Non-invasive optical imaging of brain function with fNIRS: current status and way forward

Paola Pinti[1,2], Ilias Tachtsidis[2], Paul W. Burgess[3], Antonia F. de C. Hamilton[3]

[1]Centre for Brain and Cognitive Development, Birkbeck, University of London, UK.
[2]Department of Medical Physics and Biomedical Engineering, University College London, UK.
[3]Institute of Cognitive Neuroscience, University College London, UK.

Published in:
***Encyclopedia of the Human Brain, 2nd edition: Reference Collection in* Reference Module in Neuroscience and Biobehavioral Psychology, 2023**

https://doi.org/10.1016/B978-0-12-820480-1.00028-0

Keywords

functional near-infrared spectroscopy; functional optical neuroimaging; diffuse optical tomography; wearable; wireless; cognitive neuroscience; social neuroscience; ecological; neurovascular coupling; co-registration; systemic interferences; standardization.

Objectives

- To outline the physical and physiological principles underlying functional Near Infrared Spectroscopy (fNIRS)
- To provide an overview of the current status of fNIRS for cognitive neuroscience applications
- To discuss the applications that could benefit the most from fNIRS
- To discuss the open challenges of fNIRS for the reliable assessment of brain activity

Abstract

Over the past three decades functional Near Infrared-Spectroscopy (fNIRS) has gained increased popularity as a valuable and non-invasive neuroimaging tool to study brain functioning, both inside and outside the laboratory. This has been facilitated by the rapid developments in fNIRS instrumentation and analytics, with the availability of sophisticated devices and analytical methods that enabled cognitive neuroscientists to expand investigations in more ecologically-valid settings and a wider range of populations. In this chapter, we provide an overview of the current status of the fNIRS technology and discuss the ongoing challenges as well as possible way to overcome them in the future.

Introduction

Functional Near Infrared Spectroscopy (or fNIRS) is neuroimaging technique that uses light to measure the changes in cerebral haemodynamics and oxygenation that follows neuronal activity. Because fNIRS is non-invasive, versatile for a wide range of tasks and population, portable and now mobile/wireless, it is now established as a powerful tool for cognitive neuroscientist to investigate human brain functioning, especially in "everyday-type" settings. fNIRS is particularly suitable for populations who cannot tolerate MRI, including neonates and children, elderly people, or neurodivergent people or those with psychiatric conditions. It is also valuable for tasks and applications requiring free movement including exercise and face-to-face social interactions, where other neuroimaging modalities like fMRI are not ideal. The value of fNIRS over 30 years has recently been celebrated in *Neurophotonics* with a two-volume special issue (https://www.spiedigitallibrary.org/celebrating-30-years-of-fnirs) showcasing the newest developments in hardware, applications, data analytics that have contributed and will continue to contribute to advance our understanding of the human brain.

As the field of fNIRS continues to advance, alongside opportunities there are limitations that need to be overcome to fully harness the potential of this neuroimaging method. This chapter aims to provide an overview of the possibilities and limitations of fNIRS within the field of cognitive neuroscience, particularly in in the quest for ecologically-valid scientific findings.

We start by outlining the physiological and physical principles underlying fNIRS; we follow in Section 2 with an overview of the types of NIRS hardware available to the market, and discuss what can and cannot be achieved with the available instrumentation in relation to the cognitive questions we intend to ask. Section 3 delves into the methodological challenges associated with the intrinsic limitations of fNIRS including the impact of physiological interferences and the lack of anatomical information; we review these issues and provide practical suggestions on how we believe it is best to account and overcome some of these limitations. We conclude by discussing the open challenges that still need to be addressed as the field moves toward more naturalistic investigations of brain functioning. Section 4 delves into the use of fNIRS for cognitive neuroscience experiments in settings with higher degree of ecological validity, exploring the associated complexities related to data quality, task design, individual differences, data analysis and findings interpretation; within this framework we provide examples of previous studies and suggestions on what strategies we can adopt with the resources available to use fNIRS as robustly and reliably in these contexts. Finally, we discuss the potential of fNIRS in the study of social interactions in hyperscanning configurations, describing concerns that have been raised and recommending practices to mitigate these.

1. Physiological and Physical Principles

1.1 The origin of the haemodynamic response

Neuroscientists aim to measure neural firing patterns within the brain in order to understand how the brain performs different tasks as well as why performance sometimes goes wrong. However, direct measurement of human brain activity is challenging, and most studies rely on proxy measures. Like fMRI, fNIRS measures blood flow in the brain as a proxy measure of neural activation. This works because there is a reliable *haemodynamic response* which links neural activity to blood flow changes. This arises because the development and healthy functioning of the human brain strongly relies on the functional and anatomical interdependence of neurons, astrocyte glial cells, and cerebral blood vessels (Philips et al., 2016). This is often referred to as the *NeuroVascular Unit* (NVU; Iadecola, 2017) and is

essential to ensure an adequate supply of energy substrates to the brain, mostly in the form of oxygen and glucose through the cerebral blood flow (CBF).

CBF is regulated in relation to the local level of neuronal activity in the brain: an increase in the neural activity must be followed by an increase in the blood flow reaching the activated area of the brain to meet the increased metabolic demand for oxygen and glucose. This mechanism is called *neurovascular coupling* and involves several processes that regulate the blood flow to the brain. Figure 1 provides a simplified visualization of the cascade of physiological processes that are triggered by the activation of neurons in a task. Vasoactive signals regulating CBF are generated through various pathways. Briefly, synaptic activity leads to the release of glutamate which activates receptors on the postsynaptic neurons and on the astrocytes that, in turn, trigger the increase of intracellular Ca^{2+}. This results in the production of vasoactive agents such as nitric oxide (NO) and prostanoid that act on the smooth muscle of the local blood vessels, dilating them. In parallel to this, adenosine is also produced and acts as a vasodilator on the nearby vessels as well as changes in the ion concentrations of e.g. potassium and calcium is triggered. As a result of vasodilation, an increase in blood volume is observed, to ensure an adequate supply of oxygen to the activated area.

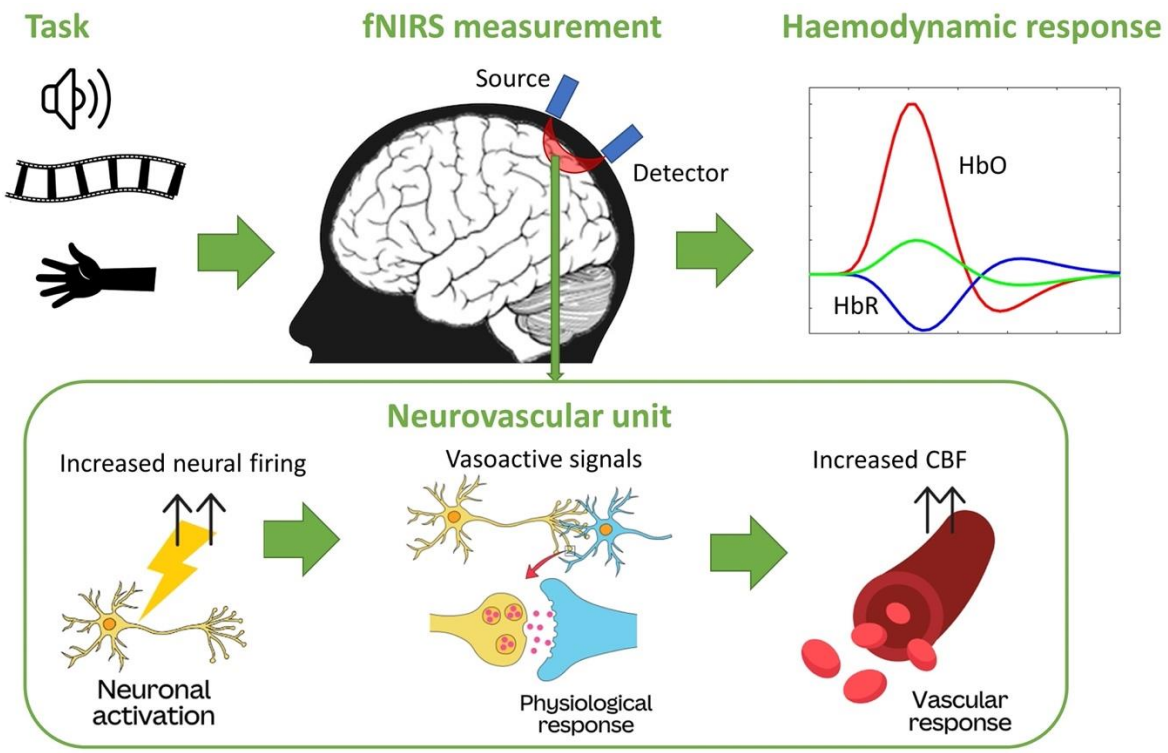

Figure 1. **Representation of the main neurovascular coupling mechanisms triggered by the processing of a certain stimulus and that leads to the hemodynamic response, detectable through fNIRS.**

A large body of work has demonstrated that the increase in the CBF that follows neuronal activity can be up to 4 times larger than the actual amount of energy demand needed to support that activity. This oversupply of CBF results in a reduction of the fraction of oxygen extraction and might act as preparatory mechanism in case surrounding neurons might become active, requiring additional energetic substrates (Mark et al., 2015). The exaggerated CBF response and the increase in blood volume and oxygenation are the basis of the modern functional neuroimaging methods such as functional Magnetic Resonance Imaging (fMRI) and functional Near Infrared Spectroscopy (fNIRS). fMRI tracks this phenomenon by measuring the so-called Blood Oxygen Level Dependent (BOLD) signal, that originates from the changes in the magnetic properties of the blood due to a decrease in concentration of deoxyhemoglobin (Kim and Ogawa, 2012); fNIRS estimates the changes in blood oxygenation through the evaluation of the different light absorption by hemoglobin in its oxygenated and deoxygenated forms (Figure 1).

By measuring blood oxygenation in the brain, fMRI and fNIRS can both therefore indirectly measure neural activation in specific brain regions, providing invaluable tools to study cognitive functions as well as cognitive impairments. By measuring the changes in brain haemodynamics and oxygenation, we can localize in time and space – with certain spatial and temporal resolutions – the patterns of brain activity linked to the behavior of interest. For instance, we can understand which brain region regulates a certain cognitive function; how different brain regions are interconnected to support a certain cognitive process; and how changes in brain functioning underpin cognitive development or cognitive dysfunctions.

This chapter explores how fNIRS can help answering some of these questions and what are its limitations, especially in those sub-areas of cognitive neuroscience that we believe would benefit the most from it. In the following section, we first provide an overview of the physical principles of fNIRS and a clarification of the terminology that will be used throughout the chapter.

1.2 Physical basis of functional Near Infrared Spectroscopy

Functional Near Infrared Spectroscopy or fNIRS is an optical neuroimaging technique based on neurovascular coupling that monitors the changes in brain hemodynamics and oxygenation through the estimation of the changes in concentration of oxygenated (oxyhaemoglobin or HbO_2) or deoxygenated (deoxyhaemoglobin or HbR) hemoglobin. These are used as an indirect measure of brain activity. More precisely, a functional haemodynamic response is defined as a concurrent increase in HbO_2 and total haemogolobin (HbT= HbO_2 +HbR) and a decrease in HbR.

fNIRS works by shining near infrared light in the range 600-1000nm into the head. This range is often referred to as *optical window into the body* and is chosen because NIR light can transmit quite well through skin and body as absorption due to water is low in this range (see Figure 2 A). In fNIRS systems, light sources are placed on the head and the emitted light is able to diffuse through the various layers of the head and reach the brain tissue. Along its path, light interacts with the biological tissues through two main processes: absorption and scattering.

Absorption is a wavelength-dependent process by which light is selectively absorbed by the atoms in the material it interacts with, reducing its intensity (i.e., light attenuation). The main absorber in the optical window of our interest is hemoglobin, that has distinct absorption properties in its oxygenated and deoxygenated forms (Figure 2); in addition, absorption due to water is minimum and increases for λ>~900 nm.

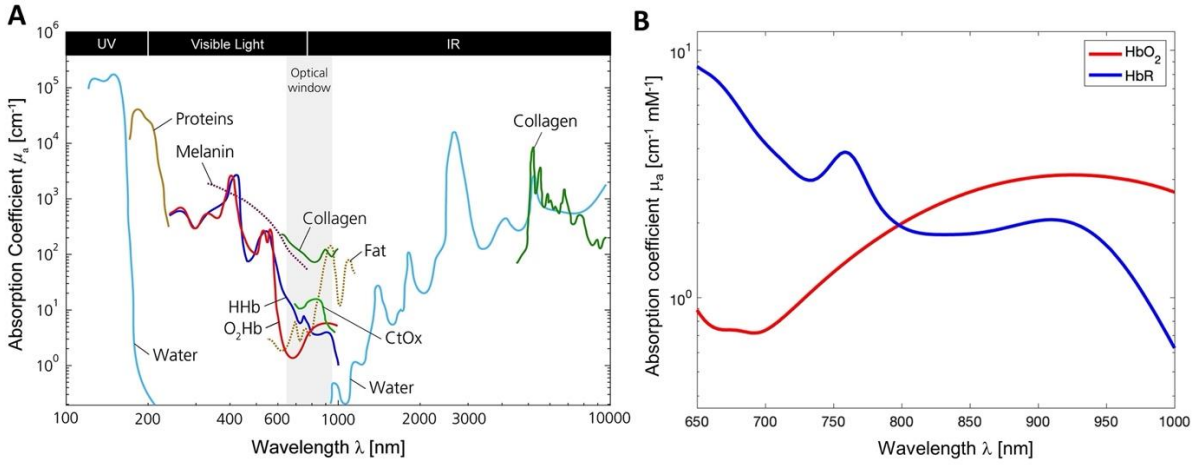

Figure 2. Absorption spectra of the main absorbing compounds in the biological tissue (A; reproduced from Scholkmann et al., (2014) under the terms of the Creative Commons CC BY license). The optical window is highlighted in grey. Panel B shows the absorption spectra of HbO_2 (red) and HbR (blue) in the optical window. Absorption coefficients are taken from https://github.com/multimodalspectroscopy/UCL-NIR-Spectra.

Other compounds like lipids and melanin may affect light absorption but their concentration can be considered constant within the limited duration of our measurements so they do not impact on measures of change in blood flow. As melanin is a strong absorber, it might make it challenging to perform NIRS measurements on darker or thicker hair (Kwasa et al., 2022) or fixed hairstyles (e.g., braids), but many studies have success in participants with dark skin tones and hair (Lloyd-Fox et al., 2014, 2019; Jasińska et al., 2023). Further research to optimize fNIRS for all populations is urgently needed. Taken together, these properties enable NIR light to travel through the scalp, skull, cerebrospinal fluid and reach the brain as well as allow us to recover the concentrations of both HbO_2 and HbR. In general, the attenuation - or optical density (OD) - of light travelling through an absorbing and non-scattering medium can be described by the Beer-Lambert law (Eq. 1.1):

$$OD(\lambda) = -log\frac{I}{I_0} = \mu_a(\lambda) \cdot d = \varepsilon(\lambda) \cdot c \cdot d \quad (1.1)$$

where I_0 and I are the intensity of the incident and transmitted light, respectively. The ratio of I to I_0 is proportional to the absorption coefficient $\mu_a(\lambda)$ [cm^{-1}] at a given wavelength λ, and the linear distance d [cm] travelled by the light (i.e., optical pathlength). In particular, $\mu_a(\lambda)$ is the probability of the photon to be absorbed by a certain compound per unit length and is proportional to the extinction coefficient $\varepsilon(\lambda)$ [cm^{-1} $molar^{-1}$] of the absorbing compound at a given wavelength (i.e., how strongly the substance attenuates light) and its concentration c [molar]. These absorption coefficients are plotting in Figure 2.

Alongside absorption, in the brain light is attenuated via the effect of scattering. Scattering is 100 times more frequent than absorption and is the process by which light interacts with particles that randomly alter the photons' path. It can be described by the scattering coefficient $\mu_s(\lambda)$ that indicates the probability of a photon to be scattered per unity

of length [cm^{-1}]. When a photon is scattered many times, the probability of being absorbed increases as it travels for a longer distance.

The account for the effects of scattering, the Beer-Lamber law (Eq. 1.1) is modified to include two factors: the Differential Pathlength Factor (or DPF), that reflects the increased distance travelled by the photon due to scattering; a scattering dependent attenuation parameter $G(\lambda)$. This takes the name of modified Beer-Lambert law (MBLL, Eq. 1.2) and defines OD as:

$$OD = \varepsilon(\lambda) \cdot c \cdot d \cdot DPF(\lambda) + G(\lambda) \tag{1.2}$$

The MDLL is difficult to solve to gather absolute measurements of OD as $G(\lambda)$ is very difficult to determine with conventional fNIRS instruments but it can be assumed time invariant under the hypothesis of high and constant scattering. Therefore, by measuring changes in OD ($\Delta OD(\lambda)$) at time t respect to t_0, $G(\lambda)$ cancels out:

$$\Delta OD(\lambda) = OD(\lambda, t) - OD(\lambda, t_0) = \varepsilon(\lambda) \cdot \Delta c \cdot d \cdot DPF(\lambda) \tag{1.3}$$

Compounds like fat, melanin or water do not change in concentration significantly during the timeframe of typical neuroimaging studies and can be assumed constant; the concentration of HbO_2 and HbR instead change and can thus be considered as the main absorbers in the NIR range of interest (Figure 2 B). $\Delta OD(\lambda)$ can be expressed as a linear combination of the attenuation of HbO_2 and HbR:

$$\Delta OD(\lambda) = (\varepsilon_{HbO_2}(\lambda) \cdot \Delta[HbO_2] + \varepsilon_{HbR}(\lambda) \cdot \Delta[HbR]) \cdot d \cdot DPF(\lambda) \tag{1.4}$$

where $\varepsilon_{HbO_2}(\lambda)$ and $\varepsilon_{HbR}(\lambda)$ are the extinction coefficients of HbO_2 and HbR at a given wavelength λ (Figure 2 B). Equation 1.4 cannot be solved as $\Delta[HbO_2]$ and $\Delta[HbR]$ are unknown. To calculate the concentration changes in HbO_2 and HbR, measures of optical density are performed at two (or more) wavelengths:

$$\Delta OD(\lambda_1) = (\varepsilon_{HbO_2}(\lambda_1) \cdot \Delta[HbO_2] + \varepsilon_{HbR}(\lambda_1) \cdot \Delta[HbR]) \cdot d \cdot DPF(\lambda_1) \quad (1.5)$$

$$\Delta OD(\lambda_2) = (\varepsilon_{HbO_2}(\lambda_2) \cdot \Delta[HbO_2] + \varepsilon_{HbR}(\lambda_2) \cdot \Delta[HbR]) \cdot d \cdot DPF(\lambda_2)$$

For example, in many systems wavelengths close to 690 nm and 830nm are used because these have distinct absorption profiles for HbO and HbR. The system of equations in 1.5 can now be solved to calculate $\Delta[HbO_2]$ and $\Delta[HbR]$:

$$\begin{bmatrix} \Delta[HbR] \\ \Delta[HbO_2] \end{bmatrix} = {}^{1}/_{d} \begin{bmatrix} \varepsilon_{HbR}(\lambda_1) & \varepsilon_{HbO_2}(\lambda_1) \\ \varepsilon_{HbR}(\lambda_2) & \varepsilon_{HbO_2}(\lambda_2) \end{bmatrix}^{-1} \begin{bmatrix} \Delta OD(\lambda_1)/DPF(\lambda_1) \\ \Delta OD(\lambda_2)/DPF(\lambda_2) \end{bmatrix} \quad (1.6)$$

The $DPF(\lambda)$ is both wavelength- and age-dependent can be calculated through the diffusion equation of light transport (Duncan et al., 1996; Scholkmann and Wolf, 2013); however, this is not possible with typical commercially available fNIRS systems (continuous wave, see Section 2.1). Recently, Scholkmann et al. proposed a general equation to estimate the DPF using the following equation (Scholkmann et al., 2013):

$$DPF(\lambda, A) = \alpha + \beta A^{\gamma} + \delta\lambda^3 + \varepsilon\lambda^2 + \zeta\lambda \quad (1.7)$$

where A is the age in years, $\alpha = 233.3$, $\beta = 0.05624$, $\gamma = 0.8493$, $\delta = -5.723 \times 10^{-7}$, $\varepsilon = 0.001245$ and $\zeta = -0.9025$. However, in some cases a precise DPF is not needed (see below).

2. Overview of available hardware

2.1. Operating principles

NIRS-based systems can be divided into three main classes: continuous-wave (CW), time domain (TD) and frequency domain (FD) instruments (Figure 3).

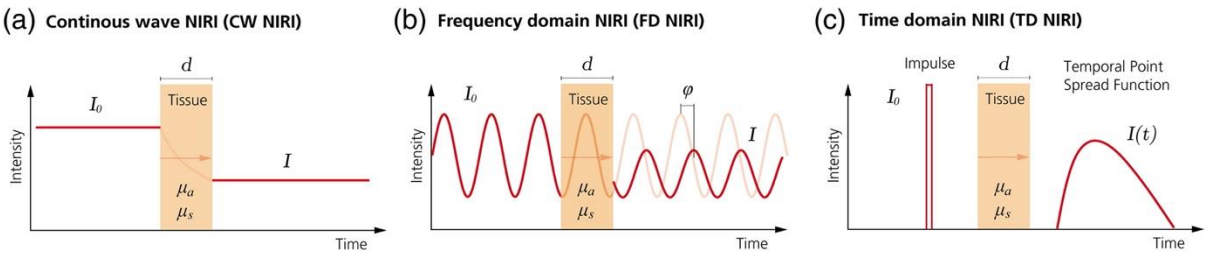

Figure 3. Types of NIRS-based devices: (a) continuous wave, (b) frequency domain, (c) time domain (reproduced from Scholkmann et al., (2014) under the terms of the Creative Commons CC BY license).

CW instruments are the simplest in terms of operation principles and are the most common among the commercially available NIRS systems. They measure changes in light absorption by shining the head with NIR light at two or more wavelengths and collecting the back-scattering light with a detector placed a few cm away from the light source on the scalp. Given that CW systems are not able to determine the absorption $\mu_a(\lambda)$ and scattering ($\mu_s(\lambda)$) coefficients nor the DPF, only the changes in concentration of HbO_2 and HbR can be estimated. Extensions of CW-NIRS technologies that include multiple source-detector distances and/or multiple 100s of NIR wavelengths can allow algorithms such as Spatially Resolved Spectroscopy to quantify absolute tissue oxygenation levels (Suzuki et al., 1999; Kovacsova et al., 2021).

FD instruments use light sources that emit NIR light modulated at a certain frequency. When this modulated light interacts with the biological tissue, its amplitude changes and its phase shifts per effect of absorption and scattering. Therefore, the detector is able to quantify both light attenuation and the phase shift of the back-scattered light, hence estimating the optical pathlength (i.e., the DPF), $\mu_a(\lambda)$ and ($\mu_s(\lambda)$ (Fantini and Sassaroli, 2020). Absolute measures of HbO_2 and HbR are thus possible as well as the calculation of the concentration changes can take into account individual scattering changes and accurate pathlength.

Finally, TD instruments are the gold standard and incorporate light sources that emit pulses of light of a few picoseconds and a fast time-resolved detectors to measure the intensity and delay in the time-of-flight (temporal spread function) of the emerging photons following the interaction in the biological tissue (i.e. absorption and scattering). The temporal spread function thus gives information about the amount of light scattering and absorption and also about the depth that the photons reached inside the tissue, enabling measures of

absolute concentration of HbO_2 and HbR. However, one of the biggest advantages of TD instruments is the possibility to get depth-resolved measures. By evaluating the distribution of times of flight of late and early photons, the deep and superficial signals coming from the brain and the extracerebral layers of the head can be separated (Wabnitz et al., 2020). This is extremely important to minimize the confounding components coming from hemodynamic changes happening in the scalp that contaminate the fNIRS signals (see Section 3.1). Further to this, it was also shown that TD-NIRS can achieve the deepest penetration depth and be less prone to absorption changes in the scalp compared to CW- and FD-NIRS (Gunadi et al., 2014).

The vast majority of commercially available NIRS instruments are CW-based as they are cheaper than TD and FD and provide enough information for neuroscientific investigations. In fact, neuroscience studies do not typically rely on absolute quantifications of HbO_2 and HbR but the concentration changes respect to a baseline are rather evaluated to infer functional brain activity. However, more recently more TD- and FD-based technologies are being brought into the market.

2.1.1 NIRS-based hardware for monitoring brain metabolism

It is worth mentioning the existence a slightly different NIRS-based instruments, namely broadband NIRS (or bNIRS) and diffuse correlation spectroscopy (DCS), that can be used to obtain information about brain metabolism.

Whilst all the technologies described above use two or three wavelengths of NIR light and hence estimate the changes in HbO_2 and HbR only, bNIRS is able to also quantify the changes brain metabolism beyond brain hemodynamics and oxygenation. This is achieved by monitoring the changes in light attenuation due to the changes in the redox state of cytochrome-c-oxidase (or oxCCO). Cytochrome-c-oxidase is another physiologically-relevant chromophore as it gives information about brain metabolism. However, it is present in lower concentration than hemoglobin thus making a small contribution toward the total tissue absorption and can only be measured with particular instruments (see Bale et al., 2016). Cytochrome-c-oxidase is one of the enzymes in the electron transport chain in the mitochondria responsible for the aerobic metabolism of glucose. During the production of energy in the form ATP, there are a series of chemical reactions that involve electrons passing

through the electron transport chain the in mitochondria. These redox reactions lead to changes in the optical properties of cytochrome-c-oxidase and the difference in the absorption spectra between its oxidized and reduced forms can be used as a proxy of mitochondrial activity. Compared to the absorption spectra of HbO_2 and HbR (Figure 2), oxCCO has a broader spectrum with a peak at approx. 830 nm. In order to resolve for changes in oxCCO and avoid cross-talk among all the chromophores, several more wavelengths are needed, with a minimum of 8. bNIRS instruments for example utilize hundreds of wavelengths in the NIR range to separate the contribution of HbO_2, HbR, and oxCCO, thus being able to uniquely monitor both brain hemodynamics and oxygenation, and metabolism.

DCS is able to measure microvascular cerebral blood flow (CBF) by looking at the temporal fluctuations of the diffused NIR In the brain tissue. In this case, only one wavelength is required. DCS is based on the principle that NIR light gets scattered by moving red blood cells (RBC) and the temporal fluctuations of the back-scattered light speckle depends on the RBC motion. In particular, the autocorrelation function of the re-emerging light (or the DCS autocorrelation function) is calculated and reflects the temporal fluctuations in light intensity due to the scattering of moving RBCs (Durduran and Yodh, 2014). By combining the DCS-derived relative changes in CBF with tissue oxygen saturation measured with NIRS, it is possible to quantify the cerebral metabolic rate of oxygen extraction ($CMRO_2$), indicator of brain metabolism. Unlike other conventional modalities to measure CBF and metabolism, such as xenon-133 computed tomography, positron emission tomography or ASL-MRI, DCS does not use ionizing radiation, is portable and suitable for a wider range of populations (Buckley et al., 2014).

However, both bNIRS and DCS devices are currently not available on the market and are at the research stage.

2.2. Measurement mode

fNIRS recordings are performed by placing a certain number of light sources and light detectors onto the head, typically housed in caps or headbands. Measurements are obtained by pairing a source with a detector and the portion of the brain tissue investigated is called channel. A channel is located approximately half-way between a source and a detector and at

a depth of half of the source-detector distance. The larger the source-detector distance, the longer the penetration depth of the light, the deeper the brain tissue under investigation. However, the more the light travels through the tissue, the more it is attenuated due to higher probability of absorption and scattering. Therefore, to get a sufficient Signal-to-Noise ratio, sources and detectors are typically placed at a 3 cm distance in adults, and 2-2.5 cm in children, yielding a penetration depth of ~1.5-2 cm. Investigation with fNIRS is thus limited to the superficial layers of the brain cortex.

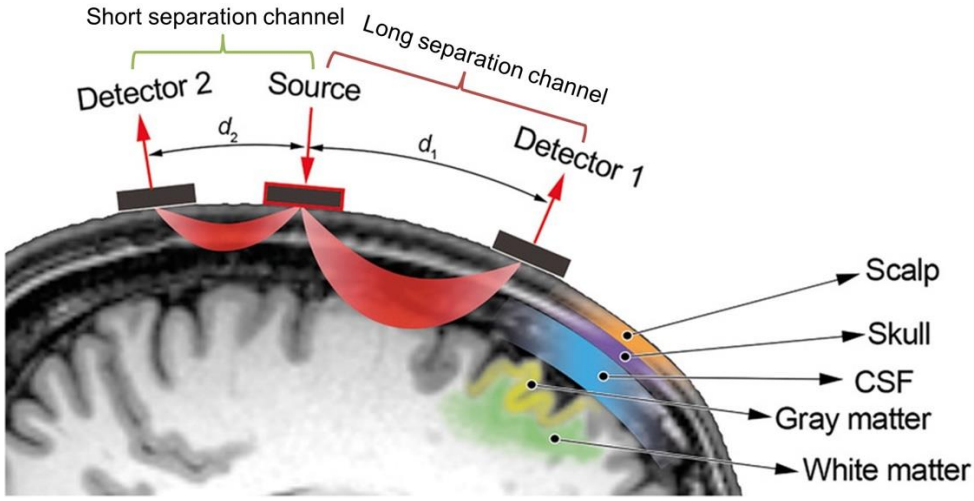

Figure 4. Path followed by the infrared light through the different layers of the head. The penetration depth of the light increase with the increase of the source-detector distance (d_1>d_2). Adapted with permission from Pinti et al., 2019.

The spatial sensitivity of NIRS-based instrumentation depends on the number of sources and detectors the system is equipped with and how these are arranged. In general, fNIRS hardware has massively improved over the past 30 years, moving from single channel instruments, which allow measurements from one point only, to multi-channel systems enabling topographic mapping of brain activity (Ferrari and Quaresima, 2012). fNIRS measures can be obtained using three main configurations: a) fixed distance; b) multi-distance; c) high-density.

Fixed distance. The most common way of collecting fNIRS data is to use sparse arrays of sources and detectors, i.e. optodes sparsely distributed at a fixed interoptode distance across the head. However, such configurations suffer from two main problems: 1) fNIRS measurements are sensitive to superficial contamination (see Section 3.1); 2) they have

limited spatial resolution and irregular spatial sensitivity. Previous work by White and Culver (2010) has demonstrated that sparse geometries may lead to mislocalizations of brain activity, providing good resolution only in the areas between close sources and detectors and hence not being suitable for detailed localization on cortical hemodynamics.

Multi-distance. The use of different source-detector distances in the configuration can help in minimizing the problem of superficial contamination in sparse arrays. As shown in Figure 4, the penetration depth of the light determines at which depth the fNIRS signal is recording from. Given that it is well documented that fNIRS signals from standard long-separation channels (e.g. Source-Detector 1 in Figure 4) are contaminated by hemodynamic changes happening in the scalp (Tachtsidis and Scholkmann, 2016), if we place a source and a detector close to each other (approx. 0.8 cm in adults and 0.2 cm in infants (Brigadoi and Cooper, 2015)), we are able to sample the hemodynamic changes in the extracerebral layers (e.g. Source-Detector2 in Figure 4). These are called short-separation channels and can be subtracted from the long separation channels to obtain a more brain-specific signal (Yücel et al., 2021). NOTE: the closest distance a source and a detector can be placed depend on the physical dimension of the optode of the particular fNIRS system. It is often not possible to place optodes at less than 1 cm from each other (e.g., Emberson et al., 2016).

High-density. When multi-distance arrays are extended to have several source-detector distances with overlapping channels, we talk about high-density fNIRS or high-density diffuse optical tomography (HD-DOT; Vidal et al., 2023). These instruments are more complex but provide denser measurements, at spatially continuous locations of the brain and at multiple depths, being able to drastically increase imaging quality while minimizing superficial contamination at the same time. The effective resolution of such dense array is good across the whole area probed by the grid, with an improvement in resolution from 3 cm of the sparse configuration to 1.3 cm of high-density grids (White and Culver, 2010). Recently several studies have further demonstrated how HD-DOT can significantly improve the localization of brain activity, the lateral and depth resolution, the SNR, and the reproducibility of the results, both in adults (Vidal et al., 2021) and infants (Frijia et al., 2021). In addition, image reconstruction can be performed to construct tomographic images of brain hemodynamics.

2.3 Hardware considerations

The range of applications of fNIRS as well as the breadth of information we can gather from fNIRS recordings strongly relies on the type of hardware we use and, to date, it is a compromise among several factors.

In general, fNIRS instrumentation can be divided in lab-based and mobile instrumentation. With *lab-based* equipment we mean all those systems that are connected to the main power supply and not battery operated, with the optical components housed within the main unit of the instrument that is not worn by the subject, and that requires the participant to be tethered to the machine via substantial cabling. With *mobile* we refer to all those devices that are substantially smaller and lighter than lab-based equipment, with optical components and a main unit which is worn and carried by the participant, allowing the participant to move freely with minor cabling or no cabling at all (i.e. wireless).

Lab-based systems are typically equipped with a larger number of sources and detectors, allowing coverage of larger portions of the head than mobile equipment. With the availability of many optodes, some can be "sacrificed" to create multi-distance configurations and include short separation channels without significantly compromising the head coverage. Since there are also no restrictions in terms of power consumption and size, these instruments can incorporate high quality optical components such as lasers and silicon photon multipliers, giving signals with a higher SNR – but also increasing the cost. High-density diffuse optical tomography (or HD-DOT) is also in principle possible as shown by Eggebrecht et al. (2014); however, to date few commercially HD-DOT systems are available.

Most of lab-based systems are based on the continuous wave technology and guide light to and from the head through optical fibers. When larger areas of the brain are monitored, the number of fiber optics increases which in turn increases the weight of the cap/headband onto the head and reduces the participant's comfort. In addition, the use of fiber bundles limits the participant's movements. Therefore, these instruments are better suited for static experiments that do not require walking or significant body movements. When used on infants and children, systems to hold the fiber bundles can be used to reduce the weight of the fibers, whilst making sure that the optical coupling between the head and the optodes is maintained without pulling the infant's head or causing discomfort. Frequency-

domain lab-based instruments are also available to the market and are able to give more accurate measures of oxy- and deoxy-hemoglobin than continuous wave; however, to achieve this, the use of sophisticated components makes these more expensive than continuous wave systems. Time domain-based fNIRS is superior in terms of the richness and quality of information it can gather; however, it also incorporates sophisticated optical components such as high-speed electronics that have challenged its miniaturization and cost. The only commercially available lab-based time domain instrument provides one measurement channel only.

Mobile instruments are generally equipped with a smaller number of sources and detectors as device miniaturization, weight and battery consumption play a critical role. In order to optimize these factors, compromises have to be made. For instance, LEDs and photodiodes are often used in continuous wave systems; these are inferior to lasers or silicon photomulitpliers (SiPM) in terms of SNR but are smaller, cheaper and require less power supply. Head coverage is generally limited to reduce the weight and power consumption and it can be difficult to include short separation channels. Nonetheless, wearability is a key feature making mobile fNIRS instruments suitable for novel and revolutionary investigations that can involve extensive body-movements as well as data collection outside the lab (Pinti et al., 2018). The majority of these instruments have the optical components directly coupled onto the head or make use of very short lightweight cables that do not compromise movements. More interestingly, the past few years have seen tremendous developments in terms of hardware miniaturization, with the first wearable continuous wave HD-DOT modular instrument and the first wireless time domain HD-DOT system made available to the market. While the continuous wave HD-DOT system is suitable for babies and children as well, time domain HD-DOT is only available in adult size. In spite of the higher costs of HD-DOT, these can uniquely provide whole-head measures with higher spatial sensitivity at multiple depths while maintaining freedom of motion. This comes with increased complexity both in the hardware, which has to handle a significantly larger number of optodes, and in the analytical tools, which have to be able to deal with increased analytical complexity and data volume (Vidal et al., 2023).

3. Methodological challenges and the way forward

Functional Near Infrared Spectroscopy is often seen as a tool that provides a good balance between the accuracy in mapping brain activity and the range of applications. In terms of measurement precision, fNIRS falls in between other popular neuroimaging techniques such as functional Magnetic Resonance Imaging (or fMRI) and electroencephalography/magnetoencephalography (or EEG/MEG), with better temporal resolution than fMRI (~tens of ms vs ~s) but lower than EEG (~ms), and better spatial resolution of EEG/MEG (~2-3 cm vs ~5-9 cm) but lower than fMRI (~3 mm). Additionally, fNIRS is portable, silent, generally tolerant to body movements, making it suitable for a wide range of applications (e.g., social interactions, exercise, sensory processing, etc) as well as populations (e.g. healthy and unhealthy, preterms and infants, elderly, etc).

However, on the other hand, there are some intrinsic limitations that can affect the reliability of the measurements obtained with fNIRS. In this section, we review some of the major methodological challenges and open questions that should be taken into account and addressed to improve results interpretation as well as the robustness of the conclusions.

3.1. The impact of systemic interferences on fNIRS signals

As described in Section 1.2, fNIRS shines infrared light into the head from a light source; the back-scattered light is then collected by a detector after photons have travelled twice through the different layers of the head (scalp, skull, CSF, grey matter, white matter). The recorded signal is thus influenced not only by the blood flow changes in the cerebral compartment, but also by the changes happening in the extra-cerebral layers of the head, where light gets absorbed to a higher extent in the skull and the scalp rather than in the grey matter (approx. 96% of the injected light is absorbed in the skin and bone and 3% in the brain; Haeussinger et al., 2011). In addition, fNIRS signals are affected by all those physiological processes that can alter the vascular tone (vasodilation/vasoconstriction) and the regulation of the blood flow, both in the brain and in the scalp. These can include but are not limited to changes in blood pressure, partial pressure of CO_2 ($PaCO_2$), respiration, heart rate. It is very

important to mention that some of these components are spontaneous and not directly evoked by the experimental task, such as the heartbeat (~1 Hz in adults, ~2 Hz in infants), respiration (~0.3 Hz), arterial blood pressure variations (or Mayer waves at ~0.1 Hz; Julien, 2006) and very low frequency modulations (VLF; <0.1 Hz); however, their amplitude can be modulated by the task itself or by posture changes. Moreover, other confounding components can be induced directly by the task. For example, speaking can lead to changes in $PaCO_2$ as well as non-neural hemodynamic changes due to the activation of the temporalis muscle with jaw movements or teeth clenching (Zimeo et al., 2018).

Because the measured fNIRS signal includes a mixture of different spontaneous and task-evoked components of neuronal and systemic origin, it is important for researchers to understand the systemic signals and how these might impact on data analysis and interpretation. Systemic physiological signals are generally considered as confounding factors in the fNIRS recordings and can lead to false positives and/or false negatives when inferring functional brain activity from fNIRS data (i.e., wrongly assess or miss the presence of a significant hemodynamic response; Tachtsidis and Scholkmann, 2016). It is thus crucial to be aware of the issue of systemic interferences in relation to each particular experiment, and to adopt certain practices to minimize their impact both at the task design and data collection stage and in the post-processing phase. Below we summarize some of the procedures that can help in enhancing the reliability of our fNIRS signals.

Pre- data collection strategies. Efforts to minimize the impact of physiological confounds can start from the moment we plan our experiment. First, when designing the task, one should identify the possible physiological responses that might be induced by the task of interest. As within the field of neuroscience we normally compare a task period to a baseline period or we compare two conditions (i.e., subtracting the activity to one from the other), we can ensure that similar amounts of systemic changes are induced by the different periods of the experiment. For instance, if a task involves participants to walk, all the experimental conditions should include walking. In this way, when we contrast conditions with each other, some of the hemodynamic changes induced by physiology may cancel out. For example, in our previous work (Burgess et al., 2022) we included two different baselines as part of the task: a cognitive only baseline (no walking), recruiting similar cognitive processes as the main task that involves cognitive+walking; a physical only baseline (no task), involving the same physical activity as the main task. These can serve to understand whether the

hemodynamic changes during the cognitive+walking task are due to the cognitive task itself or to the walk-related physiological changes and to isolate the task-evoked response.

Important improvements can also be made on the hardware side. As described in Section 2.2, hemodynamic changes in the scalp can be measured by placing a source and a detector very close to each other, to create short separation channels. fNIRS signals from the short separation channels can be subtracted from the fNIRS long separation channels to obtain a more brain-specific signal, minimizing scalp influence. Typically, this is done within the General Linear Model (GLM) framework using the short separation channels as additional regressors in the design matrix (Yücel et al., 2015) and regressing them out from the long separation channels. This is often called superficial signal regression. Several studies have demonstrated that the use of short separation channels improves the estimation of the hemodynamic response, both in task-evoked studies (e.g., Yücel et al., 2015) and resting-state functional connectivity studies (Paranawithana et al., 2022). It was also shown that the extracerebral interference is not homogenously distributed across the scalp (Gagnon et al., 2012; Kirilina et al., 2012; Erodoğan et al., 2014; Wyser et al., 2020). A better recovery of brain hemodynamic activity can thus be achieved when heterogeneity is assumed by placing short separation channels at different locations of the scalp (Gagnon et al., 2012; Wyser et al., 2020). However, this is not always possible due to a limited availability of optodes, especially in wireless and mobile fNIRS devices (see Section 2.3). DOT and TD-NIRS become advantageous in this regard, as DOT can sample fNIRS data from the scalp at multiple locations given its multidistance and overlapping optode configuration, and TD-NIRS is able to do so using the same optodes without additional ones. Therefore, in case of spare arrays, we suggest placing short separation channels to sample scalp interference from as many locations as possible, with at least at one or two locations within the probed region of interest.

Alongside superficial regression, systemic changes can be directly assessed by recording physiological data together with fNIRS. This is often referred to as systemic physiology augmented fNIRS (SPA-fNIRS), and allows to investigate the brain-body interactions. Physiological data can be used to minimize the confounding effects in the fNIRS signals as well as to investigate the relationship between brain activity and systemic physiology (Scholkmann et al., 2022). For example, in the field of social neuroscience, the availability of physiological signals can help in disentangling which factors modulate brain-to-brain coupling in hyperscanning studies (Hamilton, 2021; Guglielmini et al., 2022). Recent work by Zohdi and

colleagues has shown that the inter-subject variability in the hemodynamic responses can partially arise from the subject-specific changes in systemic physiology, further confirming the need for a SPA-fNIRS approach to better interpret the fNIRS data (Zohdi et al., 2021). SPA-fNIRS can also be enriched with other monitors of behavior, such as motion sensors. In fact, the combination of short separation channels with accelerometer signals recorded on the head significantly improves the recovery of the task-evoked hemodynamic responses (von Lühmann et al., 2020). Similarly to superficial signal regression, physiological signals can be used as regressors within the GLM in order to regress out the contribution of systemic interferences from the fNIRS time series (Kirilina et al., 2013). Additionally, von Lühmann and colleagues (2020) have proposed a GLM-based method that implements temporally embedded Canonical Correlation Analysis (tCCA) to identify the optimal set of auxiliary nuisance regressors that best denoise the fNIRS signals from systemic contamination (von Lühmann et al., 2020).

Post- data collection strategies. Although it is highly recommended to adopt the methods mentioned above (Yücel et al., 2021), sometimes it is not possible to include short separation channels in the array due to a limited number of optodes, or there might not be physiological monitors available in the lab or it might be difficult to gather those measurements on certain populations (e.g., babies) or tasks (e.g., $PaCO_2$ while walking). The easiest way to approach this issue is to remove specific components from the fNIRS signals using digital filters (Scholkmann et al., 2014). Band-pass filters can be used to minimize the low frequency oscillations related to vasomotion regulations (usually <0.01 Hz) and higher frequency physiological noise such as respiration and heart rate (usually >0.2 Hz), while ensuring that the task-related hemodynamic components are preserved (Pinti et al., 2019). However, filters are not able to remove those components that overlap with the hemodynamic activity, such as Mayer waves or blood pressure regulations (Yücel at al., 2016). In such cases, one can use more advanced signal processing approaches to try to identify the physiological confounding signals from the fNIRS data directly. Some of these algorithms are based on the separation of global components from the local task-evoked components using principal component analysis (Zhang et al., 2016). This global mean removal method was demonstrated to have comparable performance to superficial signal regression in removing the non-neural scalp interferences (Noah at al., 2022) and can be very convenient when short separation channels or auxiliary monitors are not available (Note: it requires a covered area

size of at least 9 cm and that the area of activation is smaller than the probed head area (Zhang et al., 2016)). Similarly, other methods make use of Principal Component Analysis (PCA) or Independent Component Analysis (ICA) to identify and remove the extracerebral interference to the fNIRS signals (e.g., Virtaten et al., 2009), and have been reviewed in Scholkmann et al., (2014).

In Figure 5, we summarize the approaches that can be used to account for systemic interferences and aim to inform the planning of fNIRS neuroimaging experiments to maximize the robustness of fNIRS data. In particular, we use solid lines in the decision tree based for what we recommend to maximize data quality, that is to: (1) measure and match physiological changes across the various conditions of the cognitive task; (2) use several short separation channels around the head to capture the heterogeneity of scalp changes; (3) monitor physiological signals; (4) regress out the contribution of both extracerebral changes from short separation channels and systemic confounding factors from physiological recordings to the fNIRS signals. The flow chart mainly refers to continuous wave fNIRS instruments as these are the most commonly used and available to the market.

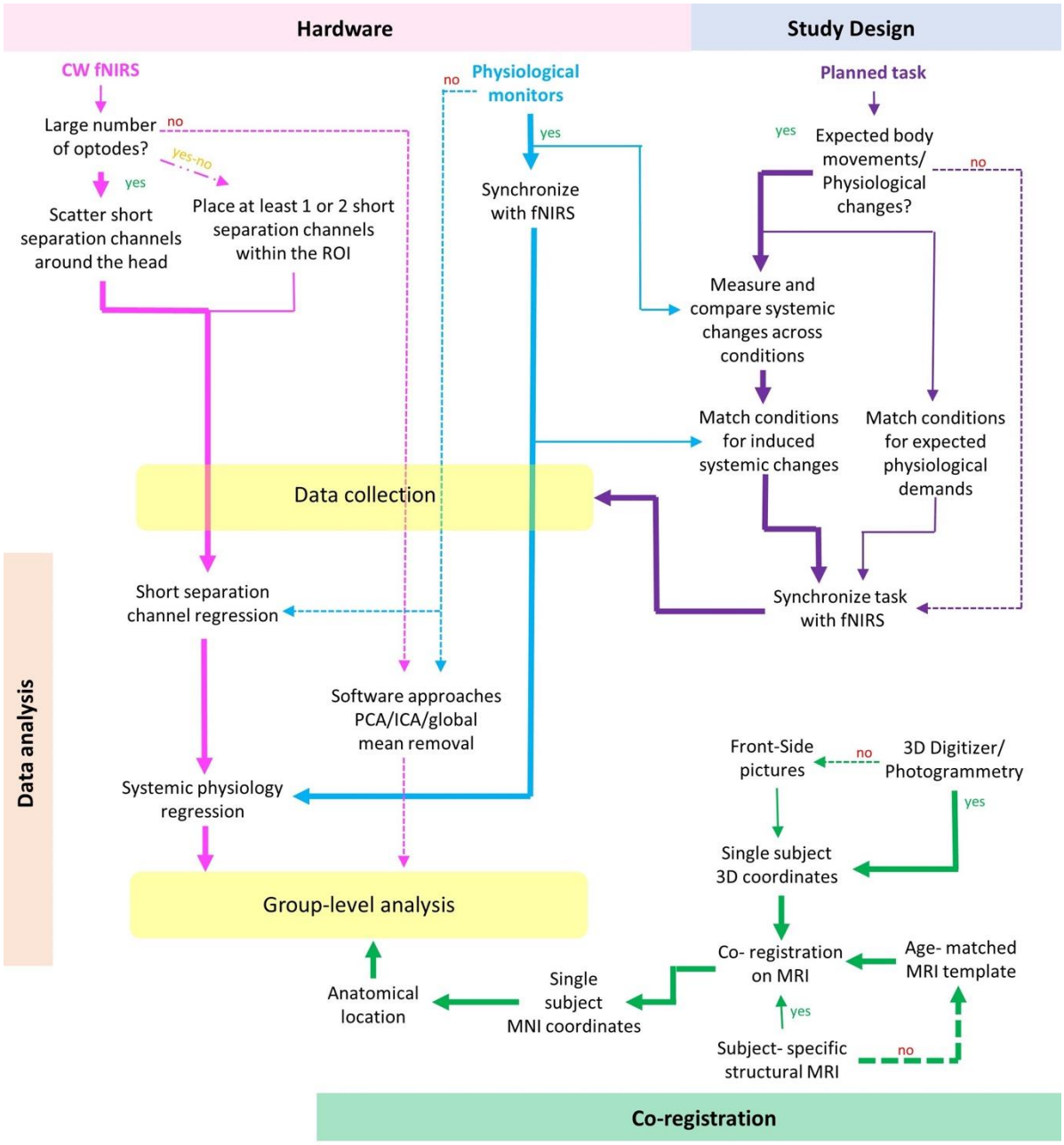

Figure 5. Summary flow chart to guide the robust acquisition and analysis of fNIRS data.

3.2. Co-registration of fNIRS channels onto underlying anatomy

One the major drawbacks of fNIRS as well as of EEG is the lack of anatomical information or anatomical structural images like MRI is able to provide. We make inferences about hemodynamic brain activity by placing sources and detectors on the scalp, but we do not know exactly the underlying anatomical location within the brain from which the hemodynamic changes we record are coming from. However, anatomical localization is

extremely important in functional neuroimaging in order to better interpret the results, compare them across studies, and to properly group the data across participants (Figure 5). Group analyses have generally the goal of assessing if a particular brain area shows significant hemodynamic changes across the group of participant; we often group channels by the 'channel-number' in an array but this does not mean that the each channel is measuring from the same brain region in all the participants. This also makes it challenging to compare results across studies or modalities (Wijeakumar et al., 2017). In Figure 6, we show an example of the challenge in channels localization in group analyses (Krishnan-Barman, 2012). Two regular optode arrays (Figure 6 A) were placed over left and right parietal / temporal cortex, with ideal cortical locations as shown in Figure 6 B.

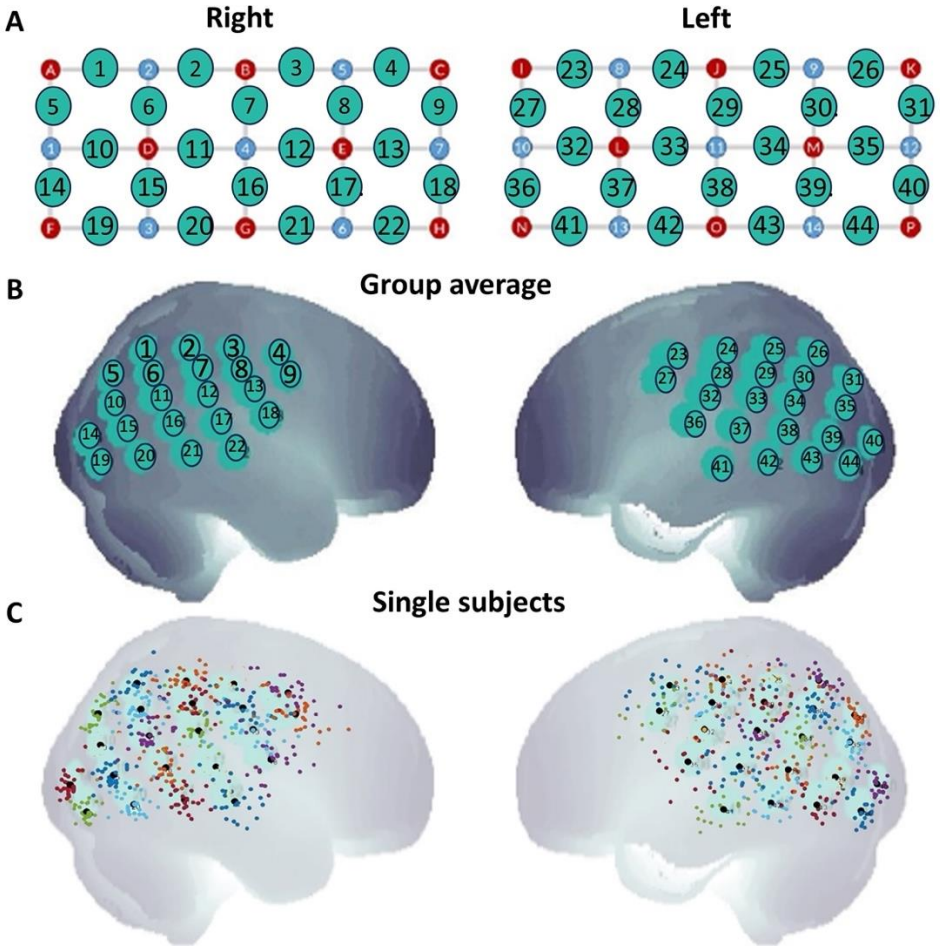

Figure 6. Example of channels localization in a group of 34 adults. Two rectangular grids (A) with 22 channels covered the right and left parietal / temporal cortex (B). Panel C shows the single-subjects channel locations: each dot represents the coordinate for one participant;

different colours are used for each channel. Substantial scatter in true channel locations due to individual differences in head size and shape can be observed.

The coordinates of the channels were recorded on each participant (N=34 adults) with a 3D digitizer. The single-subjects channel locations are show in Figure 6 C, with different colours for each channel. It is clear that there is considerable variability in true channel locations due to differences in head size between participants, with some channels of one participant overlapping the locations of different channel numbers of other participants. Therefore, if we were to group channels by their numbers, we would risk averaging together anatomically and functionally different regions. Normalisation to take these differences into account is hence important. We thus need to ensure that the data coming from the *same* brain subregion is pooled together across the cohort. This can be achieved with a two-step process: (1) reliable array placement; (2) identification of channels' anatomical locations. The overall goal is to find the correspondence between the NIRS sensors on the scalp and their underlying cortical location of where the fNIRS signal originates from (Tsuzuki and Dan, 2014).

Reliable array placement. In order to be more confident that we are grouping together anatomically homogenous fNIRS data coming from the same brain regions, we want to ensure that the fNIRS optodes are placed at the same locations across all participants. In addition, it has been recently shown that a non-standardized cap placement can affect the test-retest reliability of fNIRS data (de Rond et al., 2023). To account for differences in head size and shape between individuals, one recommendation is to place the optodes respect to the EEG 10-20 (or 10-10, 10-5, etc) electrode placement system (Yücel et al., 2021). For instance, optode X or channel X can be located onto one of the landmarks (e.g., Pinti et al. (2015) placed channel 9 onto the Fpz point and channel 9-channel 8 line aligned to Fpz-Fz). This method is quite straightforward and easy to implement; the 10-20 points are often already marked onto the neuroimaging caps (e.g., Easycaps). We can then link those locations to the underlying cortical anatomy (Jurcak et al., 2007). However, fNIRS optodes are often housed in rigid probe holders that do not adapt to different head sizes and can misoriented on the head (e.g., slight rotations from the target positioning). Stretchy caps can also be used, which are elastic but are also prone to misplacements, e.g. tilting or rotations. If these errors are larger than the fNIRS spatial resolution, e.g. misplacements larger than ~2 cm, we introduce variability in the cortical locations each channel is measuring from (Wijeakumar et al., 2017). Therefore, good

practices in probe positioning should be followed by anatomical co-registration practices (see below).

Identification of channels' anatomical locations. In order to account for between-subjects variability in channels location, either due to different head sizes or probe misplacements as described above, it is recommended to co-register channel positions onto a common brain space. Co-registration is the process by which we estimate the anatomical location of a sensor on the scalp (or the point from which our functional signal is coming from) by aligning two sets of information: the 3D coordinates of the scalp locations and an anatomical image (Chiarelli et al., 2015). Several common anatomical platforms have been developed to perform population-based neuroimaging analyses. One of the most popular is the Montreal Neurological Institute (MNI) brain template obtained by averaging 152 brains in order to provide a better representation of general population's anatomy (Tsuzuki and Dan, 2014).

There are many possible approaches to co-registration. One option is that the cortical MNI coordinates of a set of fNIRS channels can be potentially obtained by relating their position to the 10-5 positions on the scalp (Jurcak et al., 2007); however, it might be difficult and time consuming to make the link accurately, especially in whole-head setups. A second option uses a tracking system to capture the 3D coordinates of the channel locations on the head and transform them into the coordinate system of an anatomical image (i.e., the standard MNI; Aasted et al., 2015) usually by aligning fiducial points identified in both coordinate systems. Fiducial points are anatomical landmarks identified from the 10-20 electrode placement system (e.g., Nasion, Inion, Right and Left PreAuricular points, Cz). At least three non-collinear points are needed to perform the co-registration properly. The 3D coordinates of the scalp locations and the fiducial points can be recorded using a 3D digitization system (Singh et al., 2005; Aasted et al., 2015). Very popular are 3D magnetic digitizers which can give high resolution data as long as the immediate environment is free of electromagnetic interference from computers / cell phones. Some systems also require participants to stay very still during the digitization process. Recently, photogrammetry-based methods have been proposed which often do not require expensive extra hardware beyond a smartphone camera (Jaffe-Dax et al., 2020; Hu et al., 2020; Erel et al., 2021; Mazzonetto et al., 2022). For example, the method presented by Jaffe-Dax et al. (2020) reconstructs the 3D model of the head based on computer vision algorithms (structure from motion), accounting

for subject's movements. However, these can require careful calibration for each new optode configuration.

Anatomical images can be the subject-specific structural MRI scan (giving the best co-registration accuracy) or a volumetric MNI template. In the former case, scalp locations are obtained in the subject-specific MRI space and then further co-registered onto the common MNI template for group comparisons. However, MRI scans can be expensive and not feasible in certain populations, which defeats the purpose of using fNIRS. Methods have been developed to co-register the subject-specific 3D locations on the scalp to the MNI template directly (i.e., stand-alone registration; Singh et al., 2005). The co-registration of fNIRS functional data onto anatomical templates requires a further step, i.e. the projection of the co-registered scalp coordinates in the MNI space to the underlying cortex (e.g., Okamoto and Dan, 2005). Atlases can then be used to assign the anatomical label to each channel. The stand-alone co-registration is one of the most common approaches used in the fNIRS community; we have summarized the main steps of what this procedure involves in Figure 7.

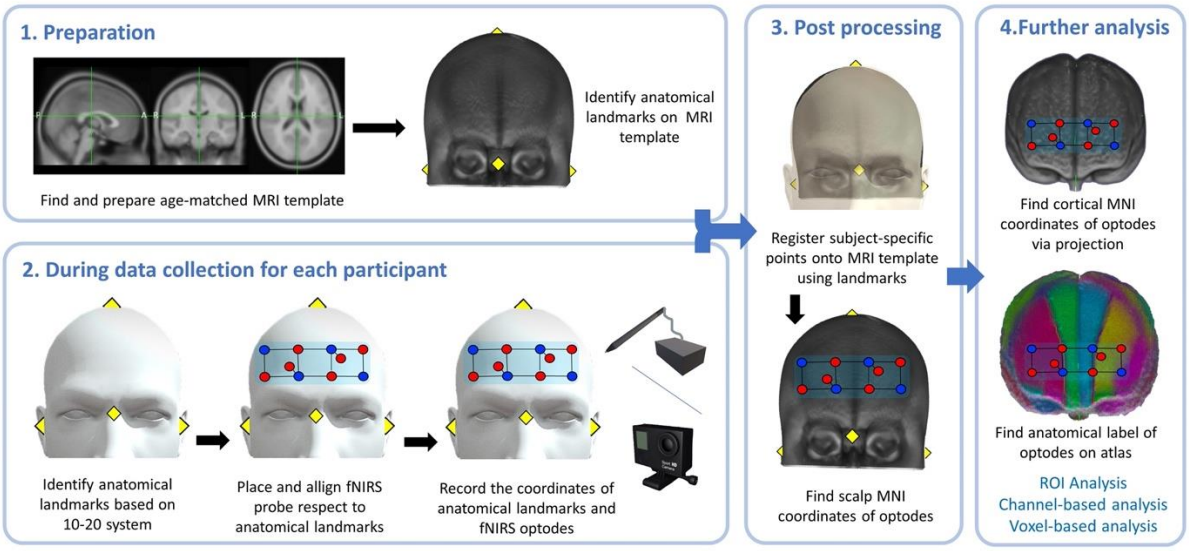

Figure 7. Main steps required to perform an accurate anatomical co-registration of fNIRS functional data.

A virtual co-registration method has been proposed as well (Tsuzuki et al., 2007) that does not require a 3D digitizer nor the MRI scan of the participant. Briefly, it simulates the probe position on the scalp based on the deformation of the holder on the subject's head by

reproducing it on a head model surface. However, this method relies on a highly accurate probe placement on the head and definition of the probe shape and deformation.

Once the cortical locations of the channels are obtained in the common MNI space for all participants, group level analyses can be performed using three different approaches: (1) channel-wise; (2) ROI analyses and (3) voxel-wise analyses. Channel-wise analyses assume that the channels positions are the same across the group and that the scalp locations have the same correspondence to the underlying cortical regions (Collins-Jones et al., 2021). However, these assumptions are often violated due to different head sized and anatomy, and misplacement of the probe, especially in whole head measurements (Tsuzuki and Dan, 2014). One way around this is to refer the MNI coordinates of the channels to anatomical atlases, such as AAL or LPB40; channels can then be grouped across the cohort based on the anatomical label rather than their number or channels belonging to the same anatomical regions can be averaged together in regions of interest (ROI).

Alternatively, voxel-based approaches have been proposed. These consist in interpolating the discrete channel-based functional data into continuous topographic images (Ye et al., 2009; Wijeakumar et al., 2017). Voxel-based analyses allow researchers to account for differences in channels locations by overlapping anatomically-corresponding voxels. However, between-subjects variability in head shape and circumference may lead to interpolated images of different size and orientation. Therefore, the borders of the images may not overlap across the group and only the central parts should be considered. In addition, interpolation introduces spatial correlation among the voxels that require the p-value to be adjusted for multiple comparisons (Tsuzuki and Dan, 2014). A similar voxel-based approach has been proposed by Tak et al. (2016) but images are interpolated on a canonical scalp surface rather than on the cortex. It is also worth mentioning that when voxel-based systems project the fNIRS data into small voxels (e.g. 1 mm cube), the resulting data has a much finer resolution than the actual fNIRS system (2-3 cm resolution) (Tsuzuki and Dan, 2014). This means that the voxel-based functional data is *interpolated* data and not strictly *real* measured data. Therefore, small activation clusters below the fNIRS resolution level should be carefully considered and appropriate corrections for multiple comparisons should also be used. Voxel-approaches which use larger voxel sizes may be more appropriate.

3.3. General considerations and open challenges

The previous sections provided an overview of the available methods to overcome some of the challenges related to the use of fNIRS in cognitive neuroscience experiments. However, some other challenges are yet to be addressed.

One is how these issues apply to developmental populations. Everything discussed so far as well as the proposed tools have been investigated, tested, and validated in adult populations. In terms of anatomical localization of fNIRS data, there are two main challenges to tackle. First, there are no co-registration algorithms or ready-to-use co-registration software developed for younger populations. The co-registration methods discussed in Section 2.3 are available in software platforms such as NIRS-SPM (Ye et al., 2009) and Homer2 (Huppert et al., 2009), but only incorporate adult MNI templates and atlases. AtlasViewer within Homer2/3 (Aasted et al., 2015) provides some head models for babies from 29 weeks to 44 weeks post-menstrual age, but none are yet available for older children. The co-registration work by Lloyd-Fox and colleagues can be used to identify subregions within the frontal and temporo-parietal areas in 4-to-7 months old infants (Lloyd-Fox et al., 2014). However, age-matched head models are important as light diffusion through the head changes with age (Fu and Richards, 2022) as well as the scalp-to-cortex correspondence (Fu and Richards, 2021). There are databases available with average MRI templates for different ages such as the Neurodevelopmental MRI Database (Richards et al., 2015). However, such images should be manipulated by the users to create appropriate and segmented volumetric masks to make them compatible with the current open-source analysis toolbox or their own co-registration code, requiring technical skills.

Second, gathering the 3D coordinates of fNIRS sensors on the scalp can be challenging in young populations, especially with 3D magnetic digitizers. Given the importance of performing the anatomical co-registration to increase the robustness of the results, this step should not be skipped in children as often happens. We recommend using the novel photogrammetric approaches that have been demonstrated to be suitable for infants (Jaffe-Dax et al., 2020; Uchitel et al., 2023) or extract the locations from front and side pictures of the probe on the head (Lloyd-Fox et al., 2014; Collins-Jones et al., 2021).

Whilst some guidelines have been set in terms of the analysis and acquisition of fNIRS data in infants (Gemignani and Gervain, 2021; Yücel et al., 2021), very little is known about

the impact of scalp interference and systemic changes in infants and children. To the best of our knowledge, only Emberson et al. (2016) has investigated the usefulness of superficial signal regression in infants and showed that this does not make a significant difference in the group level statistics. However, this finding may be specific to the particular cognitive task used and for that age range; shorter short separation channels might also have been more appropriate for infants (Brigadoi and Cooper, 2015). In Figure 8, we show an example of fNIRS signals from a short separation channel recorded on an adult (54 years old; A) and a pre-schooler (5 years and 8 months old; B). The short separation channel was placed on the prefrontal cortex, with a source detector distance of 1 cm, and both participants were performing an inhibitory control task while standing and moving in a Virtual Reality (VR) environment (https://osf.io/wyp4s/).

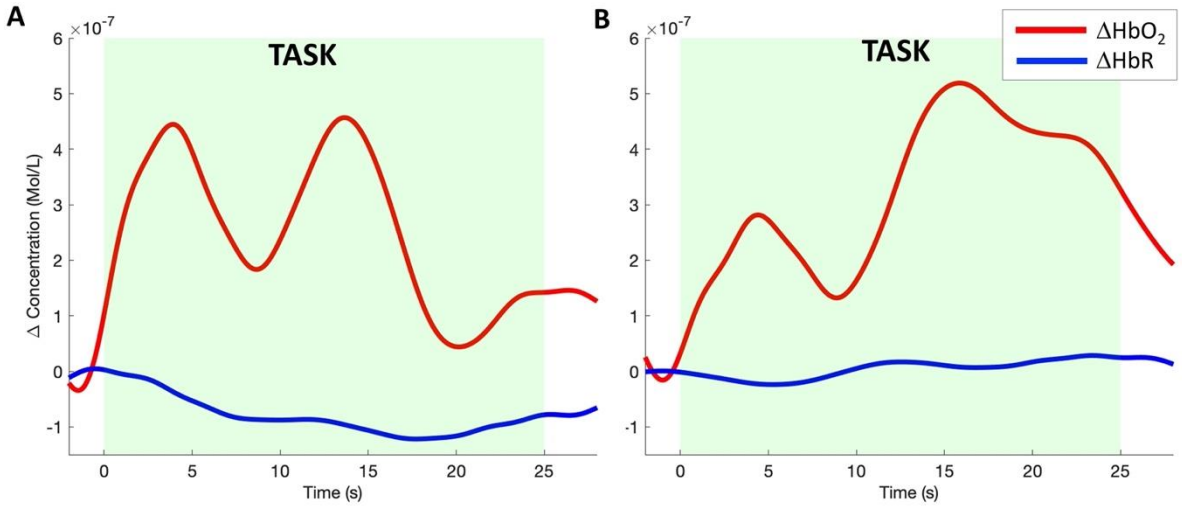

Figure 8. Mean scalp hemodynamic changes recorded from a short separation channel on the frontal lobe on a 54 year old adult (A) and a 5 years and 8 months old child (B) during an inhibitory control task in a VR environment. Red and blue lines represent the mean responses in HbO_2 and HbR respectively. The task period is indicated by the green shaded area. This illustrates similar profiles in adults and children.

Similar changes in brain hemodynamics, with increase in HbO_2 of approx. 0.5 μMol/L and decrease in HbR of approx. -0.05 μMol/L similar to those occurring during functional activation, can be observed in both participants; typical 10 s modulation components related to Mayer waves can also be seen in both cases. This suggests that extracerebral changes might be a confounding factor in fNIRS data recorded in young children, just as in adults, especially

when the participant is freely moving. Therefore, further research is needed to investigate the impact of scalp interference and physiological changes in younger populations, and to develop age-appropriate tools to account for them. In this framework, one should keep in mind that a SPA-fNIRS approach as used by Zohdi et al. (2021) might not be feasible in infants and children, and hence the platform should be restricted to the most informative physiological parameters as well as to those that can be monitored with minimally obtrusive and child-sized sensors. The same principles apply to short-separation channels and whole-head high density measurements; the overall weight of the probe should be taken into consideration when increasing the number of optodes and further investigations should assess whether there are significant scalp changes in young populations and, if there are, whether these are heterogenous and determine the minimum number of short separation channels needed to account for that.

The need for a larger number of short separation channels to capture heterogeneity also applies more generally to neuroscience experiments, especially those that involve large movements and mobile instrumentation where confounding factors might be more pronounced. Hardware-related limitations are discussed in Section 2.3. Guidelines are not yet established regarding the optimal number and the optimal placement of short separation channels. Another issue is the lack of off-the-shelf multimodal SPA-fNIRS platforms with synchronized data streams and data formats compatible with open-source analysis software, and that these require advanced technical skills from the user to handle the synchronization of individual pieces of equipment and output files manipulation (Vidal et al., 2023).

In addition, there is not yet an agreement on whether channel-based or voxel-based analyses of fNIRS data is most appropriate or in which situations one is better suited than the other. Some research teams use both approaches to determine the final group results (Hirsch et al., 2017). Previous work by Collins-Jones and colleagues (2021) has compared channel-space analysis with image-space analysis of fNIRS data, particularly in the context of infants where head size and shape variability are more pronounced than in older cohorts. They have shown that the error in array positioning has the largest impact on the difference between group-level channel-wise vs voxel-wise analysis results and that the effect of array misplacement and of different head sizes decreases as the group sample size increases (i.e. the two analyses yield more similar results for larger samples size). Based on these first investigations, we recommend following the procedures outlined in Section 3.2 to obtain an

array placement as reliable as possible across the participants, to collect subject-specific information on sensors locations to perform a more accurate anatomical co-registration, and to use an adequate sample size.

Finally, the objective assessment of data quality is a hot topic in the fNIRS community at the moment. The signal quality of channels can be deteriorated by instrumental or environmental noise, a poor optical coupling, or by the presence of significant motion artifacts. In these cases, channels producing low quality data should be excluded from the analysis to avoid misinterpretation (Hocke et al., 2018) and unreliable results. The identification of noisy channels is commonly performed by visually inspecting the signals, typically including assessment of the presence of the heart-beat component that indicates a good coupling between the optodes and the head or looking for saturated or very low intesity signals (e.g. flat lines). This, however, is highly subjective and can become very time consuming as the number of channels to examine increases. Other methods attempt to be less subjective and rely on the identification of too high (saturation) or too low (poor coupling) signals' amplitudes based on thresholds of raw intensity values (e.g., Collins-Jones et al., 2021) or on their coefficient of variation (e.g., Frijia et al., 2021). More recently, novel algorithms have been developed that evaluate data quality in time windows over the duration of the recording, allowing detection of signal quality changes over time and potential exclusion of trials rather than the whole channel. For example, the Signal Quality Index (or SQI) is a three-step process (identify very low-quality channels; identify very high-quality channels; signal quality rating) and provides a signal quality index based on the strength of the optical coupling in a non-binary numeric scale from 1 to 5 (Sappia et al., 2020). QT-NIRS (https://github.com/lpollonini/qt-nirs) instead is based on the Scalp Coupling Index (SCI) and the Power Spectral Peak (PSP) computed in time windows that reflect the strength of the heart beat component in the signal, which is used as the criterion for good data quality. Channels with a SCI and PSP below a certain threshold for a certain number of time windows, are marked as poor quality (Hernandex and Pollonini, 2020). Although progress has been made in improving the automation of signal quality assessment, these methods still rely on user-defined or empirically established thresholds and are not yet standardized.

4. Opportunities and open challenges of fNIRS in cognitive neuroscience

The last 30 years have seen tremendous developments in fNIRS hardware and data analysis techniques that have allowed the research community to drive the cognitive neuroscience field forward and to establish fNIRS as valuable tool to understand brain functioning (https://www.spiedigitallibrary.org/celebrating-30-years-of-fnirs). Every neuroimaging technique has advantages and disadvantages for different applications, and some of the attributes of fNIRS (e.g., portability, affordability, tolerance to body movements, multimodal integration, a good temporal-spatial resolution compromise) make this neuroimaging modality particularly suited than others for certain applications. In particular, with the recent advent of wearable and wireless fNIRS technologies (see Pinti et al. (2018) for a review), cognitive neuroscientists have the opportunity to expand investigations from the typical laboratory settings into contexts with higher prima facie ecological validity, gaining new knowledge of the neural mechanisms underlying cognition in new domains. However, like any new and evolving application, alongside the opportunities there are also challenges that need to be addressed and limitations that need to be explored to understand how far fNIRS can be used to *reliably* image brain activity in unconventional settings. In this section, we provide an overview of the current status of using fNIRS outside the typical laboratory while discussing the ongoing efforts and open challenges that are yet to be overcome.

4.1. Studying brain functioning in ecologically-valid settings

Functional brain activity is traditionally studied in the laboratory, very often with participants sitting in front of a computer screen and exposed to stimuli arranged in a "block-designed" fashion. Such paradigms include task periods separated by rest periods, typically 10-30 of seconds long, which are repeated multiple times throughout the experiment. In each task block, stimuli of the same type are presented several times to elicit stronger hemodynamic changes; rest periods are supposed to not recruit the same cognitive processes of the task and serve as an interval to allow the slow hemodynamic response to go back to its baseline value and to separate the task-evoked response to each stimulus type. However, rests

should be kept as short as possible as it is unknown what the participant is really doing or thinking during that time, and it is very difficult to get a real baseline with no cognitive engagement. It is more appropriate to use a low-level control condition instead of e.g. blank screens with a fixation cross when investigating higher level cognitive functions (Henson, 2015). Because of the slow nature of the hemodynamic response, block-designs are very popular in neurovascular coupling-based neuroimaging, offering a greater signal-to-noise ratio with stronger brain responses than event-related designs (Henson, 2015). The latter involves the presentation of discrete stimuli of different types individually, sometimes with little or no rest between trials. Event-related experiments are more flexible, but they provide a lower signal-to-noise ratio, require a larger number of events, and need careful consideration of study duration and the inter-trial interval.

Although these designs have always been an invaluable tool to isolate and study specific cognitive processes in a controlled way, they are clearly not a good model of how we function in our everyday life or at least for most situations. For example, if we are having a conversation with another person, we do not stop and resume the interaction every 20 seconds. Therefore, if we want to gain a better understanding on how the brain supports the cognitive abilities that we use in our daily lives, we need to move beyond the classical experimenter-driven tasks toward experiments with a higher degree of ecological validity.

Ecological validity refers to the degree to which the task performance reflects or relates to “real-world” behavior (i.e., typical behaviour or situations outside the lab). If the goal of research is to discover principles or laws of nature, then the scientist often wants to be sure that their discoveries and meaningful beyond the walls of the laboratory. Thus, the issue of ecological validity becomes key for many scientists investigating mind-brain relations. This is especially the case where equivalency between the situation presented to the participant in the lab and situations outside the lab is not obvious. There are two principal components to “ecological validity”. The first is the *predictive validity* of the task. The second is the *representativeness* of the experimental situation (Burgess et al., 2006). The predictive validity component refers to how robust the findings from the experiment will be once they are applied to “real life” situations outside the lab. The representativeness component refers to how closely the experimental situation presented to the participant mimics the situation in real life that is one is trying to understand. It is generally assumed that increasing representativeness is likely to also increase predictive validity. So to increase the ecological

validity of the task, the development of new experimental designs that are closer to real-life demands becomes crucial. There are two main approaches that can be used to answer this need: a *naturalistic* approach and a *real-world* approach. At first glance, these terms may seem to describe the same experimental framework, but they actually refer to two different concepts.

Naturalistic neuroimaging aims to implement experimental paradigms that closely mimic or approximate aspects of the real-world in an experimental setting and that capture the complexity of real-life situations with a certain degree of experimental control. They can be thought as a mid-point between the tightly controlled laboratory environment and the fully unpredictable real world. As an example, Greaves & Pinti and colleagues (2022) have taken a naturalistic neuroimaging approach to investigate the sense of self in actors while performing in the theatre. The naturalistic aspect was achieved by choosing: (1) the theatre, in order to have a realistic experimental setting where external distractions can be controlled and avoided; (2) actors as participants, as they are able to repeat the same actions with the same socio-emotional content while engaging with another actor, giving reproducible data; (3) embedding events of interest ("name calling" in this case) within the normal acting routine, preserving their normal rehearsal; (4) wearable fNIRS, to allow actors to move as they would normally do. This combination increased the degree of ecological validity compared to classical computer-based cognitive task and at the same time, the timings of the presentation of "name calling" events were under the control of the experimenter, enabling a robust task-evoked analysis of fNIRS data. Similar efforts in increasing the ecological validity while preserving some experimental control are being undertaken in the field of developmental neuroscience, where computerized stimuli can replaced by live demonstrations. For instance, this allowed Lloyd-Fox et al. (2015) to explore the brain network responding to natural infant-directed communication in infants during live face-to-face interaction with an adult. However, as recently discussed by Dopierala and Emberson (2023), although more naturalistic, this sort of paradigm is still bounded to structured designs, with fixed task-rest periods, limiting the infant's spontaneous engagement compared to truly naturalistic situations like free play.

Real-world neuroimaging refers to image brain activity in the actual real-world, where participants are immersed in their everyday life situations with minimal or no experimental control, encompassing the full unpredictability, richness, and dynamics of daily life demands. An example of real-world fNIRS study is the one by Burgess et al. (2022): participants were

asked to perform a prospective memory task while walking in the streets of London with no particular preparation of the environment and their own pace; in particular, they were asked to fist-bump a confederate (social cues) and parking meters (non-social cues) while counting the number of other items (e.g., doorbells) around the square. fNIRS data were recorded over the prefrontal cortex and results showed that medial PFC was strongly involved in maintaining both social and non-social intentions, while lateral PFC had higher activity when maintaining a social intention than a non-social one. Similarly, McKendrick and colleagues (2016) had participants navigating around a college campus with either a hand-held display or an augmented reality wearable display while performing visual perception and auditory working memory tasks. Different changes in brain hemodynamics in the prefrontal cortex were found in relation to the amount of workload induced by the two displays, with increased prefrontal activity with the augmented reality device and decreased activity with the hand-held display (McKendrick et al., 2016). The feasibility of fNIRS for naturalistic experiments was also shown while playing table-tennis, or playing musical instruments (piano and violin) or while continuously performing daily activities (Balardin et al., 2017).

In summary, naturalistic and/or real-world task designs in neuroimaging experiments provide valuable insights on how the brain works in daily life situations while at the same time they can also increase participants' engagement and hence data reliability. This becomes particularly important also in developmental populations, where engaging tasks are critical to maximize children's compliance to the task. On the other hand, ecologically-valid paradigms can introduce novel complexities that are discussed below.

4.1.2. Challenges of naturalistic and real-world neuroimaging

Naturalistic and real-world experiments differ in the degree of experimental control and of ecological validity. The former only mirror and simulate daily life situations in a relatively controlled setting; the latter are embedded into the real daily life situations with minor or no experimenter's manipulation. Whilst the higher ecological validity of real-world paradigms provides the optimal model of real-world behavior, the higher experimental control of the naturalistic approach strongly facilitates data collection and analysis. However, they both share common challenges where the severity is often proportional to the degree of ecological validity and the degree of structure imposed by the task. These challenges concern

data quality, task design, individual differences, data analysis and interpretation of findings. In discussing these issues below, we mostly refer to the work by Burgess et al. (2022) as we had first-hand experience in encountering and trying to address some of these problems.

Data quality. Naturalistic and real-world experiments can involve much larger body movements compared to classical static computer-based studies. This can have impact on fNIRS signals' quality in two ways: (1) they may increase the likelihood and the strength of motion artifacts; (2) they may lead to larger systemic interferences (see Section 3.1). Although these have not been studied systematically yet in the context of real-world neuroimaging, it is reasonable to think that these sources of noise may have a higher impact on fNIRS signals recorded in these contexts. For example, when the experimental situation becomes more realistic, also the behavior becomes more realistic; this can evoke greater emotional responses that in turn can lead to faster and greater range of movements and physiological responses. For instance, greater facial expressions may be made by the participants that in turn can lead to frowning-relating motion artifacts (Yücel et al., 2014). In the study by Greaves and Pinti et al. (2022), eyebrow raising was so frequent that the channels within medial PFC were corrupted by artifacts and had to be excluded. Moreover, preliminary investigations showed that performing a cognitive task while walking in the streets can lead to changes in heart rate and respiration of 15-20 BPM (Figure 9 A, Pinti et al., 2015) that are larger than those typically observed when participants are sitting inside the lab (e.g., approx. 1-3 BPM in Zohdi et al., 2021, Figure 9 B). In these contexts, it becomes even more important to complement fNIRS recordings with measures of systemic physiology (i.e., SPA-fNIRS), short-separation channels, and measures of head/body motion (e.g., accelerometers). These can help in explaining the variance in the fNIRS signals and enable a more robust recovery of task-evoked brain activity (von Lühmann et al., 2021).

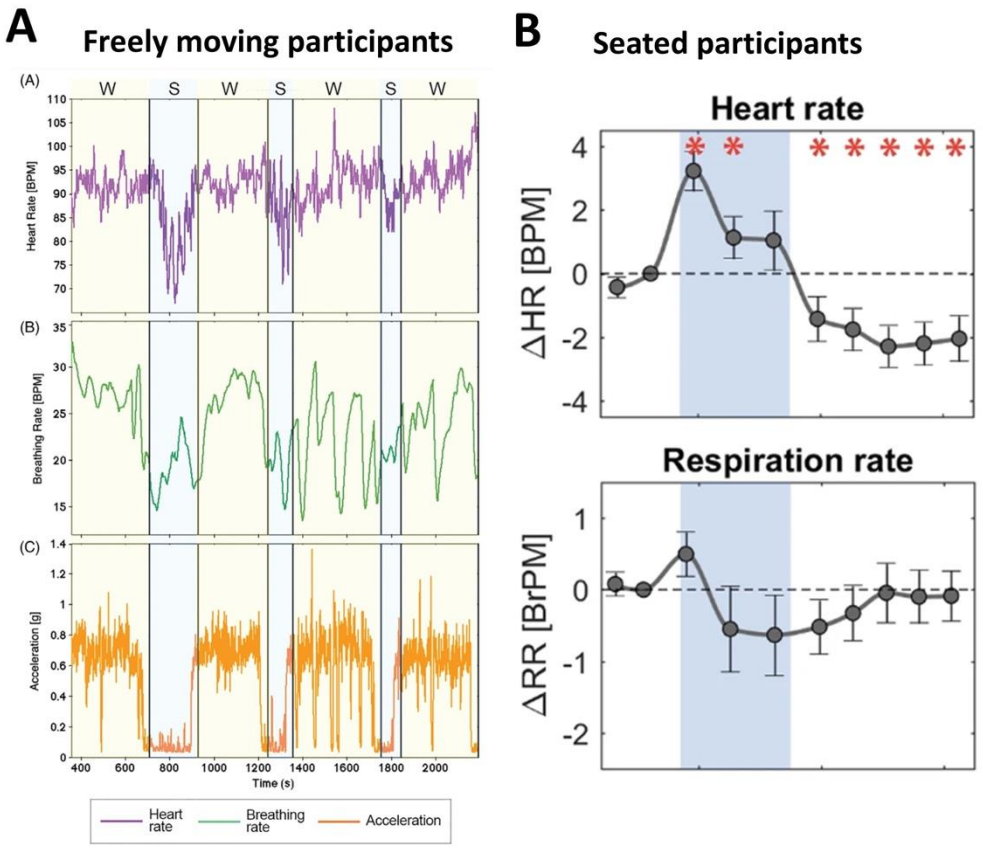

Figure 9. Examples of physiological changes recorded during fNIRS experiments. Panel A refer to data from freely moving participants (top: heart rate, middle: respiration, bottom: acceleration; reproduced with permission from Pinti et al., 2015) and panel B to participants passively exposed to colored light while sitting in the lab (top: heart rate, bottom: respiration; adapted with permission from Zohdi et al., 2021). These data show that much larger physiological changes can occur when people are freely moving compared to when they are sitting.

Task Design. Possibly the hardest challenge to overcome is the design of new paradigms that are simultaneously less structured (i.e., experimenter-directed) in order to resemble real world contexts and yet are controlled enough by the experimenter to enable the robust assessment of functional brain activity. Conventionally, experimental tasks are required to include repeated task blocks or events with the goal of averaging the fNIRS signals across repetitions to increase the SNR and the statistical power. Conditions are often separated by rest periods with minimal cognitive involvement (e.g., a fixation cross on a black screen) to evaluate task-evoked hemodynamic changes from baseline. These designs are not necessarily well-suited for monitoring brain functioning during daily life activities

continuously, where we may be particularly interested in single or specific events. Current analytical approaches are optimized to work with repeated task designs, although some efforts are currently being made to develop single-trial classification methods of brain activity (e.g., von Lühmann et al., 2020). To ensure that the fNIRS data are analyzable and a clear haemodynamic response can be detected above the background noise, there are a few strategies that can be incorporated as one increases the ecological validity of the task. An obvious one is to incorporate multiple events when possible. For instance, Burgess et al. (2022) chose objects that were naturally present at multiple locations within the environment (e.g., 6 parking meters or opening hours fixed to building that participants were required to find). However, it might be the case that fewer events of the same type might be required in naturalistic experiments, where the increased ecological validity of the task might elicit purer and clearer hemodynamic responses compared to the more artificial laboratory experiments. This is something that has yet to be investigated systematically. It is also very important that the events of interest can be easily recognizable among the abundance of other things happening (e.g., in the example from Burgess et al. (2022) parking meters were clearly identifiable), so that the particular response or behaviour can be isolated from the rest. The timings at which these events happen should be precisely monitored (e.g., via head camera video recordings) to be able to effectively time-lock the analysis to the occurrence of the event. In addition, in order to account for motion-induced confounding components, it is important to include appropriate baseline conditions, such as those involving the same level of motion but not the cognitive process of interest (e.g., the walking baseline condition in Burgess et al. (2022)). When participants are left to perform the task at their own pace, the timeline of events becomes unpredictable; it is crucial to enrich the setup with behavioral monitors, such as eye-tracking, GPS, video cameras, to precisely record the experimental session with high level of detail. However, events can sometimes still be difficult to identify. As shown in Pinti et al. (2015), it is not always true that the functional response of the brain is time-locked to the observed behavior. For instance, with the protocol used by Burgess et al. (2022), brain activity in some participants started before they reached the target of interest, while in others it was delayed. Therefore, data-driven approaches for the identification of the true occurrence of events triggering brain activity might be highly beneficial to account of the unpredictability of the response and the between-subject variability. We have developed an algorithm of the Automatic IDentification of functional Events (AIDE) that identifies the task-

evoked hemodynamic response directly from the fNIRS signals with no a-priori knowledge of the timeline of the stimuli (Pinti et al., 2017). We provide an overview of its functioning in Pinti et al. (2018) and an example of application in Burgess et al. (2022). As discussed above, it is critical to combine methods like AIDE with measures of behavior to accurately link the identified functional events to the corresponding specific actions during the experiment. Finally, when designing experiments that require participants to move, the duration of the experiment should be taken into account, since some of the fNIRS caps, especially those covering the prefrontal region, might become uncomfortable after a while, and participants may also experience physical as well as mental fatigue, leading to a loss of participant engagement in the task and the physiological responses, which may impact the occurrence/identification of the haemodynamic responses.

Individual differences. Compared to highly controlled lab-based experiments, in more naturalistic and multisensory tasks participants might have the opportunity to carry out the task at their own pace. The distinction between self-paced and experimenter-based tasks has a long history in psychometrics, and changing from one format to another, even when other characteristics are held the same, can affect behaviour and systemic reactions to the task (e.g., Steptoe et al., 1997), and very probably in the brain. For instance, there is good evidence that rostral PFC (BA10) is involved with dealing with open-ended (aka "ill-structured") cognitive tasks, which are usually inherently self-paced in format (Burgess and Wu, 2013), and the degree of open-endedness of a task will vary with participants' experience of it. So, using a more naturalistic self-paced format for your experiment can by itself potentially can lead to between-subject variability in the fNIRS data due to different strategies in performing the task, different cognitive processes being recruited at different times, the open-endedness of the task, or differences in participants' prior experiences that can influence task performance. Individual differences can also arise in terms of systemic changes related to e.g., the cardiopulmonary fitness level of the participant, different blood flow regulation mechanisms, activities performed prior the experiment, etc. It is therefore valuable in this case to track, if possible, with high resolution participants' behavior and physiology and potentially include self-reports or questionnaires to account for between-subjects differences as covariates in the analysis. Note also that the systematic study of individual differences in cognition typically requires large sample sizes so studies must be planned appropriately (Lakens and Evers, 2014).

Data analysis and interpretation. Data analysis and task design are closely linked to each other. The vast majority of extant neuroimaging experiments employed block- or event-related designed paradigms because the analysis methods currently available such as block-averaging or GLMs or preprocessing algorithms for signal denoising are able to provide a sufficient SNR for a robust recovery of the hemodynamic response only when repeated trials are available. At the same time, block- or event-related protocols have been well suited for the type of cognitive questions that we have been interested in so far and are compatible with the lab-based and wired hardware. With the development of wearable and wireless instrumentation, there are new questions that we can ask, creating the need for single-trial or more open-ended task designs. With this new need, there is also a new need for novel analytic methods that are able to deal with these data and/or instrumentation with higher sensitivity to brain specific signals. This need has not been fully met yet and new developments are currently needed that fully enable the reliable use of fNIRS in naturalistic settings. However, steps are being made to move towards that direction.

For example, more sophisticated methods based on the GLM have been developed that are able to combine multimodal behavioral and physiological signals within the analysis of fNIRS data and to identify the optimal set of nuisance signals representing confounding variables that best improve the recovery of the task-evoked haemodynamic activity (GLM combined with temporally embedded canonical correlation analysis (tCCA) by von Luhmann et al. (2020)); others can deal with motion artifacts and systemic interferences directly within the GLM with no additional hardware (Huppert et al., 2009) or, as introduced above, can identify brain activity directly from the data (AIDE, Pinti et al., 2017). Machine Learning can also be a powerful tool to improve data analytics. von Lühmann et al. (2020) for instance have shown that a GLM-based method with nuisance regressors (e.g., acceleration, scalp changes) increases the contrast to noise ratio of the data and provides more accurate features for classification of hemodynamic activity vs rest a the single-trial level (von Lühmann et al., 2020). However, these developments are still at the beginning and further improvements are expected in the future. There is a general consensus that multimodal investigations (brain+physiology+behavior) are the answer to naturalistic neuroimaging; multimodal data fusion algorithms will thus be necessary, possibly based on artificial intelligence for the automatic processing of information coming from different channels and the holistic evaluation of the neural-physiological-behavioral underpinnings of human cognition in the

real world. Accurate tracking of real-world behavior is particularly important here, and video-based systems like OpenPose and OpenFace may have an important contribution to make. As discussed above in respect to the study by Burgess et al. (2022), there are cases when it is difficult to manually link brain responses to behavior just by looking at video recordings (e.g., is the onset of brain activity when the target is in the field of view of the participant or when the participant reaches it?) or it can be nearly impossible to understand what an individual is thinking while not making any overt behaviors. Therefore, new sophisticated methods are needed to provide better automation and accuracy in the assessment of the cognitive processes involved in real-world settings.

4.2. Studying naturalistic social interactions with fNIRS-based hyperscanning

As fNIRS is a neuroimaging modality which allows free movements by participants, it is ideally suited to studying natural social interactions, where two or more people engage in a task together similarly to how they would do it in their daily lives. One way of studying the social brain networks is to measure the neural data from more than one person at the same time. This experimental configuration takes the name of 'hyperscanning' and aims to look at interpersonal neural synchrony (INS), a measure that expresses how the brain activity of one individual is similar to the brain activity the interacting partner. Data analysis usually involves the quantification of the similarity between the brain signals of the two partners (via e.g. correlation or wavelet coherence) as a proxy of the strength of the social interaction (Figure 10 A).

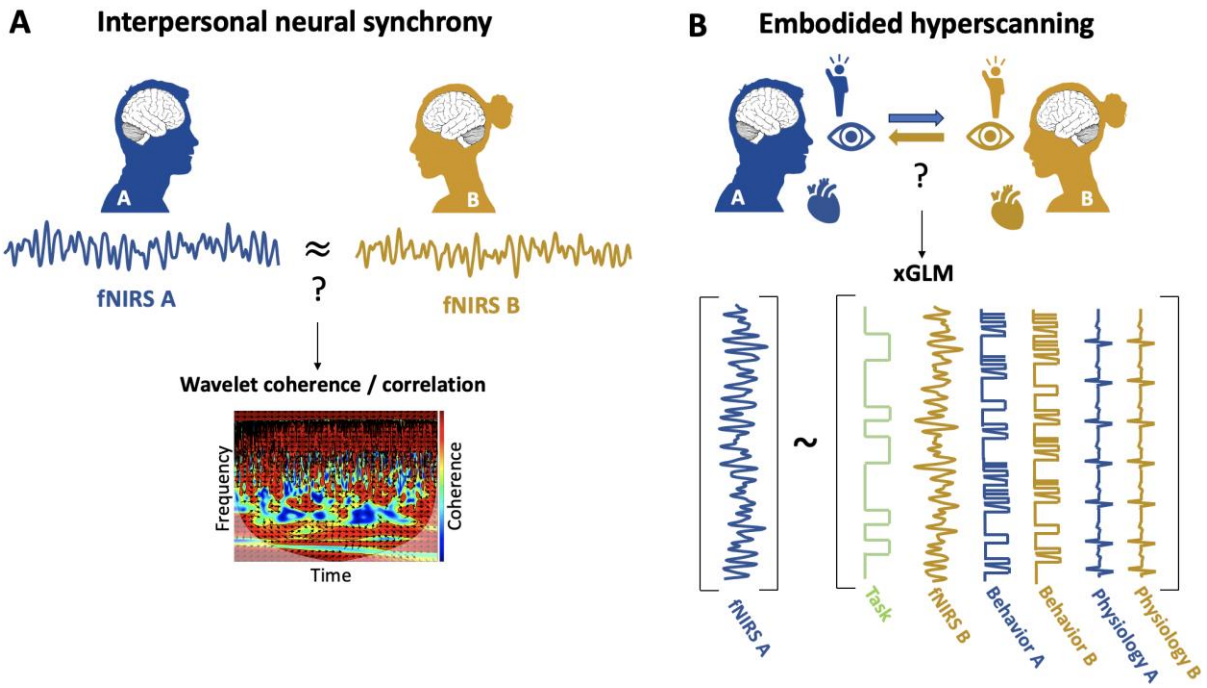

Figure 10. Example of approaches for hyperscanning investigations. Classical hyperscanning studies (A) quantify interpersonal synchrony in terms of similarity between the neural signals of two interacting partners, typically with correlation or wavelet coherence. Novel embodied approaches (B) combine multimodal information (brain, physiology, behavior) to explain the similarity in brain activities. xGLM can be used to investigate cognitive mechanisms of interpersonal synchrony, testing if brain activity of partner A can be explained as combination of predictors of self-behavior and physiology and partner's B behavior and physiology.

Purchasing multiple recording devices, recruiting multiple participants, and collecting hyperscanning data is a substantial piece of work; so what is the added value of these data? The rise in 'second person neuroscience' (Schilbach et al., 2013) is probably driving the increase in hyperscanning, and number of papers now make the claim that studying two people can tell us more than studying one person, because interaction is central to social cognition (De Jaegher et al., 2010). There are features that emerge in two person interactions that cannot be studied in traditional contexts with a single participant alone in a lab (Gvirts and Perlmutter, 2019; Hari et al., 2015; Hasson & Frith, 2016; Redcay & Schilbach, 2019). These include for example mutual adaptation in generating rhythmic actions (Konvalinka et al., 2010), audience effects (Hamilton & Lind, 2016) which arise when people respond differently to being watched by a person (Myllyneva & Hietanen, 2015), joint action when two people perform a task together (Sebanz et al., 2006) and communicative dynamics (Fusaroli et al.,

2012). Similarly, we know that different cognitive and neural mechanisms are engaged when participants believe they are interacting with a real human compared to a computer (Gallagher et al., 2002) and when they are being watched by another person (Izuma et al., 2010; Müller-Pinzler et al., 2015). Given these changes in the neurocognitive systems of a single person engaged in social interaction, it seems a logical next step that neuroimaging of both partners in a two-person social interaction should tell us more than just studying one person at a time.

Since the first NIRS hyperscanning study (Cui et al., 2012), over 80 data papers have been published and this area is growing rapidly. Many early papers provide a proof-of-principle that hyperscanning can be performed, and explore different methods for analysing the data. For example, the study by Cui et al. (2012) gave participants a computerised cooperation / competition task where they must coordinate the timings of their actions in the cooperation condition. Signals were recorded from prefrontal cortex and analysed using wavelet coherence, a popular multiresolution method to analyse hyperscanning data that measures the "correlation" between two time-series in the time-frequency space. Results showed more brain-to-brain synchrony (i.e., higher wavelet coherence) in the cooperation condition compared to competition. A number of studies have now replicated this with larger samples and additional contrasts, testing for effects of visual feedback and participant relationships (Pan et al., 2017; Reindl et al., 2018). For example, Baker and colleagues collected data from 111 dyads to test for sex differences in INS, finding more coherence in same-sex dyads than in mixed-sex dyads (Baker et al., 2016). These studies suggest that different patterns of social relationship between people can lead to different levels of brain-to-brain synchrony. A different method is to manipulate the level of interaction between two people more explicitly. For example, there is more wavelet coherence between two people who talk face to face compared to talking back-to-back (Jiang et al., 2012) and between two people who sing or hum face to face compared to back-to-back (Osaka et al., 2015). In a series of carefully controlled studies, Hirsch et al. have shown more coherence between two people making eye contact compared to looking at a picture (Hirsch et al., 2017) and between two people talking in dialogue compared to monologue (Hirsch et al., 2018). All these studies suggest that increasing the potential for a dynamic interaction between two people increases the wavelet coherence between their brains. Other groups have used tasks which involve a more dynamic face to face interaction. Using a multi-person approach, Fishburn et al. (2018)

asked participants to complete tangram puzzles in groups of 3 and found more coherence in prefrontal cortex between a pair of participants working together on the same puzzle, compared to the third person in the room who was working alone (Fishburn et al., 2018).

Overall, previous studies have made it clear that fNIRS hyperscanning is feasible and have provided examples of dual-brain analytics as well as the types of questions that can be investigated. However, to move the field onwards, it is important to consider what hyperscanning data can tell us at a more fundamental level and overcome some of the open challenges related to task design and results interpretation.

4.2.1 Methodological challenges in fNIRS hyperscanning and the way forward

The hyperscanning studies mentioned above show that, when two people are engaged in an interaction, there is an increase in 'coherence' or 'synchrony' between the two brains (or INS) which is greater than chance. The challenge is – what does this mean? As a starting point for interpreting brain-to-brain synchrony, it is useful to consider a closely related finding – inter-subject correlations in fMRI data which arise even when two participants are not interacting but they perform the same task. In a landmark work, Hasson showed individual participants in an MRI scanner a classic movie and calculated the correlation between the brain activation patterns across participants (Hasson et al., 2004). Robust inter-subject correlations in these sequential brain imaging studies were found across many visual and semantic areas, reflecting participant's consistent responses to the rich auditory-visual stimulus. Following the acronyms used by the Hasson lab, we refer to this signal as sequential inter-subject correlation (s-ISC). More recent studies using these measures of s-ISC show that it is stronger for more engaging movies (Hasson and Landesman, 2008) and when participants have a similar interpretation of the movie (Nguyen et al., 2019). s-ISC can be seen when the same story is told in different languages (Yeshurun et al., 2017) and between a speaker and a listener (Stephens et al., 2010). It is increasingly clear that s-ISC results reflect the combination of the external auditory-visual environment and the participant's conceptual understanding of the story. That it, common visual input, common auditory input and common conceptual understanding can drive s-ISC, even in the absence of any interaction between participants.

This 'common input problem' (Burgess, 2013; Holroyd, 2022) is critical to interpreting hyperscanning studies because, in any fNIRS hyperscanning task, the two participants in the same room at the same time receive many common inputs: they hear the same words or see

the same images on the screen or even feel the same rumbling as a large truck drives past the lab. This means it is very hard to know if the INS recorded in a hyperscanning study is just the result of common sensory inputs and is identical to s-ISC, or if it is different and can give us something more. The simple interpretation of INS as a measure of 'brain convergence' may be useful in applied or clinical contexts. For example, INS might predict learning (Osaka et al., 2015) or the success of therapy (Zhang et al., 2018) without any strict cognitive interpretation of where the INS is coming from. Similarly, the level of INS in group decision making seems to predict with-group coordination and between-group hostility (Yang et al., 2020). However, with this interpretation it is not always obvious how hyperscanning provides more than the fMRI movie-watching measures which are widely established.

To obtain a more detailed cognitive interpretation of the brain mechanisms involved in social interactions, it is essential to enrich brain-only approaches with additional measures of behavior and physiology (embodied hyperscanning; Figure 10 B). This is because, in a face-to-face interaction, any coordination between brains must be mediated by the actions of the body – the patterns of gaze and smiles and gestures and speech which instantiate the interaction. If these cannot be tracked and analysed, it is very hard to obtain a cognitive interpretation of the neural effects. The need for this behavioural data provides another example of the benefit of multimodal neuroimaging setups (Hamilton, 2021; Guglielmini et al., 2022).

Pioneering data from mice (Kingsbury et al., 2019) and bats (Zhang and Yartsev, 2019) suggests that brain-to-brain synchrony can be interpreted within the 'mutual prediction model', according to which coherent brain signals arise from pairs of individuals both predicting the behaviour of their partner at the same time. That is, if A is performing the action of picking up a mug, and B represents the actions of A, then the brains of both A and B are encoding the action 'lift mug' and will show coherence across brains. If A hands the mug over to B, then both brains will represent the joint action of 'A-give-mug; B-take-mug', and again there will be similar activation patterns. However, if B turns away or does an unpredictable action, the coherence pattern might be lost. Building on this, Hamilton and others have proposed the use of cross-brain general linear models (xGLMs) to understand the brain-to-brain signals more in detail. In this approach, the brain activation of a single participant (A) can be modelled as a combination of regressors describing the behaviour of A, the behaviour of partner B and the brain activity of B. Such approach was used on pairs of

mice (REF Kingsbury) and demonstrated that, when there was a clear dominant-subordinate relationship, brain activity of partner B was a stronger predictor of brain activity of A (i.e., stronger INS), and that the behaviour of the dominant rat was the strongest predictor of interbrain synchrony. Recent studies in our own lab draw on a similar approach to understanding human social interaction. For example, in Pinti et al. (2021) pairs of participants played a card game together where each round gave one participant the chance to lie (or tell the truth) and the other had to guess if their partner lied or not. A cross-brain model was built which used the signal from PFC of the task leader to predict the brain activity of the follower, and it identified channels in medial PFC with a 2 second delay between the two brains. This implies that the follower's brain is echoing the leaders, in line with the information structure of the task.

The most important outcome from this pioneering work is the recognition that brain-to-brain synchrony is mediated by behaviour. Interpersonal neural synchrony is not a telepathic connection between one brain and another, but rather one person's brain activity causes a social behaviour – a smile or gaze or word – which is then perceived by the other person and causes a change in their brain. The physical body provides the medium through which brain-to-brain coupling arises, and thus our study of the social brain must be closely linked to our study of embodied cognition. Therefore, when we study hyperscanning, it is not enough just to record brain signals; it is critical to also record the body. This includes physiological signals (e.g., heart beat, breathing) which can impact on the fNIRS signal (Tachtsidis and Scholkmann, 2016), and behaviour such as speech, gaze, facial movement and hand movement. An increasing number of motion capture technologies mean make this feasible, including wearable eyetrackers (Kredel et al., 2017) analysis of facial motion with computer vision (Baltrusaitis et al., 2016) and sensitive movement tracking (Hale et al., 2020; Pishchulin et al., 2016). Such systems could be standard in a hyperscanning lab and have the potential to strengthen our understanding of the roles brain and body play together in social interaction, describing how the meeting of minds is mediated by an interaction of bodies.

Conclusion and future outlooks

This chapter provides an overview of the working principles of fNIRS and has outlined some of the opportunities and limitations of using fNIRS for cognitive neuroscience investigations. We have concentrated in particular upon experimental contexts which aim to have a high degree of ecological validity, since this appears to be a major future opportunity for fNIRS experimentation in particular. The numerous hardware and software advances that we witnessed over the past 30 years made it possible to establish this neuroimaging modality as a valuable tool to understand human brain functioning. fNIRS is rapidly expanding in the field of cognitive neuroscience, allowing researchers to image functional brain activity across all ages and populations, both healthy and unhealthy. With the most recent hardware developments, the research community has now the opportunity to understand how brain works in more realistic contexts with mobile and wireless fNIRS devices. This means that brain activity can be measured in an even wider range of applications and environments, such in the streets or in the theatre and during naturalistic face-to-face interactions, where people are free to move around and to behave more similarly to how they would normally do in their everyday lives. This massively increases the ecological validity of cognitive tasks that now can become a better model of real-world behavior and may improve the reliability and robustness of the fNIRS-measured haemodynamic responses.

However, there remain several challenges, and further advances are needed. As we move increasingly towards more naturalistic experiments with designs that allow greater scope for the participant to decide for themselves what they do and when, and how they move physically, methodological and analytic improvements are required to be able to deal with single-trial analyses and new noise components in our signals, and hardware, to improve spatial sensitivity, cap miniaturization and stability, and overall signals quality. Given the undoubted importance of having multimodal setups that combine fNIRS with other monitors of brain activity (e.g., EEG), behavior (e.g., motion tracking, eye-tracking), physiology (e.g., ECG, respiration, blood pressure) as discussed in this chapter, we hope to see in the near future integrated platforms available to the market, with synchronized multimodal data streams, that ideally are fully wearable, lightweight, and suitable for all age ranges and for measuring from freely moving participants. This will have to be accompanied by the

development of multimodal data fusion algorithms that can process data coming from different channels and automatically feed this information into the assessment of functional brain activity recorded in unconstrained and unstructured settings. Finally, the last 30 years have been characterized by undoubted progress in hardware developments, and we expect there to be many more in the future, with even more sophisticated instruments, including but not limited to high-coverage and high-density, broadband NIRS, diffuse correlation spectroscopy, etc. These will further benefit the optics and neuroimaging community, but may also raise concerns about regulations, reliability and certification process of the equipment. Every fNIRS instrument should be tested in accordance to international norms to ensure its safety and performance before being commercialized or used in a research setting. For example, the standard IEC 80601-2-71 has been published to regulate the requirements for NIRS equipment. Other protocols such as the Noninvasive Imaging of Brain Function and Disease by Pulsed Near Infrared Light (nEUROPt) define phantom tests to characterize the performance of fNIRS instrumentation (Wabnitz et al., 2014). Such protocols are crucial to increase the comparability and reliability of fNIRS studies and should be incorporated as standard practice.

Disclosures

PP served as a paid consultant for Gowerlabs Ltd and has no competing or potential conflicts of interest.

Acknowledgments

PP was supported by the Wellcome Trust (212979/Z/18/Z).

References

Aasted, C. M., Yücel, M. A., Cooper, R. J., Dubb, J., Tsuzuki, D., Becerra, L., ... & Boas, D. A. (2015). Anatomical guidance for functional near-infrared spectroscopy: AtlasViewer tutorial. Neurophotonics, 2(2), 020801-020801.

Balardin, J. B., Zimeo Morais, G. A., Furucho, R. A., Trambaiolli, L., Vanzella, P., Biazoli Jr, C., & Sato, J. R. (2017). Imaging brain function with functional near-infrared spectroscopy in unconstrained environments. Frontiers in human neuroscience, 11, 258.

Bale, G., Elwell, C. E., & Tachtsidis, I. (2016). From Jöbsis to the present day: a review of clinical near-infrared spectroscopy measurements of cerebral cytochrome-c-oxidase. Journal of biomedical optics, 21(9), 091307-091307.

Baltrusaitis, T., Robinson, P., & Morency, L.-P. (2016). OpenFace: An open source facial behavior analysis toolkit. 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), 1–10.

Baker, J. M., Liu, N., Cui, X., Vrticka, P., Saggar, M., Hosseini, S. M. H., & Reiss, A. L. (2016). Sex differences in neural and behavioral signatures of cooperation revealed by fNIRS hyperscanning. Scientific Reports, 6(1), 26492.

Brigadoi, S., & Cooper, R. J. (2015). How short is short? Optimum source–detector distance for short-separation channels in functional near-infrared spectroscopy. Neurophotonics, 2(2), 025005-025005.

Buckley, E. M., Parthasarathy, A. B., Grant, P. E., Yodh, A. G., & Franceschini, M. A. (2014). Diffuse correlation spectroscopy for measurement of cerebral blood flow: future prospects. *Neurophotonics*, *1*(1), 011009-011009.

Burgess, A. P. (2013). On the interpretation of synchronization in EEG hyperscanning studies: a cautionary note. Frontiers in Human Neuroscience, 7, 881.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Laure, M. C., Dawson, D. R., ... & Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. Journal of the international neuropsychological society, 12(2), 194-209.

Burgess, P. W. & Wu, H-C. (2013) Rostral prefrontal cortex (Brodmann area 10): metacognition in the brain. (2013) Chapter 31 in: Principles of frontal lobe function, 2nd edition (Editors: Donald T. Stuss & Robert T. Knight) pp. 524-534.

Burgess, P. W., Crum, J., Pinti, P., Aichelburg, C., Oliver, D., Lind, F., ... & Hamilton, A. (2022). Prefrontal cortical activation associated with prospective memory while walking around a real-world street environment. NeuroImage, 258, 119392.

Chiarelli, A. M., Maclin, E. L., Low, K. A., Fabiani, M., & Gratton, G. (2015). Comparison of procedures for co-registering scalp-recording locations to anatomical magnetic resonance images. Journal of Biomedical Optics, 20(1), 016009-016009.

Collins-Jones, L. H., Cooper, R. J., Bulgarelli, C., Blasi, A., Katus, L., McCann, S., ... & BRIGHT Study Team. (2021). Longitudinal infant fNIRS channel-space analyses are robust to variability

parameters at the group-level: An image reconstruction investigation. NeuroImage, 237, 118068.

Cui, X., Bryant, D. M., & Reiss, A. L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. NeuroImage, 59(3), 2430–2437.

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? Trends in Cognitive Sciences, 14(10), 441–447.

de Rond, V., Gilat, M., D'Cruz, N., Hulzinga, F., Orban de Xivry, J. J., & Nieuwboer, A. (2023). Test-retest reliability of functional near-infrared spectroscopy during a finger-tapping and postural task in healthy older adults. Neurophotonics, 10(2), 025010-025010.

Dopierala, A. A., & Emberson, L. L. (2023). Towards imaging the infant brain at play. Communicative & Integrative Biology, 16(1), 2206204.

Duncan, A., Meek, J. H., Clemence, M., Elwell, C. E., Fallon, P., Tyszczuk, L., ... & Delpy, D. T. (1996). Measurement of cranial optical path length as a function of age using phase resolved near infrared spectroscopy. *Pediatric research*, *39*(5), 889-894.

Durduran, T., & Yodh, A. G. (2014). Diffuse correlation spectroscopy for non-invasive, micro-vascular cerebral blood flow measurement. *Neuroimage*, *85*, 51-63.

Emberson, 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(3), 031406-031406.

Eggebrecht, A. T., Ferradal, S. L., Robichaux-Viehoever, A., Hassanpour, M. S., Dehghani, H., Snyder, A. Z., ... & Culver, J. P. (2014). Mapping distributed brain function and networks with diffuse optical tomography. Nature photonics, 8(6), 448-454.

Erdoğan, S. B., Yücel, M. A., & Akın, A. (2014). Analysis of task-evoked systemic interference in fNIRS measurements: insights from fMRI. Neuroimage, 87, 490-504.

Erel, Y., Jaffe-Dax, S., Yeshurun, Y., & Bermano, A. H. (2020). STORM-Net: Simple and Timely Optode Registration Method for Functional Near-Infrared Spectroscopy (fNIRS). bioRxiv, 2020-12.

Fantini, S., & Sassaroli, A. (2020). Frequency-domain techniques for cerebral and functional near-infrared spectroscopy. Frontiers in neuroscience, 14, 300.

Ferrari, M., & Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. Neuroimage, 63(2), 921-935.

Fishburn, F. A., Murty, V. P., Hlutkowsky, C. O., MacGillivray, C. E., Bemis, L. M., Murphy, M. E., Huppert, T. J., & Perlman, S. B. (2018). Putting our heads together: interpersonal neural synchronization as a biological mechanism for shared intentionality. Social Cognitive and Affective Neuroscience, 13(8), 841–849.

Frijia, E. M., Billing, A., Lloyd-Fox, S., Rosas, E. V., Collins-Jones, L., Crespo-Llado, M. M., ... & Cooper, R. J. (2021). Functional imaging of the developing brain with wearable high-density diffuse optical tomography: a new benchmark for infant neuroimaging outside the scanner environment. NeuroImage, 225, 117490.

Fu, X., & Richards, J. E. (2021). Investigating developmental changes in scalp-to-cortex correspondence using diffuse optical tomography sensitivity in infancy. Neurophotonics, 8(3), 035003-035003.

Fu, X., & Richards, J. E. (2022). Age-related changes in diffuse optical tomography sensitivity profiles from childhood to adulthood. Journal of Biomedical Optics, 27(8), 083004-083004.

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C. D., & Tylén, K. (2012). Coming to Terms: Quantifying the Benefits of Linguistic Coordination. Psychological Science, 23(July), 931–939.

Gagnon, L., Cooper, R. J., Yücel, M. A., Perdue, K. L., Greve, D. N., & Boas, D. A. (2012). Short separation channel location impacts the performance of short channel regression in NIRS. Neuroimage, 59(3), 2518-2528.

Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. Neuroimage, 16(3 Pt 1), 814-21.

Gemignani, J., & Gervain, J. (2021). Comparing different pre-processing routines for infant fNIRS data. Developmental cognitive neuroscience, 48, 100943.

Greaves, D. A., Pinti, P., Din, S., Hickson, R., Diao, M., Lange, C., ... & Hamilton, A. F. D. C. (2022). Exploring Theater Neuroscience: Using Wearable Functional Near-infrared Spectroscopy to Measure the Sense of Self and Interpersonal Coordination in Professional Actors. Journal of Cognitive Neuroscience, 34(12), 2215-2236.

Guglielmini, S., Bopp, G., Marcar, V. L., Scholkmann, F., & Wolf, M. (2022). Systemic physiology augmented functional near-infrared spectroscopy hyperscanning: a first evaluation investigating entrainment of spontaneous activity of brain and body physiology between subjects. Neurophotonics, 9(2), 026601.

Gunadi, S., Leung, T. S., Elwell, C. E., & Tachtsidis, I. (2014). Spatial sensitivity and penetration depth of three cerebral oxygenation monitors. *Biomedical optics express*, *5*(9), 2896-2912.

Gvirts, H. Z., & Perlmutter, R. (2019). What Guides Us to Neurally and Behaviorally Align With Anyone Specific? A Neurobiological Model Based on fNIRS Hyperscanning Studies. The

Neuroscientist, 107385841986191.

Haeussinger, F. B., Heinzel, S., Hahn, T., Schecklmann, M., Ehlis, A. C., & Fallgatter, A. J. (2011). Simulation of near-infrared light absorption considering individual head and prefrontal cortex anatomy: implications for optical neuroimaging. PloS one, 6(10), e26377.

Hamilton, A. F. D. C. (2021). Hyperscanning: beyond the hype. Neuron, 109(3), 404-407.

Hamilton, A. F. de C., & Lind, F. (2016). Audience effects: what can they tell us about social neuroscience, theory of mind and autism? Culture and Brain, 4(2), 159–177.

Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of Social Interaction in Human Brain Function. Neuron, 88(1), 181–193.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. Science, 303(5664), 1634–1640.

Hasson, U., & Landesman, O. (2008). Neurocinematics: The neuroscience of film. …, 2(1), 1–26.

Hasson, U., & Frith, C. D. (2016). Mirroring and beyond: Coupled dynamics as a generalized framework for modelling social interactions. In Philosophical Transactions of the Royal Society B: Biological Sciences (Vol. 371, Issue 1693).

Henson, R. (2015). https://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency

Hernandez, S. M., & Pollonini, L. (2020, April). NIRSplot: a tool for quality assessment of fNIRS scans. In *Optics and the Brain* (pp. BM2C-5). Optica Publishing Group.

Hirsch, J., Zhang, X., Noah, J. A., & Ono, Y. (2017). Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. NeuroImage, 157, 314-330.

Hirsch, J., Adam Noah, J., Zhang, X., Dravida, S., & Ono, Y. (2018). A cross-brain neural mechanism for human-to-human verbal communication. Social Cognitive and Affective Neuroscience, 13(9), 907–920.

Holroyd, C. B. (2022). Interbrain synchrony: on wavy ground. *Trends in Neurosciences*, *45*(5), 346-357.

Hu, X. S., Wagley, N., Rioboo, A. T., DaSilva, A. F., & Kovelman, I. (2020). Photogrammetry-based stereoscopic optode registration method for functional near-infrared spectroscopy. Journal of Biomedical Optics, 25(9), 095001-095001.

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. Applied optics, 48(10), D280-D298.

Jaffe-Dax, S., Bermano, A. H., Erel, Y., & Emberson, L. L. (2020). Video-based motion-resilient reconstruction of three-dimensional position for functional near-infrared spectroscopy and electroencephalography head mounted probes. Neurophotonics, 7(3), 035001-035001.

Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., & Lu, C. (2012). Neural synchronization during face-to-face communication. Journal of Neuroscience, 32(45), 16064–16069.

Julien, C. (2006). The enigma of Mayer waves: facts and models. Cardiovascular research, 70(1), 12-21.

Jurcak, V., Tsuzuki, D., & Dan, I. (2007). 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. Neuroimage, 34(4), 1600-1611.

Kim, S. G., & Ogawa, S. (2012). Biophysical and physiological origins of blood oxygenation level-dependent fMRI signals. Journal of Cerebral Blood Flow & Metabolism, 32(7), 1188-1206.

Kingsbury, L., Huang, S., Wang, J., Gu, K., Golshani, P., Wu, Y. E., & Hong, W. (2019). Correlated neural activity and encoding of behavior across brains of socially interacting animals. Cell, 178(2), 429-446.

Kirilina, E., Jelzow, A., Heine, A., Niessing, M., Wabnitz, H., Brühl, R., ... & Tachtsidis, I. (2012). The physiological origin of task-evoked systemic artefacts in functional near infrared spectroscopy. Neuroimage, 61(1), 70-81.

Kirilina, E., Yu, N., Jelzow, A., Wabnitz, H., Jacobs, A. M., & Tachtsidis, I. (2013). Identifying and quantifying main components of physiological noise in functional near infrared spectroscopy on the prefrontal cortex. Frontiers in human neuroscience, 7, 864.

Kovacsova, Z., Bale, G., Mitra, S., Lange, F., & Tachtsidis, I. (2021). Absolute quantification of cerebral tissue oxygen saturation with multidistance broadband NIRS in newborn brain. Biomedical Optics Express, 12(2), 907-925.M

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. Quarterly Journal of Experimental Psychology (2006), 63(11), 2220–2230.

Kredel, R., Vater, C., Klostermann, A., & Hossner, E.-J. (2017). Eye-Tracking Technology and the Dynamics of Natural Gaze Behavior in Sports: A Systematic Review of 40 Years of Research. Frontiers in Psychology, 8(OCT), 1845.

Krishnan-Barman, S. (2021) Adults imitate to send a social signal. Doctoral thesis (Ph.D), UCL (University College London).

Kwasa, J., Peterson, H. M., Jones, L., Karrobi, K., Parker, T., Nickerson, N., & Wood, S. (2022). Demographic Reporting and Phenotypic Exclusion in fNIRS. bioRxiv, 2022-11.

Jasińska, K. K., Nematova, S., & Yang, X. (2023). Neural Activation during Phonological Processing in Primary-School Children with Limited Reading Experience: Insights from Rural Côte d'Ivoire.

Iadecola, C. (2017). The neurovascular unit coming of age: a journey through neurovascular coupling in health and disease. Neuron, 96(1), 17-42.

Izuma, K., Saito, D. N., & Sadato, N. (2010). The roles of the medial prefrontal cortex and striatum in reputation processing. Social Neuroscience, 5(2), 133–147.

Lakens, D., & Evers, E. R. (2014). Sailing from the seas of chaos into the corridor of stability: Practical recommendations to increase the informational value of studies. Perspectives on psychological science, 9(3), 278-292.

Lloyd-Fox, S., Papademetriou, M., Darboe, M. K., Everdell, N. L., Wegmuller, R., Prentice, A. M., ... & Elwell, C. E. (2014). Functional near infrared spectroscopy (fNIRS) to assess cognitive function in infants in rural Africa. Scientific reports, 4(1), 1-8.

Lloyd-Fox, S., Richards, J. E., Blasi, A., Murphy, D. G., Elwell, C. E., & Johnson, M. H. (2014). Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants. Neurophotonics, 1(2), 025006-025006.

Lloyd-Fox, S., Széplaki-Köllőd, B., Yin, J., & Csibra, G. (2015). Are you talking to me? Neural activations in 6-month-old infants in response to being addressed during natural interactions. Cortex, 70, 35-48.

Lloyd-Fox, S., Blasi, A., McCann, S., Rozhko, M., Katus, L., Mason, L., ... & Bright Project Team. (2019). Habituation and novelty detection fNIRS brain responses in 5-and 8-month-old infants: The Gambia and UK. Developmental Science, 22(5), e12817.

Mark, C. I., Mazerolle, E. L., & Chen, J. J. (2015). Metabolic and vascular origins of the BOLD effect: Implications for imaging pathology and resting-state brain function. Journal of Magnetic Resonance Imaging, 42(2), 231-246.

Mazzonetto, I., Castellaro, M., Cooper, R. J., & Brigadoi, S. (2022). Smartphone-based photogrammetry provides improved localization and registration of scalp-mounted neuroimaging sensors. Scientific Reports, 12(1), 10862.

McKendrick, R., Parasuraman, R., Murtza, R., Formwalt, A., Baccus, W., Paczynski, M., & Ayaz, H. (2016). Into the wild: neuroergonomic differentiation of hand-held and augmented reality wearable displays during outdoor navigation with functional near infrared spectroscopy. Frontiers in human neuroscience, 216.

Müller-Pinzler, L., Gazzola, V., Keysers, C., Sommer, J., Jansen, A., Frässle, S., Einhäuser, W., Paulus, F. M., & Krach, S. (2015). Neural pathways of embarrassment and their modulation by social anxiety. NeuroImage, 119, 252–261.

Myllyneva, A., & Hietanen, J. K. (2015). The dual nature of eye contact: to see and to be seen. Social Cognitive and Affective Neuroscience, nsv075-.

Nguyen, M., Vanderwal, T., & Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. NeuroImage, 184, 161–170.

Noah, J. A., Zhang, X., Dravida, S., DiCocco, C., Suzuki, T., Aslin, R. N., ... & Hirsch, J. (2021). Comparison of short-channel separation and spatial domain filtering for removal of non-neural components in functional near-infrared spectroscopy signals. Neurophotonics, 8(1), 015004-015004.

Osaka, N., Minamoto, T., Yaoi, K., Azuma, M., Shimada, Y. M., & Osaka, M. (2015). How Two Brains Make One Synchronized Mind in the Inferior Frontal Cortex: fNIRS-Based Hyperscanning During Cooperative Singing. Frontiers in Psychology, 6, 1811.

Okamoto, M., & Dan, I. (2005). Automated cortical projection of head-surface locations for transcranial functional brain mapping. Neuroimage, 26(1), 18-28.

Pan, Y., Cheng, X., Zhang, Z., Li, X., & Hu, Y. (2017). Cooperation in lovers: An fNIRS-based hyperscanning study. Human Brain Mapping, 38(2), 831–841.

Paranawithana, I., Mao, D., Wong, Y. T., & McKay, C. M. (2022). Reducing false discoveries in resting-state functional connectivity using short channel correction: an fNIRS study. Neurophotonics, 9(1), 015001-015001.

Pinti, P., Aichelburg, C., Lind, F., Power, S., Swingler, E., Merla, A., ... & Tachtsidis, I. (2015). Using fiberless, wearable fNIRS to monitor brain activity in real-world cognitive tasks. JoVE (Journal of Visualized Experiments), (106), e53336.

Pinti, P., Merla, A., Aichelburg, C., Lind, F., Power, S., Swingler, E., ... & Tachtsidis, I. (2017). A novel GLM-based method for the Automatic IDentification of functional Events (AIDE) in fNIRS data recorded in naturalistic environments. Neuroimage, 155, 291-304.

Pinti, P., Aichelburg, C., Gilbert, S., Hamilton, A., Hirsch, J., Burgess, P., & Tachtsidis, I. (2018). A review on the use of wearable functional near-infrared spectroscopy in naturalistic environments. Japanese Psychological Research, 60(4), 347-373.

Pinti, P., Scholkmann, F., Hamilton, A., Burgess, P., & Tachtsidis, I. (2019). Current status and issues regarding pre-processing of fNIRS neuroimaging data: an investigation of diverse signal filtering methods within a general linear model framework. Frontiers in human neuroscience, 12, 505.

Pinti, P., Devoto, A., Greenhalgh, I., Tachtsidis, I., Burgess, P. W., & de C Hamilton, A. F. (2021). The role of anterior prefrontal cortex (area 10) in face-to-face deception measured with fNIRS. Social Cognitive and Affective Neuroscience, 16(1-2), 129-142.

Pishchulin, L., Insafutdinov, E., Tang, S., Andres, B., Andriluka, M., Gehler, P., & Schiele, B. (2016). DeepCut: Joint Subset Partition and Labeling for Multi Person Pose Estimation.

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. Nature Reviews Neuroscience, 20(8), 495–505.

Reindl, V., Gerloff, C., Scharke, W., & Konrad, K. (2018). Brain-to-brain synchrony in parent-child dyads and the relationship with emotion regulation revealed by fNIRS-based hyperscanning. NeuroImage, 178, 493–502.

Richards, J. E., Sanchez, C., Phillips-Meek, M., & Xie, W. (2016). A database of age-appropriate average MRI templates. Neuroimage, 124, 1254-1259.

Sappia, M. S., Hakimi, N., Colier, W. N., & Horschig, J. M. (2020). Signal quality index: an algorithm for quantitative assessment of functional near infrared spectroscopy signal quality. *Biomedical Optics Express*, *11*(11), 6732-6754.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience1. Behavioral and brain sciences, 36(4), 393-414.

Scholkmann, F., & Wolf, M. (2013). General equation for the differential pathlength factor of the frontal human head depending on wavelength and age. Journal of biomedical optics, 18(10), 105004-105004.

Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Pavia, J. M., Wolf, U., & Wolf, M. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. Neuroimage, 85, 6-27.

Scholkmann, F., Tachtsidis, I., Wolf, M., & Wolf, U. (2022). Systemic physiology augmented functional near-infrared spectroscopy: a powerful approach to study the embodied human brain. Neurophotonics, 9(3), 030801-030801.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. Trends Cogn Sci, 10(2), 70–76.

Singh, A. K., Okamoto, M., Dan, H., Jurcak, V., & Dan, I. (2005). Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. Neuroimage, 27(4), 842-851.

Steptoe, A., Evans, O., & Fieldman, G. (1997). Perceptions of control over work: psychophysiological responses to self-paced and externally-paced tasks in an adult population sample. *International Journal of Psychophysiology*, *25*(3), 211-220.

Suzuki, S., Takasaki, S., Ozaki, T., & Kobayashi, Y. (1999, July). Tissue oxygenation monitor using NIR spatially resolved spectroscopy. In *Optical tomography and spectroscopy of tissue III* (Vol. 3597, pp. 582-592). SPIE.

Tachtsidis, I., & Scholkmann, F. (2016). False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. Neurophotonics, 3(3), 031405-031405.

Tak, S., Uga, M., Flandin, G., Dan, I., & Penny, W. D. (2016). Sensor space group analysis for fNIRS data. Journal of neuroscience methods, 264, 103-112.

Tsuzuki, D., & Dan, I. (2014). Spatial registration for functional near-infrared spectroscopy: from channel position on the scalp to cortical location in individual and group analyses. Neuroimage, 85, 92-103.

Tsuzuki, D., Jurcak, V., Singh, A. K., Okamoto, M., Watanabe, E., & Dan, I. (2007). Virtual spatial registration of stand-alone fNIRS data to MNI space. Neuroimage, 34(4), 1506-1518.

Yang, J., Zhang, H., Ni, J., De Dreu, C. K. W., & Ma, Y. (2020). Within-group synchronization in the prefrontal cortex associates with intergroup conflict. Nature Neuroscience, 1–7.

Ye, J. C., Tak, S., Jang, K. E., Jung, J., & Jang, J. (2009). NIRS-SPM: statistical parametric mapping for near-infrared spectroscopy. Neuroimage, 44(2), 428-447.

Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C. J., & Hasson, U. (2017). Same Story, Different Story. Psychological Science, 095679761668202.

Yücel, M. A., Selb, J., Boas, D. A., Cash, S. S., & Cooper, R. J. (2014). Reducing motion artifacts for long-term clinical NIRS monitoring using collodion-fixed prism-based optical fibers. Neuroimage, 85, 192-201.

Yücel, M. A., Selb, J., Aasted, C. M., Lin, P. Y., Borsook, D., Becerra, L., & Boas, D. A. (2016). Mayer waves reduce the accuracy of estimated hemodynamic response functions in functional near-infrared spectroscopy. Biomedical optics express, 7(8), 3078-3088.

Yücel, M. A., Lühmann, A. V., Scholkmann, F., Gervain, J., Dan, I., Ayaz, H., ... & Wolf, M. (2021). Best practices for fNIRS publications. Neurophotonics, 8(1), 012101-012101.

Wabnitz, H., Jelzow, A., Mazurenka, M., Steinkellner, O., Macdonald, R., Milej, D., ... & Pifferi, A. (2014). Performance assessment of time-domain optical brain imagers, part 2: nEUROPt protocol. *Journal of biomedical optics*, *19*(8), 086012-086012.

Wabnitz, H., Contini, D., Spinelli, L., Torricelli, A., & Liebert, A. (2020). Depth-selective data analysis for time-domain fNIRS: moments vs. time windows. Biomedical Optics Express, 11(8), 4224-4243.

White, B. R., & Culver, J. P. (2010). Quantitative evaluation of high-density diffuse optical tomography: in vivo resolution and mapping performance. Journal of biomedical optics, 15(2), 026006-026006.

Wijeakumar, S., Huppert, T. J., Magnotta, V. A., Buss, A. T., & Spencer, J. P. (2017). Validating an image-based fNIRS approach with fMRI and a working memory task. NeuroImage, 147, 204-218.

Wyser, D., Mattille, M., Wolf, M., Lambercy, O., Scholkmann, F., & Gassert, R. (2020). Short-channel regression in functional near-infrared spectroscopy is more effective when considering heterogeneous scalp hemodynamics. Neurophotonics, 7(3), 035011-035011.

Uchitel, J., Blanco, B., Collins-Jones, L., Edwards, A., Porter, E., Pammenter, K., ... & Austin, T. (2023). Cot-side imaging of functional connectivity in the developing brain during sleep using wearable high-density diffuse optical tomography. NeuroImage, 265, 119784.

Vidal-Rosas, E. E., Zhao, H., Nixon-Hill, R. W., Smith, G., Dunne, L., Powell, S., ... & Everdell, N. L. (2021). Evaluating a new generation of wearable high-density diffuse optical tomography technology via retinotopic mapping of the adult visual cortex. Neurophotonics, 8(2), 025002-025002.

Vidal-Rosas, E. E., von Lühmann, A., Pinti, P., & Cooper, R. J. (2023). Wearable, high-density fNIRS and diffuse optical tomography technologies: a perspective. Neurophotonics, 10(2), 023513-023513.

Virtanen, J., Noponen, T., & Meriläinen, P. (2009). Comparison of principal and independent component analysis in removing extracerebral interference from near-infrared spectroscopy signals. Journal of biomedical optics, 14(5), 054032-054032.

von Lühmann, A., Ortega-Martinez, A., Boas, D. A., & Yücel, M. A. (2020). Using the general linear model to improve performance in fNIRS single trial analysis and classification: a perspective. Frontiers in human neuroscience, 14, 30.

von Lühmann, A., Li, X., Müller, K. R., Boas, D. A., & Yücel, M. A. (2020). Improved physiological noise regression in fNIRS: a multimodal extension of the general linear model using temporally embedded canonical correlation analysis. NeuroImage, 208, 116472.

von Lühmann, A., Zheng, Y., Ortega-Martinez, A., Kiran, S., Somers, D. C., Cronin-Golomb, A., ... & Yücel, M. A. (2021). Toward Neuroscience of the Everyday World (NEW) using functional near-infrared spectroscopy. Current opinion in biomedical engineering, 18, 100272.

Zhang, X., Noah, J. A., & Hirsch, J. (2016). Separation of the global and local components in functional near-infrared spectroscopy signals using principal component spatial filtering. Neurophotonics, 3(1), 015004-015004.

Zhang, Y., Meng, T., Hou, Y., Pan, Y., & Hu, Y. (2018). Interpersonal brain synchronization associated with working alliance during psychological counseling. Psychiatry Research: Neuroimaging, 282, 103-109.

Zhang, W., & Yartsev, M. M. (2019). Correlated neural activity across the brains of socially interacting bats. Cell, 178(2), 413-428.

Zimeo Morais, G. A., Scholkmann, F., Balardin, J. B., Furucho, R. A., de Paula, R. C. V., Biazoli Jr, C. E., & Sato, J. R. (2018). Non-neuronal evoked and spontaneous hemodynamic changes in the anterior temporal region of the human head may lead to misinterpretations of functional near-infrared spectroscopy signals. Neurophotonics, 5(1), 011002-011002.

Zohdi, H., Egli, R., Guthruf, D., Scholkmann, F., & Wolf, U. (2021). Color-dependent changes in humans during a verbal fluency task under colored light exposure assessed by SPA-fNIRS. Scientific reports, 11(1), 9654.