Optimizing short-channel regression in fNIRS: an empirical evaluation in a naturalistic multimodal paradigm

Y. Lemaire[1], P. Barone[1], K. Strelnikov[1,2], O. Deguine[1,3], A. Caclin[4], J. Ginzburg[4]

1. Université de Toulouse, CNRS, Centre de Recherche Cerveau et Cognition CerCo UMR5549, Toulouse, France

2. Centre for Cognitive and Brain Sciences, University of Macau, Taipa, Macau Special Administrative Regions of China

3. Service d'Oto Rhino Laryngologie (ORL), Otoneurologie et ORL Pédiatrique, Hôpital Pierre-Paul-Riquet, CHU de Toulouse, site Purpan, Toulouse, France

4. Université Claude Bernard Lyon 1, INSERM, CNRS, Centre de Recherche en Neurosciences de Lyon CRNL U1028 UMR5292, Bron, France

Corresponding author: Yann Lemaire, yann.lemaire@cnrs.fr

Abstract:

Functional Near-Infrared Spectroscopy (fNIRS) is increasingly favored for its portability and suitability for ecological paradigms, yet methodological standardization remains a challenge, especially regarding the optimal use of short-separation channels (SC) to remove systemic physiological noise. While SC signals are widely recognized for isolating extracerebral hemodynamics such as cardiac and respiratory artifacts, no consensus exists on how best to incorporate them into analysis pipelines, limiting comparability across studies. Here, we systematically investigate six SC regression strategies within a Generalized Linear Model (GLM) framework, using fNIRS data acquired from temporal and occipital cortices of 16 healthy adults during naturalistic auditory, visual, and audiovisual passive stimulation. These strategies varied in the number of SCs used, anatomical specificity, and the application of dimensionality reduction. Our results show that for oxygenated hemoglobin signals, processing data without SC regression yields uninterpretable outcomes, with widespread negative beta estimates and no detectable task-related activation, underscoring the crucial role of SC correction. Among correction methods, approaches pooling all SC signals without anatomical constraints consistently outperformed spatially constrained techniques mimicking limited SC availability. Furthermore, applying orthogonalization to the full set of SCs provided additional improvements by efficiently capturing shared systemic variance and reducing redundancy, enhancing detection of modality-specific cortical activations and condition contrasts. Finally, although the deoxygenated hemoglobin signal is inherently less sensitive to systemic noise, it benefited from the same SC regression strategies, supporting their applicability across chromophores. Overall, our findings highlight the essential role of SC regression in recovering physiologically meaningful signals in fNIRS and recommend including all available SC channels within the GLM, coupled with orthogonalization techniques, as a robust and generalizable best practice for denoising across diverse experimental designs and hardware configurations.

Keywords:

Functional Near-Infrared Spectroscopy (fNIRS); Short-separation channels; Systemic physiological noise; Naturalistic stimuli; Analysis strategy; Modality-specific cortical responses

1. Introduction

Functional Near-Infrared Spectroscopy (fNIRS) is an increasingly prominent neuroimaging modality, valued for its portability, affordability, and non-invasiveness (Yücel et al., 2024). By measuring relative changes in oxy- (HbO) and deoxy- (HbR) hemoglobin concentrations in cortical tissue, fNIRS enables the investigation of a wide range of cognitive functions, including perception, attention, memory, language, and motor processing (Highton et al., 2023). Its ease of use and tolerance for motion makes it particularly well suited for studying populations traditionally underserved by neuroimaging research, such as infants, children, older adults, and clinical groups. Importantly, fNIRS is especially promising in the field of auditory neuroscience. Unlike fMRI, it is silent—an essential feature when studying auditory processing. Compared to electroencephalography (EEG) and magnetoencephalography (MEG), fNIRS is unaffected by hearing aids or cochlear implants, offering a unique advantage for studying individuals with hearing impairments. As such, fNIRS has been used in a growing body of research on auditory perception, including investigations into the processing of physical features of sound like intensity levels or auditory-spatial cues (Bauernfeind et al., 2018; Weder et al., 2018; van de Rijt et al., 2018; Chen et al., 2015), as well as research on auditory cognitive processes like speech processing, listening effort, language development, development of auditory cortical function, and auditory short-term/working memory (e.g., Bauernfeind et al., 2018; Gervain et al., 2008; Anderson et al., 2019; Ginzburg et al., 2024). It has also shown promise for clinical applications such as in the study of cortical reorganisation after impaired sensory input or in tinnitus and during subsequent rehabilitation with cochlear implants (Anderson et al., 2017; 2019).

Yet despite these many advantages, several key methodological challenges still limit the full potential of fNIRS. One major challenge lies in the diversity of preprocessing practices across research groups and toolboxes, which makes it difficult to compare findings and reach generalizable conclusions (Yücel et al., 2024). A central issue in fNIRS data analysis concerns the removal of systemic physiological noise, particularly through the use of short separation channels (SCs), which has yet to reach methodological consensus. The present study aims to address this issue by systematically evaluating different strategies to use SC information for systemic signal removal, in the context of passive, ecologically valid auditory, visual, and audiovisual stimulation, and across cortical regions.

1.1. Ecological stimulation: an underused opportunity

A growing body of neuroscience research calls for using naturalistic, ecological stimuli to better approximate real-world perception (Sonkusare et al., 2019). In this regard, fNIRS is uniquely positioned to bridge the gap between laboratory experiments and everyday cognition, given its robustness to motion and its flexibility across settings (Pinti et al., 2018). However, few fNIRS studies have taken full advantage of this by incorporating complex, real-life auditory or audiovisual stimuli such as movies or audio stories. Naturalistic paradigms allow for the study of perception under more realistic

conditions, capturing cognitive processing in a way that more controlled experiments cannot (Somech et al., 2022). Despite this potential, to our knowledge, no study has yet systematically examined the effectiveness of different analysis strategies when applied to fNIRS data acquired during passive, ecological stimulation involving auditory, visual, and audiovisual content.

1.2. Methodological variability and the need for standardization

As a relatively new neuroimaging tool, fNIRS is still undergoing rapid methodological evolution. In auditory research, its adoption has trailed behind more established techniques such as fMRI, EEG, and MEG. As a result, considerable variability exists in both experimental protocols and analytical pipelines across studies (Yücel et al., 2024). This lack of standardization complicates replication, interpretation, and cross-study comparisons. Auditory fNIRS studies most commonly use block or event-related designs. In block-design trials, a single stimulus is presented continuously for an extended duration (e.g., 10 seconds), followed by an inter-stimulus interval sufficient for the hemodynamic response to return to approximately baseline levels (Amaro et al., 2006). Despite being widely used, there is disagreement about the best analysis techniques for this experimental design; new algorithms and techniques are frequently released (Paranawhitana et al., 2022; Luke et al., 2021; Brigadoi, 2015; Santosa, 2020; Zhou, 2020) but there remains a significant need for thorough validation of these tools.

1.3. Challenges in signal quality and systemic noise

While fNIRS is praised for its practical advantages, it is not without limitations. One of the main challenges lies in the contamination of neural signals by systemic physiological noise including heartbeat, respiration, low-frequency blood pressure oscillations, Mayer waves, and spontaneous neural activity (Scholkmann et al., 2010; Kirilina et al., 2013; Tachtsidis & Scholkmann, 2016; Yücel et al., 2016). These confound can originate from both superficial extracerebral layers and deeper cortical tissues (Habermehl et al., 2012), and can significantly obscure the measurement of true cerebral responses. While certain components of systemic physiological noise (i,e., low-frequency blood pressure oscillations, heart rate, and respiratory signals) fall outside the frequency range of the hemodynamic response and can be effectively attenuated using temporal filtering (e.g., band-pass filtering, Pinti et al., 2020), this approach has clear limitations. In particular, it cannot eliminate low-frequency oscillations in blood flow, such as Mayer waves (~0.05 to 0.15 Hz), which overlap spectrally with the expected signal in block design paradigms (Huppert et al., 2009; Kirilina et al., 2013; Yücel et al., 2016). These slow vascular fluctuations, driven by systemic physiology rather than localized neural activity, remain a major confound for interpreting fNIRS data and need targeted methods for removal.

1.4. The unresolved issue of short-channels in fNIRS

SCs are a powerful tool for isolating superficial hemodynamic signals, particularly those arising from extracerebral systemic activity such as cardiac, respiratory, or Mayer waves. By positioning detectors at short distances from the source (typically ~8 mm), SCs primarily capture these superficial signals, which can then be regressed from the long-distance channel signals (that measure both the cerebral and superficial

signal) to improve the specificity of the cortical response. Before SCs became widely available from most fNIRS manufacturers, researchers relied on alternative methods to mitigate systemic contamination (e.g., independent component analysis, Kohno et al., 2007; HbO-HbR-correlation based correction, Cui et al., 2010; Yamada et al., 2012). While these methods provided partial solutions, the introduction of SCs marked a significant advance by offering a direct, empirical estimate of extracerebral physiology—an approach that has been shown to outperform other correction methods in terms of removing systemic artifacts and improving signal specificity (Luke et al., 2021; Klein et al., 2022; Wyser et al., 2020; Sato et al., 2016; Santosa et al., 2020; Zhou et al., 2020).

Despite growing recognition of their importance, the use and implementation of SCs in fNIRS studies remain inconsistent. In a large multi-lab collaborative study, Yücel et al. (2024) brought together 102 researchers from 38 international teams to analyze two fNIRS datasets. Their findings revealed a striking disparity in analysis pipelines and outcomes, underscoring the lack of consensus within the field and the urgent need for standardized preprocessing approaches. Specifically, 79% of the teams adopted a generalized linear model (GLM) framework, which enables the inclusion of SC signals as regressors. Among these, 46% used individual SC signals as single regressors, 6% used the first principal component of all SC signals, 39% applied a multiple-regressor approach using either raw SC signals or their principal components. For the remaining 9%, the methodology could not be determined from the available information. However, none of the teams reported the number or placement of SCs used, nor did they justify the rationale behind their chosen implementation.

These inconsistencies are also reflected in the main fNIRS signal processing toolboxes available to the community: MNE-NIRS (Luke et al., 2021), Homer2/Homer3 (Huppert et al., 2009), and NIRS brain AnalyzIR (Santosa et al., 2018). All three toolboxes offer methods for including SCs within a GLM framework, which is currently considered best practice, especially in block-design paradigms. However, each proposes a different implementation, and none has emerged as the standard approach. This mirrors the findings of Yücel et al. (2024), who pointed out the heterogeneity and methodological ambiguity around SC use in the literature.

When SCs are available, two main scenarios typically arise: (1) there is one SC for each source used in the montage, or (2) only a subset of sources are paired with SCs, due to hardware or design constraints. Within each of these scenarios, multiple strategies exist for regressing SCs from the long-channel signals. One approach favors spatial specificity, regressing each long channel with the SC signal nearest to its source-detector pair. The alternative, suggested by Santosa et al. (2020), is to use all available SCs to extract global components (e.g., via principal component analysis, PCA) and regress them from all long channels, thereby targeting systemic noise that is more global but potentially less spatially precise. This question becomes especially important in multi-lobar studies, where systemic interference might differ across anatomical regions—such as between temporal and occipital lobes—and decisions about spatial specificity versus generality in SC regression could critically impact the interpretation of results.

1.6. Objectives of the current study

The current study aims to systematically address the unresolved questions surrounding the use of SCs in fNIRS preprocessing, with a particular focus on their role

in block-design GLM analyses. We recorded fNIRS signals over the temporal and occipital lobes while administering passive, naturalistic auditory, visual, and audiovisual stimuli to healthy participants. Importantly, our montage was designed to include one SC per source, allowing us to explore a variety of SC correction strategies within the same dataset. We tested six distinct preprocessing pipelines:

1. No SCs (baseline condition)
2. One SC per lobe (n = 3), regressed using:
    1. spatial-specific SCs (nearest to the long channel)
    2. non-spatial-specific SCs (all available)
3. All SCs available (n = 8), regressed using:
    1. spatial-specific SCs (nearest to the long channel)
    2. non-spatial-specific with all available SC raw signals
    3. non-spatial-specific with all available SC components after PCA

This design enables us to evaluate not only the general effect of including SCs, but also the impact of spatial specificity, SC coverage, and signal summarization methods (e.g., PCA), both on HbO and HbR signal. Our findings aim to provide practical guidance and help move the field toward a more standardized and evidence-based use of SC regression in fNIRS analysis.

2. METHODS

2.1. Participants

Sixteen participants, aged 24 to 34 years (M = 27.1, SD = 2.7; 8 females), were recruited for this study. All participants self-reported having no auditory or uncorrected visual impairments, and none of the participants had any history of neurological or psychiatric conditions, nor had they been diagnosed with any neurodevelopmental disorders. The study was approved by the OUEST IV ethics committee (approval number: 2019-A00295-52) and was conducted in compliance with the Declaration of Helsinki. Informed consent was obtained from all participants before their involvement in the study.

2.2. Stimuli

Visual stimuli were presented on a 27” screen (DELL, P2719H) with a 60 Hz refresh rate. Auditory stimuli were delivered via speakers (KAIWIN SP 690N), calibrated to produce a sound level of approximately 60 dB SPL (A weighted). We exclusively used ecologically valid stimuli, comprising video clips or sounds of real-life objects and events (e.g., moving cars, a ringing phone). Fifteen distinct concepts (e.g., 'bird', 'sheep', 'boat') served as the basis for stimulus generation. For each concept, three corresponding stimuli were created: an auditory-only version (e.g., the sound of birds singing presented with a black screen), a visual-only version (e.g., a peaceful meadow), and an audiovisual version (e.g., a video of birds singing in a meadow with the corresponding sound). All stimuli were unique and sourced from royalty-free content websites. Each individual stimulus presentation lasted for 10 seconds.

2.3. fNIRS Montage and Data Acquisition

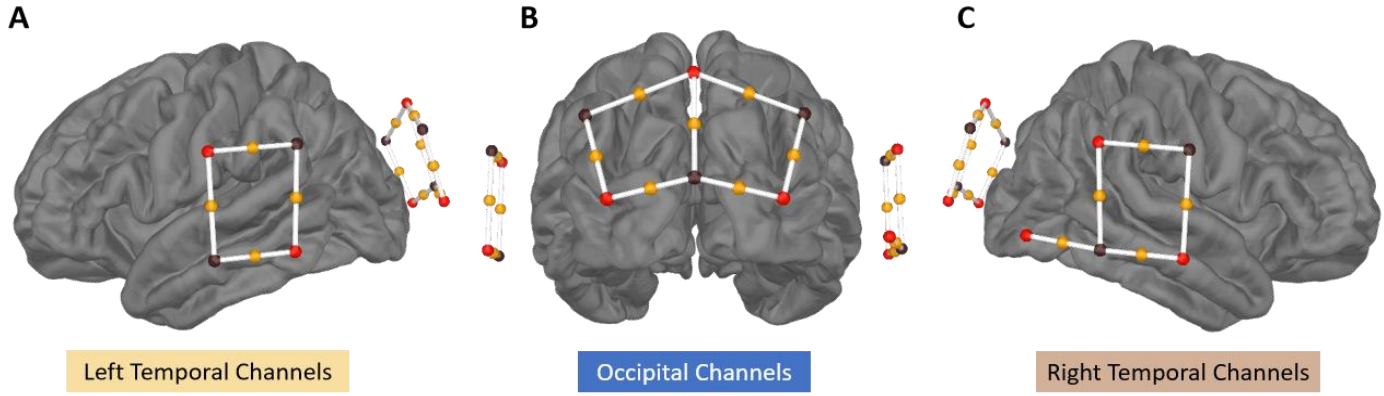


Figure 1: Location of sources (black spheres) and detectors (red spheres) and their corresponding midpoint channels (orange spheres with white lines connecting the source/detector pair for each channel) on (A) left; (B) front; and (C) right views of the brain. Channel 16, located in the left-most part of the right temporal montage, is the only channel without a corresponding counterpart on the left side (see also Table 1). It was excluded from further analysis due to this asymmetry. The visualization was obtained using the *mne-nirs* library (Luke et al., 2021).

Near-infrared light absorption was measured at 760 nm and 850 nm wavelengths at a sampling frequency of 7.81 Hz using a continuous wave NIRScout instrument (NIRx Medical Technologies, LLC). Data were collected using the NIRStar 15.2. acquisition software. Eight light sources and eight light detectors were attached to a cap with a 10-20 system to the probe location (see Figure 1). In addition, eight short 8-mm channels (one for each source) recorded systemic signals. The montage was created using fOLD (fNIRS Optodes Location Resolution, Zimeo Morais et al., 2018), which allows placing the optodes in the 10-20 international system to cover selected anatomical areas. We used the Brodmann atlas (Strotzer, 2009) to generate a montage covering the auditory and visual cortex. For the auditory cortex, we targeted the left and right Superior Temporal Gyrus (STG) and the Middle Temporal Gyrus (MTG). For the visual cortex, we targeted the Primary Visual Cortex (V1) and the Visual Association Cortex (V2). One channel located over the right posterior temporal area (channel 16) was excluded from the analysis because it did not have a left-hemisphere counterpart. The final montage thus included fifteen measurement channels with source-detector distances ranging from 2.5 cm to 4 cm . In total, 7 channels were covering the occipital area and 4 channels per hemisphere were covering bilateral temporal areas. The detail of the location of each channel and its two highest specificities to Brodmann areas are presented in Table 1.

Lobe	Channel number	Source	Detector	Brodmann area	Specificity (%)
Left Temporal Lobe	1	C5	T7	22 - Superior Temporal Gyrus	42,20
				21 - Middle Temporal gyrus	37,02
	2	C5	CP5	48 - Retrosubicular area	36,01
				22 - Superior Temporal Gyrus	22,25
	3	TP7	T7	21 - Middle Temporal gyrus	49,28
				20 - Inferior Temporal gyrus	47,32
	4	TP7	CP5	22 - Superior Temporal Gyrus	35,29
				21 - Middle Temporal gyrus	34,85
Occipital Lobe	5	O1	PO3	18 - Visual Association Cortex (V2)	48,62
				17 - Primary Visual Cortex (V1)	40,66
	6	O1	Oz	17 - Primary Visual Cortex (V1)	78,86
				18 - Visual Association Cortex (V2)	16,25
	7	POz	PO3	18 - Visual Association Cortex (V2)	64,83
				19 - V3	17,65
	8	POz	Oz	17 - Primary Visual Cortex (V1)	60,56
				18 - Visual Association Cortex (V2)	30,10
	9	POz	PO4	18 - Visual Association Cortex (V2)	67,80
				19 - V3	15,69
	10	O2	Oz	17 - Primary Visual Cortex (V1)	67,60
				18 - Visual Association Cortex (V2)	26,74
	11	O2	PO4	17 - Primary Visual Cortex (V1)	48,05
				18 - Visual Association Cortex (V2)	45,92
Right Temporal Lobe	12	T8	C6	22 - Superior Temporal Gyrus	47,33
				21 - Middle Temporal gyrus	38,02
	13	T8	TP8	37 - Fusiform gyrus	68,11
				20 - Inferior Temporal gyrus	18,12
	14	CP6	C6	22 - Superior Temporal Gyrus	31,60
				48 - Retrosubicular area	26,83
	15	CP6	TP8	22 - Superior Temporal Gyrus	40,43
				21 - Middle Temporal gyrus	35,40
	16	P8	TP8	21 - Middle Temporal gyrus	54,96
				20 - Inferior Temporal gyrus	41,52

Table 1: *Channel specificity for Brodmann areas: each of the 16 recording channels is composed of a source and a detector (located in standard 10-20 positions). We reported in the table the two corticals structures for each channel with the highest specificity (in percent) according to the fOLD software (Zimeo Morais et al., 2018). Channels localised over the occipital lobe are highlighted in light blue. Channels over the left temporal area are highlighted in light yellow, and channels over the right temporal area are highlighted in light red. Channel 16, highlighted in light grey, was excluded from statistical analysis because it did not have a left-hemispheric counterpart.*

2.4. Procedure

The experimenter measured the participant's head circumference before the session to establish the proper cap size. The cap alignment was meticulously checked and adjusted to place the Cz optode halfway between the two tragi and between the nasion and inion. The participant was then brought into a dimly lit, sound-attenuated booth. If necessary, their hair was gently moved using a thin wooden stick to allow unobstructed access to the optode locations. Optodes were positioned according to the pre-defined montage (see Figure 1 and Table 1). The fNIRS signal was calibrated and monitored for quality, and optode placement was adjusted as needed until satisfactory signal quality was achieved before continuing. Participants were instructed to limit their movements, avoid talking, listen carefully, and keep their eyes on the screen during the experiment while they were seated comfortably. Participants were given a 60-second baseline period to focus on a central crosshair before the experiment started. Passive listening and viewing only were necessary for the task itself.

Participants completed a continuous sequence of 45 trials. This sequence consisted of 15 auditory-only (A) trials, 15 visual-only (V) trials, and 15 audiovisual (AV) trials. The presentation order was pseudo-randomized, adhering to the constraint that each consecutive group of three trials contained one trial from each sensory modality (A, V, AV), with the specific order within each triplet being random. Following each 10-second stimulus presentation, there was an inter-stimulus interval (ISI) with a duration jittered between 7 and 10 seconds. During this ISI, participants were instructed to fixate their gaze on a static black-and-white image of a cartoon planet Earth displayed in the centre of the screen. The total duration of the experimental task was 12 to 15 minutes.

2.5. fNIRS data preprocessing

fNIRS data were preprocessed using the NIRS Brain-AnalyzIR toolbox (Santosa et al., 2018) along with custom-written scripts on Matlab R2021a (available at https://github.com/yanndavlem/nirs_article). We first converted raw intensity signals to changes in optical density. To address motion artifacts caused by excessive head movements, we used Temporal Derivative Distribution Repair (TDDR, Fishburn et al., 2019), a robust regression method aimed at removing large fluctuations in the optical density signal while preserving smaller fluctuations indicative of hemodynamic activity. The corrected optical density was then band-pass filtered between 0.01 and 0.12 Hz to remove cardiac activity (~1.2 Hz) and respiratory activity (~0.25 Hz). Subsequently, the corrected and filtered optical densities were transformed into (de)oxygenated hemoglobin concentrations using the modified Beer-Lambert Law. After visually inspecting the preprocessed data, no data were excluded from subsequent analyses.

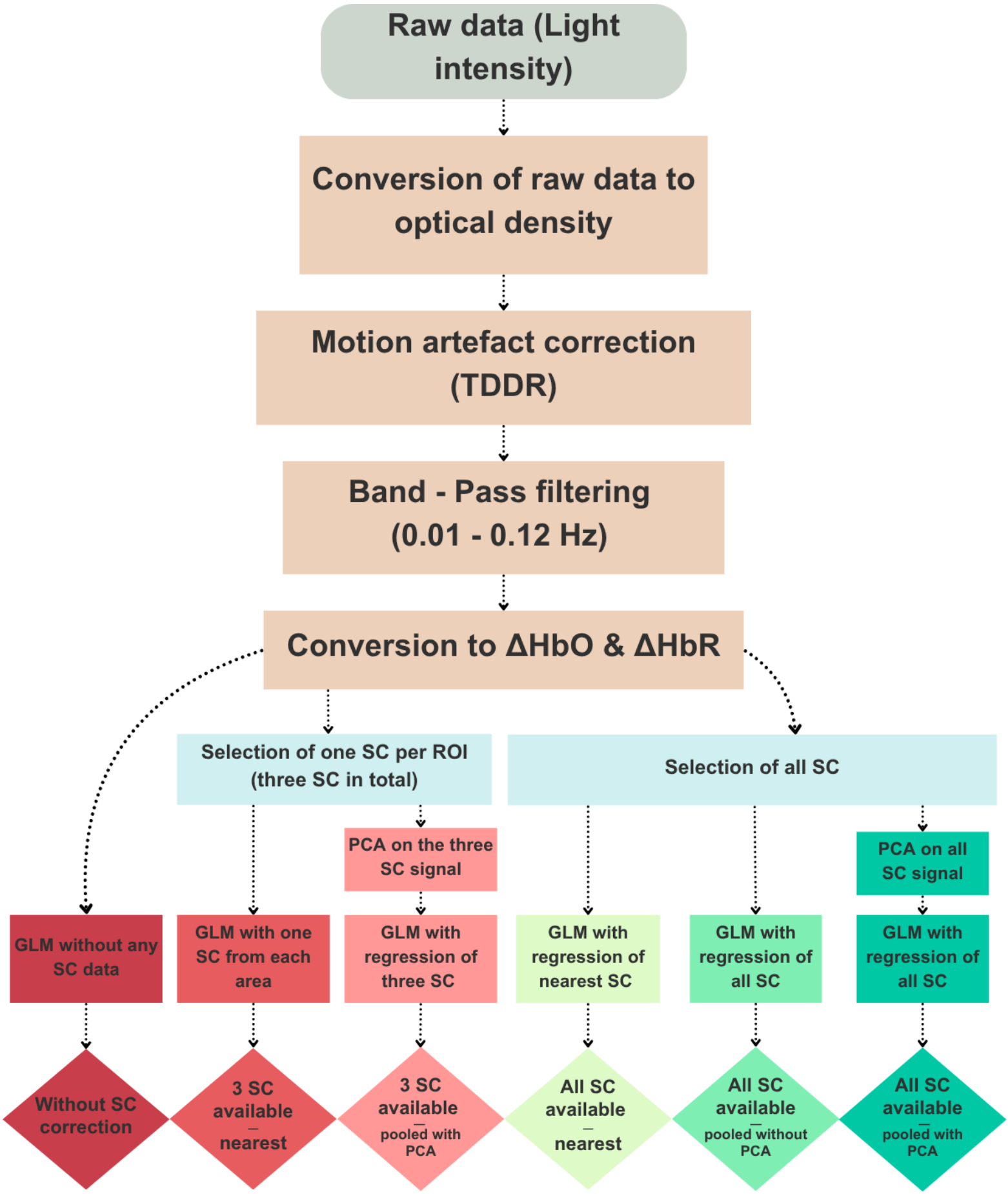

Raw data (Light intensity)
Conversion of raw data to optical density
Motion artefact correction (TDDR)
Band - Pass filtering (0.01 - 0.12 Hz)
Conversion to ΔHbO & ΔHbR
Selection of one SC per ROI (three SC in total)
Selection of all SC
PCA on the three SC signal
PCA on all SC signal
GLM without any SC data
GLM with one SC from each area
GLM with regression of three SC
GLM with regression of nearest SC
GLM with regression of all SC
GLM with regression of all SC
Without SC correction
3 SC available – nearest
3 SC available – pooled with PCA
All SC available – nearest
All SC available – pooled without PCA
All SC available – pooled with PCA

Figure 2: *Preprocessing and analysis pipelines of the fNIRS data. A standard pipeline was used to convert raw signals to HbO and HbR concentration changes (ΔHbO, ΔHbR). The main statistical analysis (GLM) was performed using either no SC correction, three SCs (to simulate potential hardware limitations), or all eight available SCs. Two correction approaches were applied: a spatial-specific method, in which each LC was regressed using the signal from the nearest SC, and a non-spatial-specific method, in which signals from all available SCs were pooled to regress each LC. In the latter case, we additionally assessed the impact of applying a PCA to the SC signal. This resulted in six distinct pipelines: (1) No SC correction, (2) 3 SCs – nearest (spatial-specific), (3) 3 SCs – pooled with PCA (non-spatial-specific), (4) All SCs – nearest (spatial-specific), (5) All SCs – pooled without PCA (non-spatial-specific), and (6) All SCs – pooled with PCA (non-spatial-specific). SC: short-channels; LC: long-channels; PCA: principal component analysis; GLM: generalized linear model. TDDR: temporal derivative distribution repair.*

2.6. Regression of systemic signal with short-channels

We conducted six types of Generalized Linear Model (GLM) analyses to evaluate the impact of SC regression on fNIRS signals (Figure 2). All analyses were performed using Ordinary Least Squares (OLS), a fundamental method for estimating the relationship between a dependent variable and one or more independent variables by minimizing the residual sum of squares (von Eye et al., 2023). This approach allowed us to obtain beta coefficients that quantify the relationship between the experimental conditions and the hemodynamic responses.

The predictor variables were defined using a boxcar function convolved with the canonical Hemodynamic Response Function (HRF), allowing us to generate a Statistical Parametric Mapping of the HRF (SPM-HRF; Santosa et al., 2018) for each of the three conditions (A, V, and AV). In all analyses, a GLM was applied to each long-channel and each chromophore (HbO, HbR) with one HRF-convolved regressor for each condition. The inter-stimulus interval (ISI) served as the implicit baseline in all analyses.

In addition to the three regressors for each condition, the number and nature of the regressors varied according to the SC strategy used, ranging from 0 (no SC regression) to up to 16 regressors (when including all available SCs of both chromophores, with or without PCA). To evaluate different strategies for SC correction, we compared two main scenarios: one in which only three SCs were available (to simulate hardware limitations) and another using all eight available SCs (one per source). Within each scenario, we applied either a spatial-specific approach, where each long channel was corrected using the signal from its anatomically closest SC to account for region-specific systemic fluctuations, or a non-spatial-specific approach, where all SC signals were pooled to model widespread physiological interference across the montage. In the latter case, when all SC signals were available, the GLM was run with and without applying PCA to evaluate its impact on capturing global physiological components shared across sources.

Six SC correction strategies were thus tested:

1. **No SCs (baseline condition)**

   This reference analysis involved no SC regression. Only standard pre-processing steps and a GLM with the three condition regressors were applied.

2. Limited SC availability (n = 3)

For the next two analyses, we simulated scenarios in which only three SCs were available, reflecting potential hardware limitations or experimental constraints. In this configuration, one SC was selected per lobe based on its central location relative to surrounding long channels (LCs): the SC near source C5 was used for the left temporal region (channels 1–4), the SC at POz for the occipital region (channels 5–11), and the SC at CP6 for the right temporal region (channels 12–15).

a. Spatial-specific approach

For each long channel and each chromophore, the signal from the anatomically nearest SC for the same chromophore (i.e., HbO regressed with HbO, HbR with HbR) was included as a regressor of no interest in the GLM. This chromophore-specific approach follows the implementation of the nearest-SC inclusion in a GLM in Homer3 (Huppert et al., 2009). The nearest SC was defined as the one selected for the same lobe in which the long channel was located.

b. Non-spatial-specific approach

For each long channel and each chromophore, the signals from both chromophores (HbO and HbR) of the three available SCs (i.e., 3 HbO and 3 HbR signals) were combined and orthogonalized using PCA, following the NIRS brain AnalyzIR (Santosa et al., 2018) implementation. The resulting six components were then included as regressors of no interest in the GLM.

3. Full SC availability (n = 8):

In the next three analyses, the signals from all available SCs (n = 8, one per source) were used.

a. Spatial-specific approach

For each long channel and each chromophore, the signal from the anatomically nearest SC, i.e., the one physically attached to the optode source of the long channel, for the same chromophore was used as a regressor of no interest in the GLM.

b. non-spatial-specific approach without PCA

For each long channel and each chromophore, the signals from both chromophores (HbO and HbR) of the eight available SCs (i.e., 8 HbO and 8 HbR signals) were combined and used as regressors of no interest in the GLM. Note that no PCA was applied to the SCs' signal.

c. non-spatial-specific approach with PCA

For each long channel and each chromophore, the signals from both chromophores (HbO and HbR) of the eight available SCs (i.e., 8 HbO and 8 HbR signals) were combined and orthogonalized using PCA. The resulting sixteen components were then included as regressors of no interest in the GLM.

2.7. Statistical analysis

All analyses described above yielded a beta value for each condition regressor (A, V, AV), for each subject, long channel, and chromophore. These beta estimates were then submitted to Bayesian t-tests, either one-sample tests against zero or two-sample tests comparing conditions (see next section for details), to quantify a degree of logical support or belief for an effect of interest. We report Bayes Factors (BF_{10}) as a relative measure of evidence in favor of the alternative hypothesis compared to the null model (Wagenmakers et al., 2018). Traditionally, a BF_{10} between 1 and 3 is considered weak evidence for the tested model, between 3 and 10 as positive evidence, between 10 and 100 as strong evidence, and greater than 100 as decisive evidence. Similarly, to interpret the strength of evidence in favor of the null model, a BF_{10} between 0.33 and 1 is considered weak evidence, a BF between 0.01 and 0.33 as positive evidence, a BF between 0.001 and 0.01 as strong evidence, and a BF lower than 0.001 as decisive evidence (Lee & Wagenmakers, 2014).

2.8. Pipeline quality metrics and validation

To compare the different analysis pipelines, we computed two evaluation metrics for each:

- Metric 1: emergence tests; we counted the number of channels showing evidence for an effect in any condition (relative to baseline) in the expected direction.
  We ran emergence tests in each channel and for each condition by performing one-tailed, one-sample Bayesian t-tests against 0 across subjects. A channel was considered as showing evidence for an effect in the expected direction as follows:
  - **For the temporal lobes:**
    BF_{10} > 3 for condition A compared to 0 (A > 0 for HbO, A < 0 for HbR)
    OR BF_{10} > 3 for condition AV compared to 0 (AV > 0 for HbO, AV < 0 for HbR)
  - **For the occipital lobe:**
    BF_{10} > 3 for condition V compared to 0 (V > 0 for HbO, V < 0 for HbR)
    OR BF_{10} > 3 for condition AV compared to 0 (AV > 0 for HbO, AV < 0 for HbR)
- Metric 2: between-condition difference tests; we counted the number of channels where there was evidence for a difference between conditions in the expected direction.
  Comparisons between conditions were only conducted in channels that showed evidence of an effect for at least one condition relative to the baseline in the expected direction (as defined in Metric 1).
  Specifically, we performed pairwise comparisons between conditions in each channel selected by Metric 1, using one-tailed, two-sample Bayesian t-tests across subjects. Channels were considered showing an effect in the expected direction as follows:

- **For the temporal lobes**:
  BF_{10} > 3 for condition AV vs V (AV > V for HbO, AV < V for HbR)
  OR BF_{10} > 3 for condition A vs V (A > V for HbO, A < V for HbR)
- **For the occipital lobe**:
  BF_{10} > 3 for condition AV vs A (AV > A for HbO, AV < A for HbR)
  OR BF_{10} > 3 for condition A vs V (V > A for HbO, V < A for HbR)

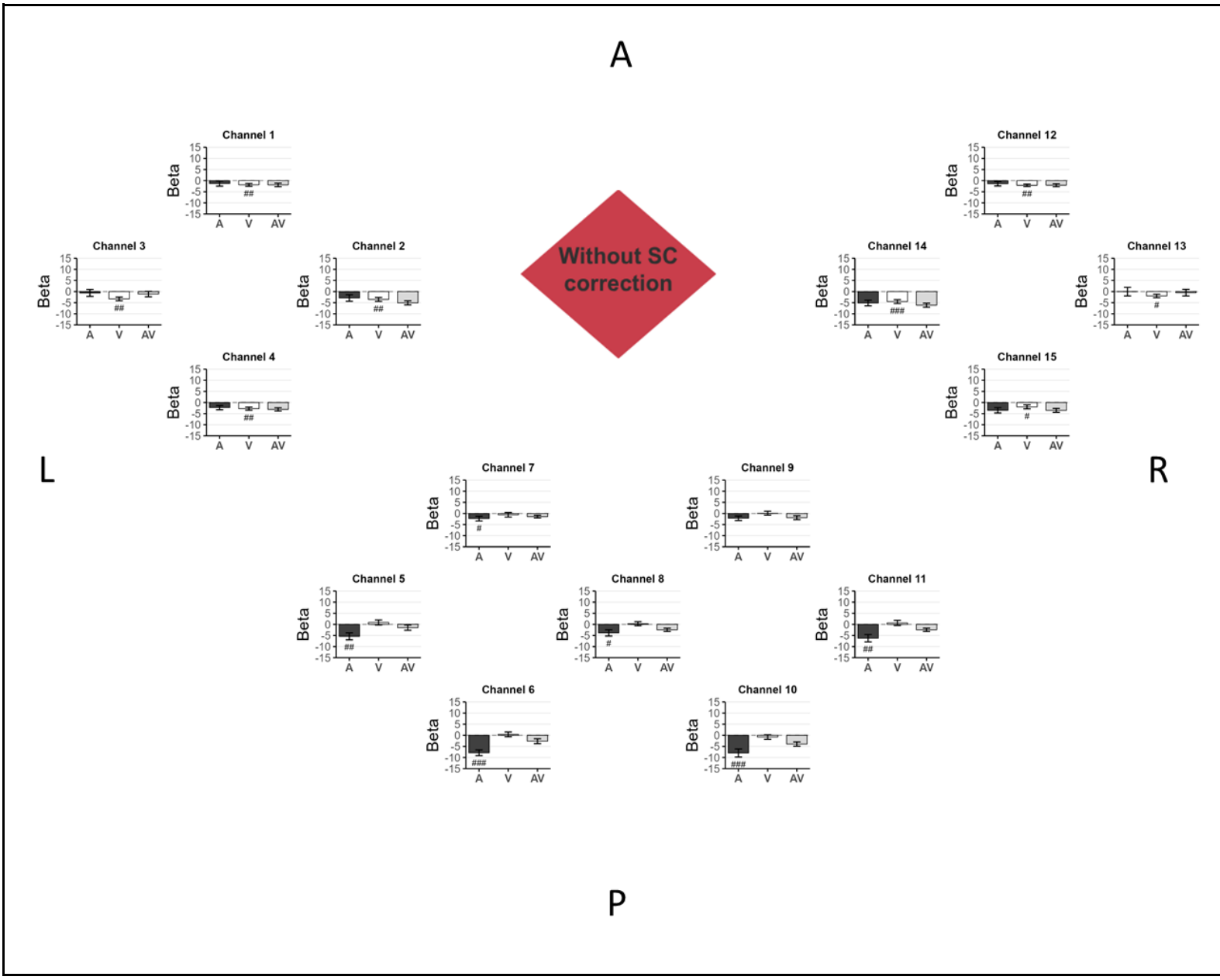
3. Results

We report only the HbO results in the main text. HbO results for all channels are presented in Figures 3 to 8. A summary of the validation metrics is provided in Figure 9. HbR results are available in the Supplementary Materials and discussed in the main text.

3.1. Baseline: no SC correction

Beta values for the A, V, AV conditions from the regression of HbO signals in the analysis without using SCs in the GLM analysis are represented in Figure 3 for all channels (HbR results for the corresponding analysis are represented in Supplementary Figure S1). The summary of the validation metrics is presented in Figure 9 for HbO and Figure S7 for HbR.

Most channels exhibited negative beta values in response to auditory, visual, and audiovisual stimulation. No channel in the temporal lobes or the occipital lobe showed emergence of activity in the expected direction (metric 1). As no channel showed evidence of emergence in the expected direction (metric 1), we did not proceed with testing for differences between conditions (metric 2).



<u>Figure 3:</u> Baseline: no SC correction (HbO results). *In this bird's-eye view of the fNIRS montage, channels are positioned according to their approximate location on the scalp, with labels indicating the anterior (A), posterior (P), left (L), and right (R) regions of the montage. Channel 1 to 4 are located above the left temporal auditory areas, Channel 5 to 11 are located above the occipital area, channel 12 to 15 are located above the right temporal auditory area. Each graph represents the beta estimated from HbO signal for the auditory (A, black color), visual (V, white color) and audiovisual (AV, grey color) stimulation, averaged across the 16 participants. For each channel and each condition, a # symbol is displayed when the Bayes Factor (BF_{10}) from a one-tailed, one-sample Bayesian t-test against 0 exceeds 3 in the expected direction (see method section 2.8), ## when $BF_{10} > 10$, and ### when $BF_{10} > 100$. Error bars indicate the standard error of the mean for each condition.*

3.2. Limited short-channel Availability (n = 3)

The next two sections simulate a scenario where not all SCs are available. Specifically, we examined a configuration with only one SC per lobe (n = 3), located in the left temporal, right temporal, and occipital areas. These sections present results from two approaches under this constraint: a spatial-specific regression using the nearest SC for each long channel, and a non-spatial-specific approach using PCA-transformed components of all available SC signals.

3.2.1. Spatial-specific approach

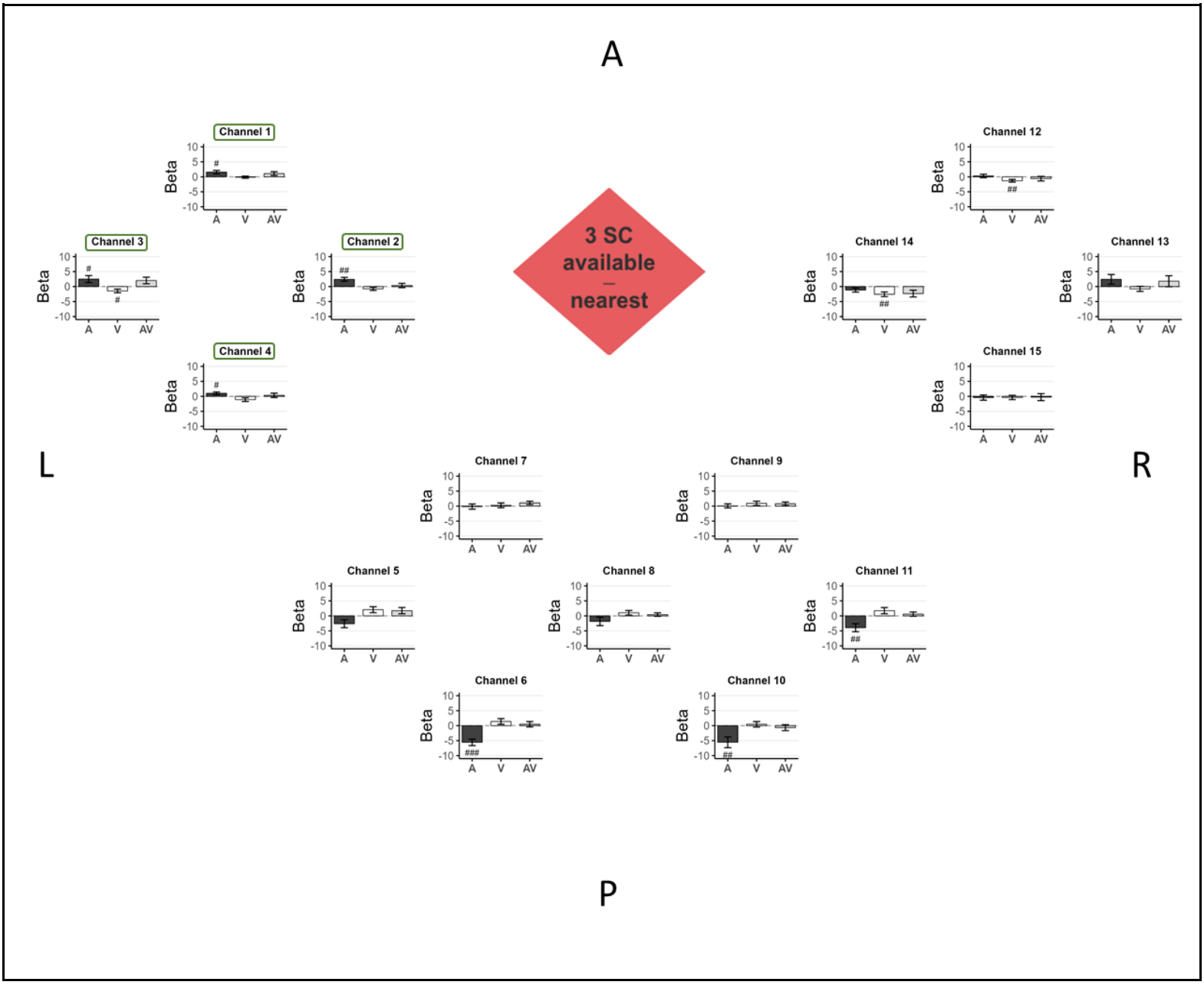
Beta values for the A, V, and AV conditions from the regression of HbO signals under the limited SC availability scenario, using the spatial-specific analysis with the nearest available SC for each long-channel regression in the GLM, are shown in Figure 4. HbR results for the corresponding analysis are presented in Supplementary Figure S2. The summary of the validation metrics is presented in Figure 9 for HbO and Figure S7 for HbR.

In the left temporal area, all four channels showed evidence for emergence of activity in the expected direction (metric 1) in condition A vs. 0 ($BF_{10} > 3$; see Figure 4). For between-condition differences (metric 2), all of those channels showed positive to strong evidence for higher beta values in condition A vs. V ($7.55 < BF_{10} < 77.03$) and one (channel 3) showed strong evidence for higher beta values in condition AV vs. V ($BF_{10} = 10.42$).

In the right temporal area, no channel showed emergence of activity in the expected direction.

In the occipital area, no channel showed emergence of activity in the expected direction.

As no channels showed evidence of emergence in the expected direction (metric 1) for the occipital area and right temporal area, we did not proceed with testing for differences between conditions for these regions of interest (metric 2).



<u>Figure 4:</u> Limited short-channel availability - spatial-specific approach (HbO results). *Same as in Figure 3, but for the scenario simulating the availability of one short-channel per lobe. The analysis used the nearest short-channel to each long-channel for regression. A green box surrounding the channel label indicates that the channel showed evidence for an expected effect based on metric 1 described in Section 2.8.*

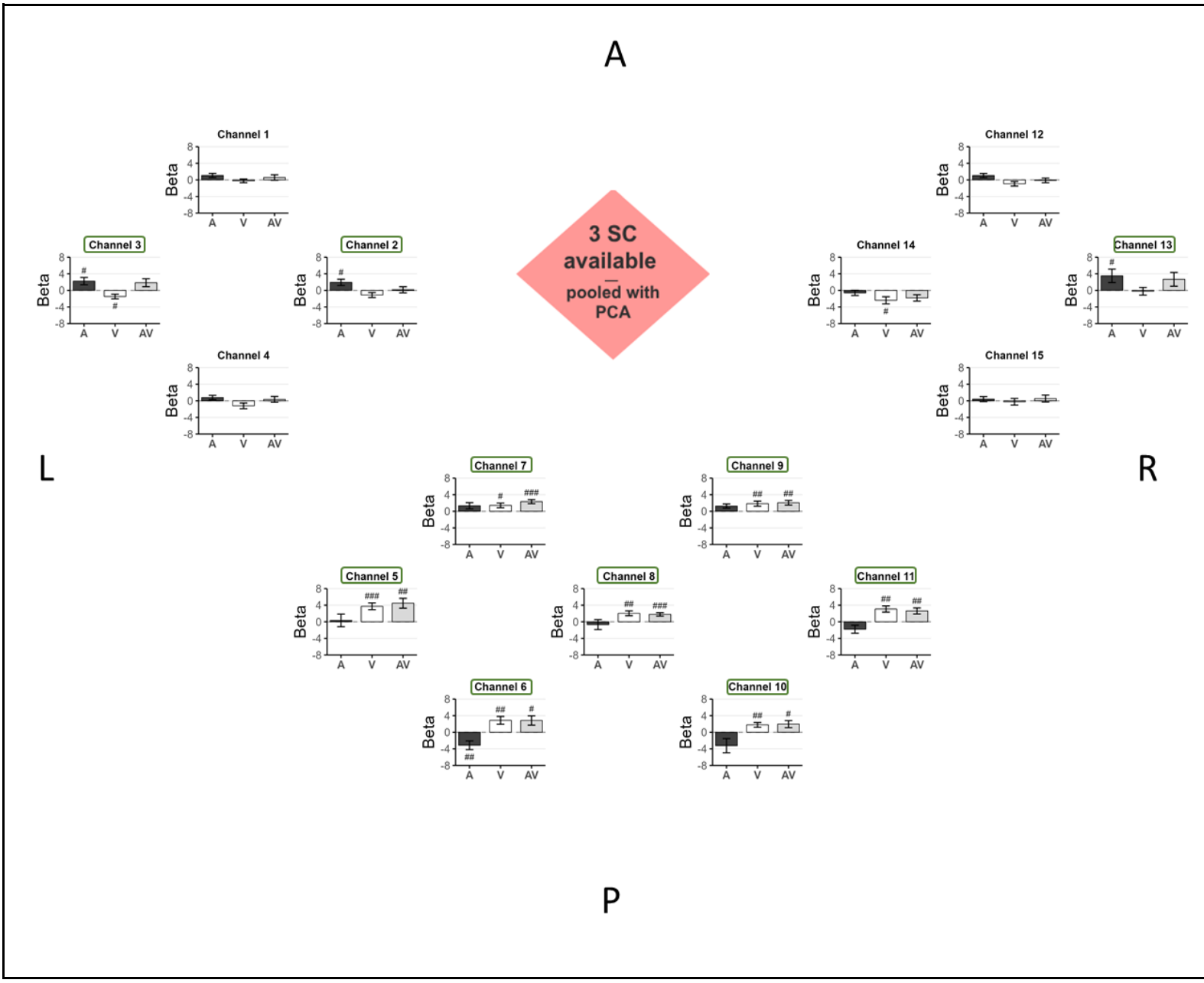
<u>3.2.2. Non-spatial-specific approach</u>

Beta values for the A, V, and AV conditions from the regression of HbO signals under the limited SC availability scenario, using the non-spatial-specific analysis with PCA components of all available SCs for each long-channel regression in the GLM, are shown in Figure 5. HbR results for the corresponding analysis are provided in Supplementary Figure S3. The summary of the validation metrics is presented in Figure 9 for HbO and Figure S7 for HbR.

In the left temporal area, two channels (channels 2 and 3) showed evidence for emergence of activity in the expected direction (metric 1) in condition A vs. 0 ($BF_{10} > 3$; see Figure 5). For the between-condition differences (metric 2), those two channels showed strong evidence for higher beta values in condition A vs. V ($22.3 < BF_{10} < 53.4$), and one (channel 3) showed strong evidence for higher beta values in condition AV vs. V ($BF_{10} = 15.9$).

In the right temporal area, one channel (channel 13) showed evidence for emergence of activity in the expected direction (metric 1) in condition A vs. 0 ($BF_{10} > 3$; see Figure 5). However, this channel did not show evidence of between-condition differences (metric 2).

In the occipital area, all seven channels showed evidence for the emergence of activity in the expected direction (metric 1) in conditions V vs. 0 and AV vs. 0 ($BF_{10} > 3$; see Figure 5). Among them, for the between-condition differences (metric 2), four channels (channels 6, 8, 10 and 11) showed positive to decisive evidence for higher beta in condition V vs. A ($3.03 < BF_{10} < 245.73$), and four channels (channels 5, 6, 10 and 11) showed positive to decisive evidence for higher beta in condition AV vs. A ($3.57 < BF_{10} < 105.89$).



<u>Figure 5:</u> Limited short-channel availability - non-spatial-specific approach (HbO results). *Same as in Figure 4, but showing results from the analysis using PCA-transformed components of all available short-channel signals for regression.*

3.3. Full short-channel availability (n = 8)

The next three sections present results based on analyses where all SCs were available for regression, using (1) a spatial-specific approach, (2) a non-spatial-specific approach with raw SC signals, and (3) a non-spatial-specific approach with PCA-transformed components of all SC signals.

3.3.1 Spatial-specific approach

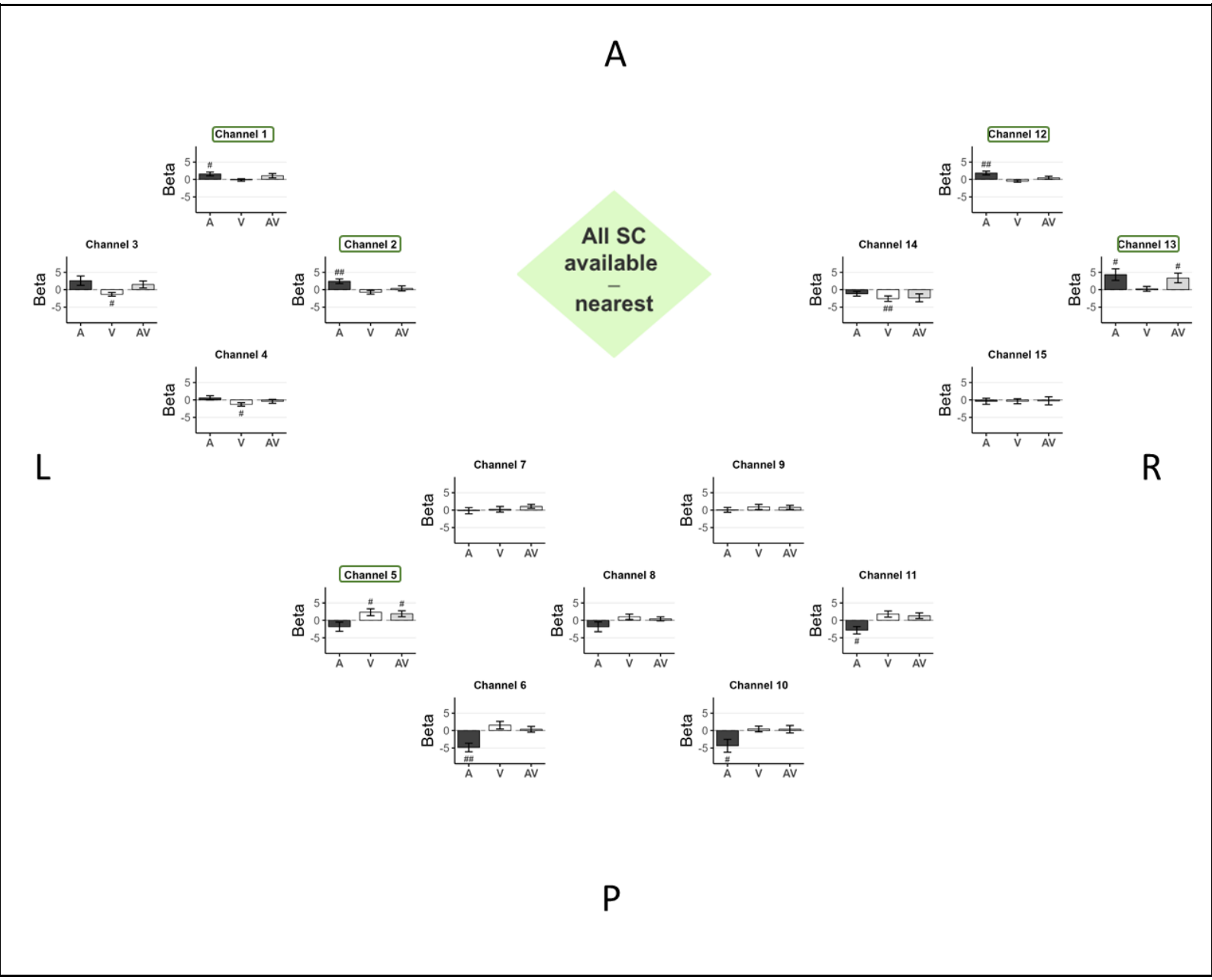
Beta values for the A, V, and AV conditions from the regression of HbO signals under the full SC availability scenario, using the spatial-specific analysis with the nearest

available SC for each long-channel regression in the GLM, are shown in Figure 6. HbR results for the corresponding analysis are provided in Supplementary Figure S4. The summary of the validation metrics is presented in Figure 9 for HbO and in Figure S7 for HbR.

In the left temporal area, two channels (channels 1 and 2) showed evidence for emergence of activity in the expected direction (metric 1) in condition A vs. 0 (BF_{10} > 3; see Figure 6). Only channel 3 showed positive evidence for the emergence of activity (metric 1) in the expected direction in condition AV vs 0 (BF10 > 3; see Figure 6). For between-condition differences (metric 2), those two channels showed positive to strong evidence for higher beta values in condition A vs. V (7.55 < BF_{10} < 77.03).

In the right temporal area, two channels (channels 12 and 13) showed evidence for emergence of activity in the expected direction (metric 1) in conditions A vs. 0 and/or AV vs. 0 (BF_{10} > 3; see Figure 6). For between-condition differences (metric 2) those two channels showed positive to strong evidence for higher beta values in condition A vs. V (4.33 < BF_{10} < 53.4).

In the occipital area, one channel (channel 5) showed evidence for the emergence of activity in the expected direction (metric 1) in conditions V vs. 0 and AV vs. 0 (BF_{10} > 3; see Figure 6). For between-condition differences (metric 2), this channel showed positive evidence for higher beta values in condition V vs. A (BF_{10} = 6.67) as well as AV vs. A (BF_{10} = 5.11).



<u>Figure 6:</u> Full short-channel availability - spatial-specific approach (HbO results). *Same as in Figure 4, but for the scenario with all short-channel availability. The analysis used the nearest short-channel to each long-channel for regression.*

<u>3.3.2. Non-spatial-specific approach without PCA</u>

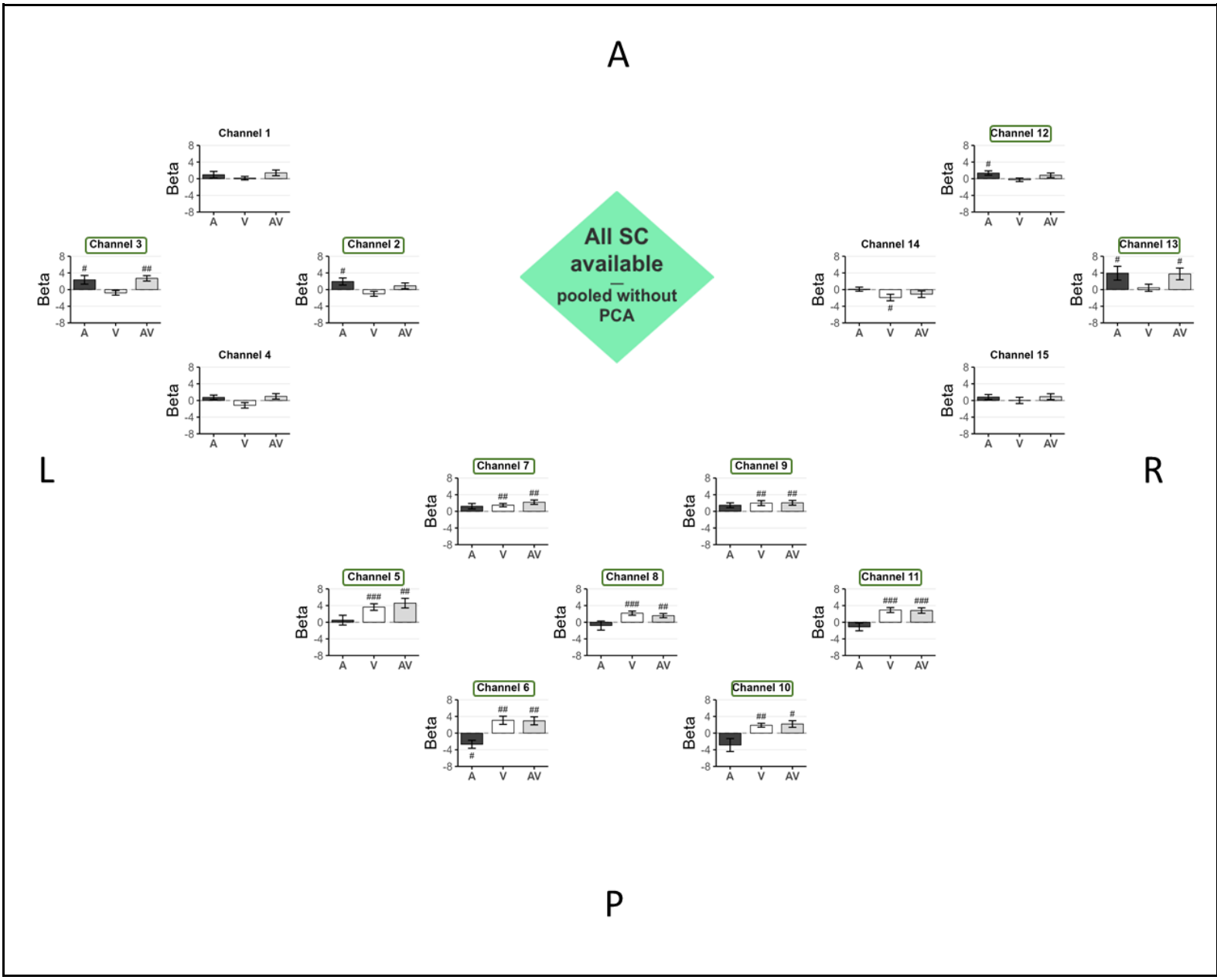
Beta values for the A, V, and AV conditions from the regression of HbO signals under the full SC availability scenario, using the non-spatial-specific analysis with the raw signal of all available SCs for each long-channel regression in the GLM, are shown in Figure 7. HbR results for the corresponding analysis are provided in Supplementary Figure S5. The summary of the validation metrics is presented in Figure 9 for HbO and in Figure S7 for HbR.

In the left temporal area, two channels (channels 2 and 3) showed positive evidence for emergence of activity in the expected direction (metric 1) in conditions A vs. 0 and/or AV vs. 0 (BF_{10} > 3; see Figure 7). For between-condition differences (metric 2), those

two channels showed positive to strong evidence for higher beta in condition A vs. V ($8.47 < BF_{10} < 12.53$), and positive to decisive evidence for higher beta in condition AV vs. V ($3.39 < BF_{10} < 172.15$).

In the right temporal area, two channels (channels 12 and 13) showed positive evidence for emergence of activity in the expected direction (metric 1) in conditions A vs. 0 and/or AV vs. 0 ($BF_{10} > 3$; see Figure 7). For between-condition differences (metric 2) channel 12 showed positive evidence for higher beta in condition AV vs. V ($BF_{10} = 5.87$).

In the occipital area, all seven channels showed evidence for the emergence of activity in the expected direction (metric 1) in conditions V vs. 0 and AV vs. 0 ($BF_{10} > 3$; see Figure 7). Among them, for between-condition differences, five channels (channels 5, 6, 8, 10 and 11) showed positive to decisive evidence for higher beta in condition V vs. A ($4.04 < BF_{10} < 217.87$), and among them, four channels (channel 5, 6, 10 and 11) showed positive to decisive evidence for higher beta in condition AV vs. A ($6.25 < BF_{10} < 188.82$).



<u>Figure 7:</u> Full short-channel availability – non-spatial-specific approach without PCA (HbO results). *Same as in Figure 4 but showing results from the analysis using raw signal of all available short-channel signals for regression.*

<u>3.3.3. Non-spatial-specific approach with PCA</u>

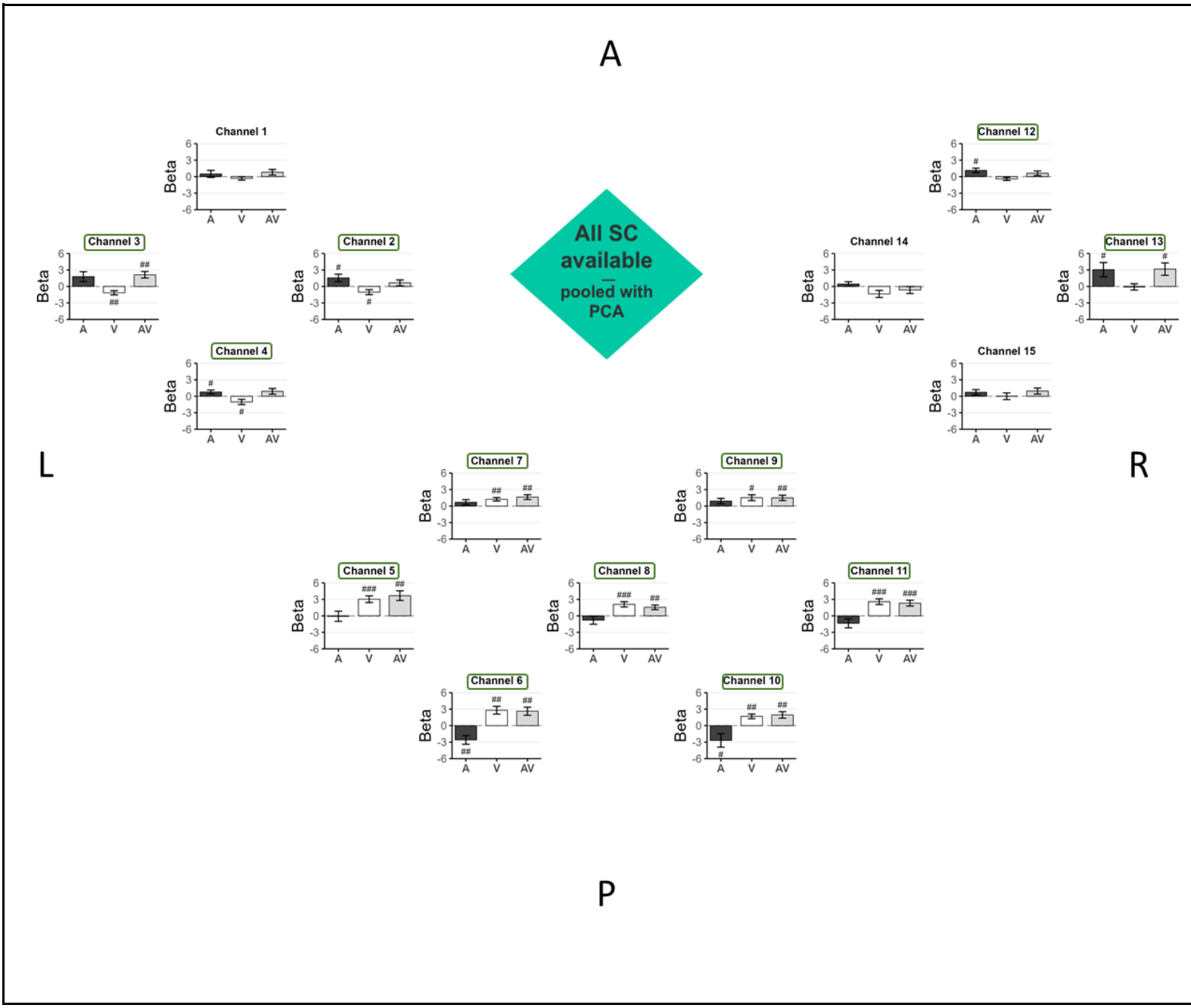
Beta values for the A, V, and AV conditions from the regression of HbO signals under the full SC availability scenario, using the non-spatial-specific analysis with PCA components of all available SCs for each long-channel regression in the GLM, are shown in Figure 8. HbR results for the corresponding analysis are provided in Supplementary Figure S6. The summary of the validation metrics is presented in Figure 9 for HbO and in Figure S7 for HbR.

In the left temporal area, two channels (channels 2 and 4) showed evidence for emergence of activity in the expected direction (metric 1) in conditions A vs. 0 and/or AV vs. 0 (BF_{10} > 3; see Figure 8). For between-condition differences (metric 2), those three channels showed strong evidence for higher beta values in condition A vs. V

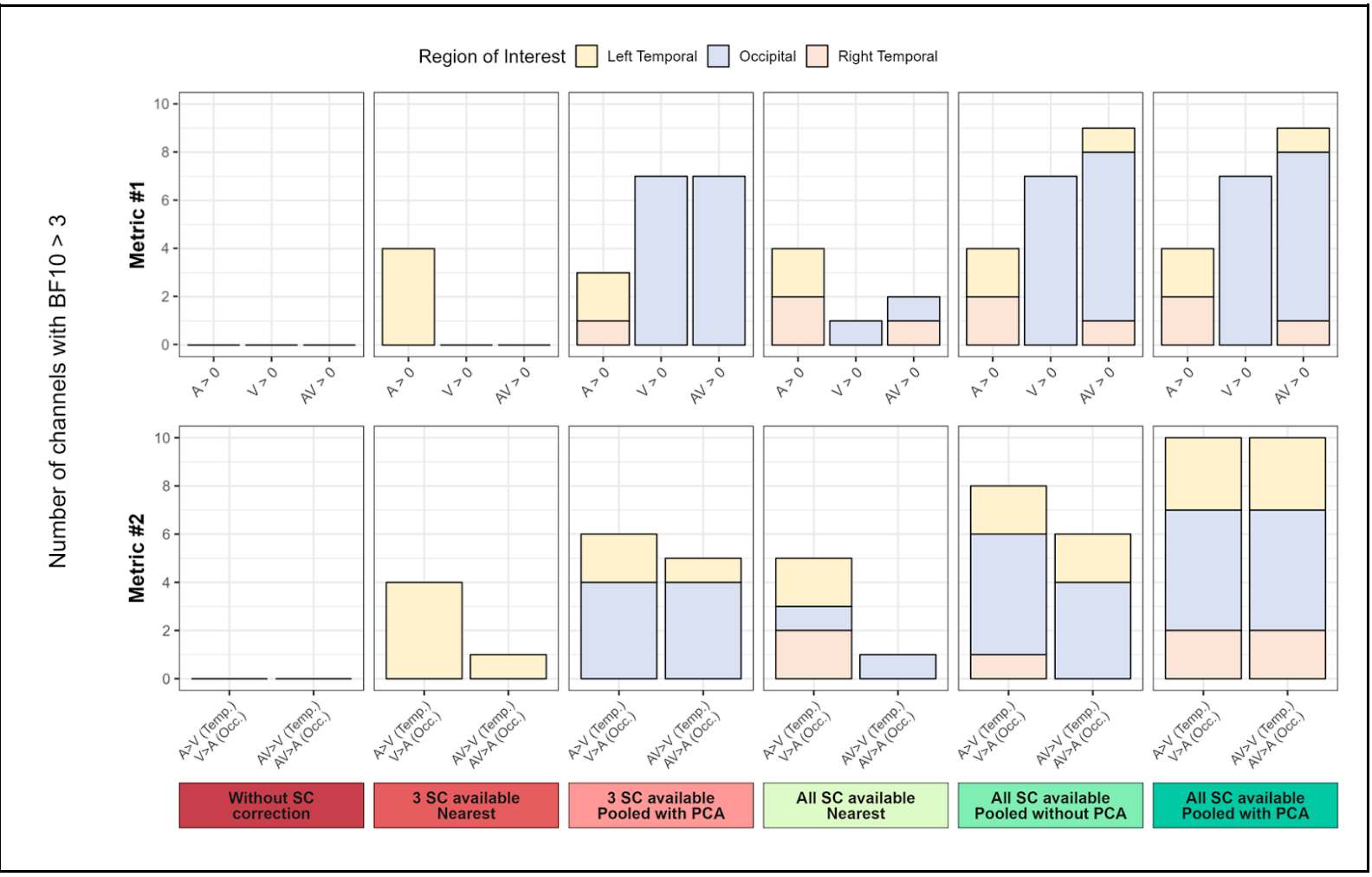
(16.88 < BF_{10} < 24.75), and all three showed positive to decisive evidence for higher beta values in condition AV vs. V (5.65 < BF_{10} < 625.71).

In the right temporal area, two channels (channels 12 and 13) showed evidence for emergence of activity in the expected direction (metric 1) in condition A vs. 0 and/or AV vs. 0 (BF_{10} > 3; see Figure 8). For between-condition differences (metric 2), those two channels showed positive to strong evidence for higher beta values in condition A vs. V (3.97 < BF_{10} < 23.65), and the same two channels showed positive evidence for higher beta values in condition AV vs. V (3.03 < BF_{10} < 7.04).

In the occipital area, all seven channels showed evidence for the emergence of activity in the expected direction (metric 1) in condition V vs. 0 and AV vs. 0 (BF_{10} > 3; see Figure 8). Among them, for between-condition differences (metric 2), five channels (channels 5, 6, 8, 10 and 11) showed strong to decisive evidence for higher beta values in condition V vs. A (27.43 < BF_{10} < 1765.72), and the same five channels showed positive to decisive evidence for higher beta values in condition AV vs. A (6.25 < BF_{10} < 188.82).



<u>Figure 8:</u> Full short-channel availability – non-spatial-specific approach with PCA (HbO results). *Same as in Figure 4, but showing results using PCA-transformed components of all available short-channel signals for regression.*



<u>Figure 9:</u> Summary of validation metrics (HbO results). *The top panels display, for each analysis pipeline and across all regions of interest, the number of channels showing evidence of activation ($BF_{10} > 3$) based on a one-tailed, one-sample Bayesian t-test against zero (metric 1). Conditions included Auditory (A), Visual (V), and Audiovisual (AV) stimulation, and tests were directed toward positive effects (A > 0, V > 0, AV > 0). Channels were retained for further analysis only if they showed a $BF_{10} > 3$ in the expected direction—specifically, A > 0 or AV > 0 in temporal areas, and V > 0 or AV > 0 in occipital areas. The bottom panels display, for each analysis pipeline and across regions of interest, the number of channels (among those showing expected emergence in metric 1) with evidence of differential activation ($BF_{10} > 3$) based on one-tailed, two-sample Bayesian t-tests between conditions. The tests were directed as follows: A > V and AV > V in temporal areas, and V > A and AV > A in the occipital area.*

<u>4. Discussion</u>

The goal of the present study was to systematically assess the impact of short-channel (SC) regression on statistical outcomes in functional near-infrared spectroscopy (fNIRS) studies, using a naturalistic audiovisual paradigm designed to elicit distinct activation patterns in the temporal and occipital cortices. By comparing six GLM-based preprocessing strategies that varied in the number, anatomical specificity, and dimensionality reduction of SC signals, we aimed to identify the most effective pipeline for isolating neural activity from extracerebral physiological noise captured with SCs.

To robustly evaluate each strategy's efficiency, we relied on two complementary Bayesian metrics. The first metric assessed the emergence of task-related activity in the expected direction, allowing us to identify channels that exhibited within-condition activation consistent with the experimental design. The second metric, applied only to the subset of channels identified via the first, quantified the strength and reliability of between-condition differences. With the HbO signal, our results revealed that the absence of SC regression led to widespread negative beta estimates and no detectable emergence of neural activity in expected regions, highlighting the critical impact of physiological noise on the fNIRS signal. In contrast, most pipelines that included SC regression produced zero-centered, interpretable beta distributions, enabling detection of the expected task-related activity. When comparing a spatially specific approach (using the nearest SC) to a non-spatially specific approach (pooling all SCs together) in the GLM, we found that the non-spatially specific method yielded more robust performance across both Bayesian metrics, regardless of whether a limited number of SCs or all available SCs were used. Finally, under conditions of full SC availability, applying a dimensionality reduction technique to the SC signals yielded a slight additional improvement in results, suggesting that identifying shared components across SCs, rather than leaving them mingled in variable proportions, leads to a more accurate estimation of extra-cerebral signals and enhance denoising without compromising spatial specificity. Importantly, while the HbR signal was overall less impacted by the presence or absence of SC correction, similar trends were observed, and the same SC regression strategy that maximized HbO signal interpretability also yielded the most robust and consistent results for HbR.

4.1. The impact of omitting SC regression

Our results demonstrate that omitting SC regression has a profound impact on the interpretability and reliability of fNIRS data. When no SC correction was applied to the HbO signal, beta estimates were systematically negative across the cortex, with no detectable emergence of neural activity in regions where activation was expected based on the task design (Figure 3). Critically none of the channels passed the criteria of our first metric (i.e., emergence in the expected direction) when the SC signal was not included in the regression (Figure 9). This indicates a complete failure of the uncorrected pipeline to recover expected neural activation from the data.

This failure of the uncorrected pipeline is consistent with previous studies showing that signal cleaning, particularly using SC regression, is essential for recovering task-related neural activity. SC regression consistently outperforms uncorrected pipelines across a variety of paradigms. Whether based on averaging, individual SC modeling, or principal component decomposition of SC signals, regressing out superficial components using SC reliably improves signal quality and interpretability (Luke et al., 2021; Wyser et al., 2020; Santosa et al., 2020). It also surpasses alternative correction methods such as common average referencing or global channel regression, in the case of motor tasks targeting sensorimotor areas (Klein et al., 2022). When SCs are not available, applying any regression method is still markedly better than leaving the signal uncorrected. Global signal regression (Zhou et al., 2020), ICA-based modeling of global scalp hemodynamics (Kohno et al., 2007), or spatial filtering technique (Zhang et al., 2009) all result in improved signal quality compared to raw data (Sato et al., 2016). Although our study did not directly compare SC regression with other correction methods, our results, combined with extensive prior evidence, support the conclusion

that SC regression remains the most robust and reliable approach to isolating neural activity in fNIRS.

4.2. Spatial-specific vs. non-spatial SC regression under limited and full SC availability

We compared two GLM-based approaches to SC regression: a spatial-specific method that assigns the nearest SC to each long channel, and a non-spatial-specific method that models the shared variance of all SCs using principal component analysis (PCA). While the spatial-specific approach assumes that local SCs best reflect physiological noise affecting nearby long channels, the PCA-based method assumes that common noise components can be estimated more globally across SCs. In the spatial-specific method, we regressed the long channel signal with the signal from the nearest SC in a chromophore-specific manner (i.e., HbO signals were regressed using the nearest HbO SC, and similarly for HbR) following the implementation used in Homer 3 (Huppert et al., 2009). Notably, most published studies applying SC regression do not specify whether their regressors are chromophore-specific, and reporting this methodological detail would be a valuable practice to adopt in future work.

Under limited SC availability (n = 3; one SC per lobe), the spatial-specific method recovered auditory-related responses only in the left temporal region, with no channel showing emergence in right temporal or occipital areas. In contrast, the non-spatial-specific method not only replicated left temporal effects (though slightly weaker) but also showed emerging activity in the right temporal lobe and robust responses across the occipital cortex, with four out of seven channels meeting strong to decisive evidence thresholds for between-condition effects (V > A, AV > A). These results indicate that even with minimal SC availability, pooling SC signals and using their shared variance across channels can more effectively capture and remove systemic noise than relying solely on local proximity.

Under full SC coverage (n = 8), the non-spatial PCA-based approach led to the clearest and most widespread emergence of task-related responses across all regions. It revealed strong between-condition effects in the left and right temporal cortices as well as in the occipital lobe, showing broad spatial coverage and consistent activation patterns. Notably, the spatial-specific method under full SC availability failed to reach the performance of the non-spatial PCA-based method under full availability and even fell short of the non-spatial PCA-based method under limited availability. These findings suggest that the spatial proximity of SCs may not sufficiently capture distributed physiological noise sources, and that aggregating shared variance across all SCs provides a more robust correction. To our knowledge, the only study that directly compared spatial-specific and non-spatial-specific approaches in the sensorimotor cortex during a hand-grasping task reported comparable GLM performance between the two methods (Wyser et al., 2020). However, their non-spatial analysis relied on a single PCA component, which likely limited its effectiveness, as previous work has shown that including multiple PCA components is critical for capturing a broader range of systemic noise (Santosa et al., 2020; Luke et al., 2021).

When comparing limited vs. full SC availability, the non-spatial PCA-based approach recovered qualitatively similar activation patterns in the occipital cortex across both conditions, indicating that visual-related effects can be detected even with only three SCs. However, the temporal lobes, particularly the right hemisphere, showed a marked improvement with full SC coverage: more channels passed our statistical criterion, and

auditory-driven responses were stronger and more spatially extensive. This observation suggests that while limited SC availability can be sufficient to recover robust, widespread effects, full coverage becomes essential when targeting weaker or more spatially focal responses (here the auditory activations), in line with Klein et al. (2022) who showed that including all SCs in a non-spatial GLM yielded better signal quality than using just two SCs (one per hemisphere) during a motor task.

Taken together, our results indicate that non-spatial-specific PCA regression consistently outperforms spatial-specific approaches, both when SCs are limited and when coverage is full. While using multiple SCs is always preferable, even minimal SC configurations can be effective if shared variance is appropriately modeled. These findings support existing recommendations to incorporate multiple sources of systemic signal when applying SC correction (Santosa et al., 2020; Luke et al., 2021; Wyser et al., 2020), and suggest that global physiological noise is more effectively addressed by modeling shared variance across multiple SCs rather than relying solely on spatial proximity.

4.3. PCA in non-Spatial SC regression

To our knowledge, this study is the first to directly address whether applying PCA to SC signals improves the effectiveness of non-spatial-specific GLM-based denoising. While recent work has widely adopted PCA-transformed SC regressors to account for systemic noise (Luke et al., 2021; Zhou et al., 2020; Klein et al., 2022; Wyser et al., 2020; Sato et al., 2016; Santosa et al., 2020), no prior study has explicitly compared PCA-based and raw-signal approaches using the same dataset and analysis pipeline. Our results show that PCA-based regression outperforms the use of raw SC signals, in terms of the number of channels showing condition-specific effects and the statistical strength of those effects. The raw-signal approach did recover expected activity, especially in the occipital region, but in fewer channels and with reduced evidence strength in the temporal lobes. These results suggest that directly entering all SC signals into the GLM may introduce redundancy and multicollinearity, limiting sensitivity. Crucially, PCA offers an orthogonal representation of the shared systemic signal variance across SCs, enhancing the model's ability to isolate task-related neural signals. An alternative strategy has been proposed and implemented in the MNE-NIRS library (Luke et al., 2021), which involves including the mean HbO and HbR signal across all SC channels as regressors. This method has shown comparable performance to PCA-based regression in passive auditory tasks, offering an alternative implementation option.

Our findings support the use of PCA as the preferred method for non-spatial-specific SC regression when full SC coverage is available and robust denoising is desired, while also recognizing that alternatives such as the mean-signal approach may provide reasonable approximations in specific contexts.

4.4. Correction strategies for HbR signal

The HbR signal, analyzed using the same procedures as for HbO (except for directionality of Bayesian tests), showed consistent patterns of negative task-related responses across SC correction scenarios. Without SC correction, occipital channels showed the strongest effects, with six channels exhibiting emergence (metric 1), and up to five channels showing strong to decisive evidence for V > A or AV > A (metric 2).

With limited SC availability (n=3), both spatial- and non-spatial-specific approaches recovered similar emergence patterns, additionally uncovering between-condition differences for one channel over the left temporal area. Full SC availability (n=8) revealed similar emergence detection but better between-condition differences detection with up to five channels in the occipital cortex for V < A and AV < A comparisons and one channel in the left temporal area for both A < V and AV < V comparisons. Overall, as with HbO, the analysis pipeline that yielded the best results was the one including all principal components of all SCs in the case of full SC availability; however, the advantage is markedly less pronounced than for HbO, as the expected effects can still be retrieved even without SC regression. Our findings are consistent with a substantial body of research showing that HbR is inherently less affected by systemic physiological noise than HbO (Kirilina et al., 2012, 2013; Abdalmalak et al., 2022; Santosa et al., 2020; Tachtsidis & Scholkmann, 2016). HbR exhibits more linear and predictable response characteristics (Wobst et al., 2001), likely due to the greater sympathetic innervation of arterioles compared to venules (Tachtsidis & Scholkmann, 2016) and its lower sensitivity to corpuscular volume changes in the vasculature (Wobst et al., 2001). As a result, SC regression techniques tend to produce more pronounced improvements for HbO, while HbR typically shows less need for such correction (Santosa et al., 2020; Abdalmalak et al., 2022). Nonetheless, our results indicate that SC regression still provides a modest but consistent improvement for HbR, and that the pipeline yielding the best results for HbO likewise offers the most robust outcomes for HbR. These findings support the recommendation that, despite its lower sensitivity to systemic noise, HbR should be corrected using the same SC regression strategy as HbO to ensure optimal data quality and interpretability.

4.5. Recommendations and conclusion

To our knowledge, this is the first study to provide concrete recommendations for SC regression in fNIRS preprocessing using a naturalistic paradigm combining passive auditory, visual, and audiovisual stimulation within the same participants. This cross-modal, multi-condition design enhances the generalizability of our findings across sensory modalities and cortical regions.

Our results confirm that for HbO signal, omitting SC regression leads to completely uninterpretable GLM outcomes, with no detectable task-related activation, underscoring that SC correction is paramount for recovering meaningful neural signals. When SC channels are limited, pooling all available SC signals (rather than relying solely on spatially nearest channels) improves detection of expected activations. Furthermore, when full SC coverage is available, applying dimensionality summarization techniques such as PCA to the pooled SC signals further enhances the ability to isolate task-related activity by capturing shared systemic signal variance and reducing redundancy. The same SC regression strategy should also be applied to the HbR signal, which, although less affected by systemic interference, still benefits from such correction.

Taken together, these findings, alongside other studies, support the recommendation to include all available SC signals in the GLM and apply orthogonalization techniques such as PCA, providing a robust and effective approach for denoising fNIRS data across diverse experimental contexts and hardware setups.

Data and Code Availability

The anonymized preprocessed data and analytic code necessary to reproduce the analyses presented in this paper is publicly accessible at the following URL: https://github.com/yanndavlem/nirs_article.

Author Contributions

YL: Conceptualization, Methodology, Software, Investigation, Resources, Writing - Original Draft, Writing - Review & Editing, Visualization, Formal analysis. PB: Investigation, Writing - Review & Editing, Resources, Supervision, Conceptualization, Funding acquisition. KS: Software. OD: Resources, Funding acquisition. AC: Conceptualization, Supervision, Methodology, Resources, Software, Writing–review & editing, Funding acquisition, Formal analysis, Visualization. JG: Conceptualization, Supervision, Methodology, Software, Investigation, Resources, Writing - Original Draft, Writing - Review & Editing, Visualization, Formal analysis.

Funding

This work was supported by an ANRT (Association Nationale de la Recherche et de la Technologie) - Advanced Bionics CIFRE (Conventions industrielles de formation par la recherche) PhD funding (N°2022/0341) to YL, the AgeHear ANR funding (ANR-20-CE28-0016) to PB, the recurrent funding of the CNRS (Centre national de la recherche scientifique), France, and the DRCI of the CHU Hospital Toulouse. This work was also funded by a “Pack Ambition Recherche” (COGAUDYS project) from the Auvergne-Rhône-Alpes Region, awarded to AC.

Declaration of Competing Interests

The authors declare no conflict of interest.

Acknowledgement:

We thank all of our subjects for their participation in the study, and the ENT team at Purpan hospital, Toulouse, for their help. The authors would also like to acknowledge Laure Arnold and the AB team for their support.

References:

Abdalmalak, A., Novi, S. L., Kazazian, K., Norton, L., Benaglia, T., Slessarev, M., Debicki, D. B., Lawrence, K. St., Mesquita, R. C., & Owen, A. M. (2022). Effects of Systemic Physiology on Mapping Resting-State Networks Using Functional Near-Infrared Spectroscopy. *Frontiers in Neuroscience*, *16*, 803297. https://doi.org/10.3389/fnins.2022.803297

Amaro, E., & Barker, G. J. (2006). Study design in fMRI : Basic principles. *Brain and Cognition*, *60*(3), 220-232. https://doi.org/10.1016/j.bandc.2005.11.009

Anderson, C. A., Wiggins, I. M., Kitterick, P. T., & Hartley, D. E. H. (2017). Adaptive benefit of cross-modal plasticity following cochlear implantation in deaf adults. *Proceedings of the National Academy of Sciences*, *114*(38), 10256-10261. https://doi.org/10.1073/pnas.1704785114

Anderson, C. A., Wiggins, I. M., Kitterick, P. T., & Hartley, D. E. H. (2019). Pre-operative Brain Imaging Using Functional Near-Infrared Spectroscopy Helps Predict Cochlear Implant Outcome in Deaf Adults. *Journal of the Association for Research in Otolaryngology*, *20*(5), 511-528. https://doi.org/10.1007/s10162-019-00729-z

Bauernfeind, G., Wriessnegger, S. C., Haumann, S., & Lenarz, T. (2018). Cortical activation patterns to spatially presented pure tone stimuli with different intensities measured by functional near-infrared spectroscopy. *Human Brain Mapping*, *39*(7), 2710-2724. https://doi.org/10.1002/hbm.24034

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. https://doi.org/10.1117/1.NPh.2.2.025005

Chen, L.-C., Sandmann, P., Thorne, J. D., Herrmann, C. S., & Debener, S. (2015). Association of Concurrent fNIRS and EEG Signatures in Response to Auditory and Visual Stimuli. *Brain Topography*, *28*(5), 710-725. https://doi.org/10.1007/s10548-015-0424-8

Cui, X., Bray, S., & Reiss, A. L. (2010). Speeded Near Infrared Spectroscopy (NIRS) Response Detection. *PLOS ONE*, *5*(11), e15474. https://doi.org/10.1371/journal.pone.0015474

Fishburn, F. A., Ludlum, R. S., Vaidya, C. J., & Medvedev, A. V. (2019). Temporal Derivative Distribution Repair (TDDR): A motion correction method for fNIRS. *NeuroImage*, *184*, 171-179. https://doi.org/10.1016/j.neuroimage.2018.09.025

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, *105*(37), 14222-14227. https://doi.org/10.1073/pnas.0806530105

Ginzburg, J., Cheylus, A., Collard, E., Ferreri, L., Tillmann, B., Moulin, A., & Caclin, A. (2024). Role of the prefrontal cortex in musical and verbal short-term memory: A functional near-infrared spectroscopy study. *Imaging Neuroscience*, *2*, 1-23. https://doi.org/10.1162/imag_a_00168

Habermehl, C., Holtze, S., Steinbrink, J., Koch, S. P., Obrig, H., Mehnert, J., & Schmitz, C. H. (2012). Somatosensory activation of two fingers can be discriminated with ultrahigh-density diffuse optical tomography. *NeuroImage*, *59*(4), 3201-3211. https://doi.org/10.1016/j.neuroimage.2011.11.062

Highton, D., Boas, D. A., Minagawa, Y., Mesquita, R. C., & Gervain, J. (2023). Special Section Guest Editorial: Thirty Years of Functional Near-Infrared Spectroscopy. *Neurophotonics*, *10*(2), 023501. https://doi.org/10.1117/1.NPh.10.2.023501

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. https://doi.org/10.1364/AO.48.00D280

Jian-Wen, W. (2009). Tunneling conductance in quantum wire/insulator/d_{x2-y2} + id_{xy} mixed wave superconductor junctions. *Chinese Physics B*, *18*(10), 4479-4485. https://doi.org/10.1088/1674-1056/18/10/063

Kirilina, E., Jelzow, A., Heine, A., Niessing, M., Wabnitz, H., Brühl, R., Ittermann, B., Jacobs, A. M., & Tachtsidis, I. (2012). The physiological origin of task-evoked

systemic artefacts in functional near infrared spectroscopy. *NeuroImage*, *61*(1), 70-81. https://doi.org/10.1016/j.neuroimage.2012.02.074

Kirlilna, 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*. https://doi.org/10.3389/fnhum.2013.00864

Klein, F., Debener, S., Witt, K., & Kranczioch, C. (2022). fMRI-based validation of continuous-wave fNIRS of supplementary motor area activation during motor execution and motor imagery. *Scientific Reports*, *12*(1), 3570. https://doi.org/10.1038/s41598-022-06519-7

Kohno, S., Miyai, I., Seiyama, A., Oda, I., Ishikawa, A., Tsuneishi, S., Amita, T., & Shimizu, K. (2007). Removal of the skin blood flow artifact in functional near-infrared spectroscopic imaging data through independent component analysis. *Journal of Biomedical Optics*, *12*(6), 062111. https://doi.org/10.1117/1.2814249

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling : A Practical Course*. Cambridge University Press.

Luke, R., Larson, E., Shader, M. J., Innes-Brown, H., Van Yper, L., Lee, A. K. C., Sowman, P. F., & McAlpine, D. (2021). Analysis methods for measuring passive auditory fNIRS responses generated by a block-design paradigm. *Neurophotonics*, *8*(02). https://doi.org/10.1117/1.NPh.8.2.025008

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*(01). https://doi.org/10.1117/1.NPh.9.1.015001

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. https://doi.org/10.1111/jpr.12206

Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., & Burgess, P. W. (2020). The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience. *Annals of the New York Academy of Sciences*, *1464*(1), 5-29. https://doi.org/10.1111/nyas.13948

Santosa, H., Zhai, X., Fishburn, F., & Huppert, T. (2018). The NIRS Brain AnalyzIR Toolbox. *Algorithms*, *11*(5), Article 5. https://doi.org/10.3390/a11050073

Santosa, H., Zhai, X., Fishburn, F., Sparto, P. J., & Huppert, T. J. (2020). Quantitative comparison of correction techniques for removing systemic physiological signal in functional near-infrared spectroscopy studies. *Neurophotonics*, *7*(03). https://doi.org/10.1117/1.NPh.7.3.035009

Sato, T., Nambu, I., Takeda, K., Aihara, T., Yamashita, O., Isogaya, Y., Inoue, Y., Otaka, Y., Wada, Y., Kawato, M., Sato, M., & Osu, R. (2016). Reduction of global interference of scalp-hemodynamics in functional near-infrared spectroscopy using short distance probes. *NeuroImage*, *141*, 120-132. https://doi.org/10.1016/j.neuroimage.2016.06.054

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic Stimuli in Neuroscience : Critically Acclaimed. *Trends in Cognitive Sciences*, *23*(8), 699-714. https://doi.org/10.1016/j.tics.2019.05.004

Strotzer, M. (2009). One Century of Brain Mapping Using Brodmann Areas*. *Clinical Neuroradiology*, *19*(3), 179-186. https://doi.org/10.1007/s00062-009-9002-3

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. https://doi.org/10.1117/1.NPh.3.3.031405

van de Rijt, L. P. H., van Wanrooij, M. M., Snik, Ad. F. M., Mylanus, E. A. M., van Opstal, A. J., & Roye, A. (2018). Measuring Cortical Activity During Auditory Processing with Functional Near-Infrared Spectroscopy. *Journal of hearing science*, *8*(4), 9-18. https://doi.org/10.17430/1003278

von Eye, A., & Wiedermann, W. (Éds.). (2023). OLS Parameter Estimation. In *The General Linear Model : A Primer* (p. 5-9). Cambridge University Press. https://doi.org/10.1017/9781009322164.004

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I : Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35-57. https://doi.org/10.3758/s13423-017-1343-3

Weder, S., Zhou, X., Shoushtarian, M., Innes-Brown, H., & McKay, C. (2018). Cortical Processing Related to Intensity of a Modulated Noise Stimulus—A Functional Near-Infrared Study. *Journal of the Association for Research in Otolaryngology*, *19*(3), 273-286. https://doi.org/10.1007/s10162-018-0661-0

Wobst, P., Wenzel, R., Kohl, M., Obrig, H., & Villringer, A. (2001). Linear Aspects of Changes in Deoxygenated Hemoglobin Concentration and Cytochrome Oxidase Oxidation during Brain Activation. *NeuroImage*, *13*(3), 520-530. https://doi.org/10.1006/nimg.2000.0706

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. https://doi.org/10.1117/1.NPh.7.3.035011

Yamada, T., Umeyama, S., & Matsuda, K. (2012). Separation of fNIRS Signals into Functional and Systemic Components Based on Differences in Hemodynamic Modalities. *PLOS ONE*, *7*(11), e50271. https://doi.org/10.1371/journal.pone.0050271

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. https://doi.org/10.1364/BOE.7.003078

Yücel, M., Luke, R., Mesquita, R., von Lühmann, A., Mehler, D., Lührs, M., Gemignani, J., Abdalmalak, A., Albrecht, F., Almeida, I., Artemenko, C., Ashton, K., Augustynowicz, P., Bajracharya, A., Bannier, E., Barth, B., Bayet, L., Behrendt, J., Khani, H., … Zemanek, V. (2024). *The fNIRS Reproducibility Study Hub (FRESH): Exploring Variability and Enhancing Transparency in fNIRS Neuroimaging Research*. OSF. https://doi.org/10.31222/osf.io/pc6x8_v1

Zhang, Q., Strangman, G. E., & Ganis, G. (2009). Adaptive filtering to reduce global interference in non-invasive NIRS measures of brain activation: How well and when does it work? *NeuroImage*, *45*(3), 788-794. https://doi.org/10.1016/j.neuroimage.2008.12.048

Zhou, X., Sobczak, G., McKay, C. M., & Litovsky, R. Y. (2020). Comparing fNIRS signal qualities between approaches with and without short channels. *PLOS ONE*, *15*(12), e0244186. https://doi.org/10.1371/journal.pone.0244186

Zimeo Morais, G. A., Balardin, J. B., & Sato, J. R. (2018). fNIRS Optodes' Location Decider (fOLD): A toolbox for probe arrangement guided by brain regions-of-interest. *Scientific Reports*, *8*(1), 3341. https://doi.org/10.1038/s41598-018-21716-z