heinzel, S., Ehlis, A.C., Pauli, P., Fallgatter, A.J., 2007. Model-based analysis of rapid event-related functional near-infrared spectroscopy (fNIRS) data: a parametric validation study. *Neuroimage* 35, 625–634.
- Quaresima, V., Bisconti, S., Ferrari, M., 2012. A brief review on the use of functional near-infrared spectroscopy (fNIRS) for language imaging studies in human newborns and adults. *Brain Lang.* 121, 79–89.
- Richards, J.E., Xie, W., 2015. Brains for all the ages: structural neurodevelopment in infants and children from a life-span perspective. *Advances in Child Development and Behavior*, 48. Elsevier Inc.
- Richards, J.E., Sanchez, C., Phillips-Meek, M., Xie, W., 2016. A database of age-appropriate average MRI templates. *Neuroimage* 124, 1254–1259.
- Sakatani, K., Chen, S., Lichty, W., Zuo, H., Wang, Y.-P., 1999. Cerebral blood oxygenation changes induced by auditory stimulation in newborn infants measured by near infrared spectroscopy. *Early Hum. Dev.* 55.
- Sanchez, C.E., Richards, J.E., Almli, C.R., 2012. Neurodevelopmental MRI brain templates for children from 2 weeks to 4 years of age. *Dev. Psychobiol.* 54, 77–91.
- Scholkmann, F., Spichtig, S., Muehleemann, T., Wolf, M., 2010. How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation. *Physiol. Meas.* 31, 649–662.
- Scholkmann, F., et al., 2014. A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage* 85, 6–27.
- Schweiger, M., Arridge, S., 2014. The Toast++ software suite for forward and inverse modeling in optical tomography. *J. Biomed. Opt.* 19, 040801.
- Verriotti, M., et al., 2016. Mapping cortical responses to somatosensory stimuli in human infants with simultaneous near-infrared spectroscopy and event-related potential recording. *eNeuro* 3, 663–673.
- Von Lüthmann, A., et al., 2020. Towards Neuroscience in the Everyday World: progress in wearable fNIRS instrumentation and applications. *OSA Technical Digest (Optical Society of America, 2020)*. <https://www.osapublishing.org/abstract.cfm?uri=BRAIN-2020-BM3C.2>.
- White, B.R., Culver, J.P., 2010. Quantitative evaluation of high-density diffuse optical tomography: in vivo resolution and mapping performance. *J. Biomed. Opt.* 15, 026006.
- Wyser, D., Lamercy, O., Scholkmann, F., Wolf, M., Gassert, R., 2017. Wearable and modular functional near-infrared spectroscopy instrument with multidistance measurements at four wavelengths. *Neurophotonics* 4, 1.
- Zhao, H., Cooper, R.J., 2017. Review of recent progress toward a fiberless, whole-scalp diffuse optical tomography system. *Neurophotonics* 5, 1.
- Zhao, H., et al., 2019. Advances in wearable high-density diffuse optical tomography: first applications of a new commercial technology and development of an infant-specific research device. In: Dehghani, H., Wabnitz, H. (Eds.). *Diffuse Optical Spectroscopy and Imaging VII*, 2. SPIE doi:10.1117/12.2526976.
- Zimmermann, B.B., et al., 2019. Development of a Wearable fNIRS System Using Modular Electronic Optodes for Scalability. In: *Proceedings of the Biophotonics Congress: Optics in the Life Sciences Congress 2019 (BODA, BRAIN, NTM, OMA, OMP)* BW1A.3 (OSA doi:10.1364/BRAIN.2019.BW1A.3).