

Supplemental Material

Research Articles | Development/Plasticity/Repair

Maturation of neuronal activity in the human cortex exhibits robust spatial gradients across the birth transition

<https://doi.org/10.1523/JNEUROSCI.0610-25.2025>

Received: 26 March 2025

Revised: 4 June 2025

Accepted: 4 July 2025

Copyright © 2025 the authors

1 Data preparation, artefact removal and rejection

2 All data were first filtered with 3 filters (a 4th order Butterworth notch filter to remove 60Hz noise, a 4th
3 order high pass filter with a cutoff frequency of 0.5Hz, and a 4th order Butterworth low pass filter with a
4 cutoff frequency of 32Hz) and then re-sampled to 256Hz. All recording were initially pre-screened for
5 data size and EEG quality. Any EEG recording less than 200MB in size was excluded (data duration was
6 too short). Then, a series of EEG features were calculated on each EEG recording channel ($n = 125$), if
7 these features exceeded conservative thresholds of EEG quality in more than 20% of EEG recording
8 channels ($n = 25$) then a recording was excluded. EEG features include a robust estimate of the spectral
9 slope (the so-called aperiodic component of the spectrum between 1 and 30Hz), measures of relative
10 spectral power (low frequency power between 1 and 4 Hz and high frequency power between 8 and
11 32Hz), measures of the EEG amplitude (methods of feature extraction can be found on Github -
12 https://github.com/nstevensonUH/spatial_maturation_preterm_brain). The channel thresholds were
13 defined as

- 14 • Recording channels with a spectral slope less than or equal to -3 or greater than or equal to -1;
15 or with greater than 10% of the spectrum rejected as outliers during robust estimation (implying
16 a periodic component in the EEG) were excluded; or
- 17 • Recording channels with relative low frequency power greater than 40% or less than 20%, were
18 excluded; or
- 19 • Recordings with relatively high frequency power greater than 40% or less than 10% were
20 excluded
- 21 • EEG amplitude greater than 200 μ V for 25% of the recording duration were excluded

22 After passing the initial screening process, EEG recordings were then subject to an independent
23 component analysis (ICA) based artefact removal procedure. A continuous section of EEG recording (at
24 least 10 minutes in duration) that was minimally contaminated by large amplitude artefacts ($> 200 \mu$ V)
25 was extracted. The multi-channel signal segment was then decomposed into independent components
26 and the same set of qEEG features as calculated above (except for EEG amplitude which is used at a later
27 stage) from each independent component ($n = 124$), an outlier detection threshold was then applied to
28 each feature. The outlier detection threshold was a conservative 5 scaled median absolute deviation
29 (MAD) and was applied to the absolute value of the median centered feature. Outlier components were
30 removed and the signal was reconstructed (if more than 10 components were found then only the top
31 10 components were removed). High amplitude (greater than 200 μ V) artefacts were detected but not
32 removed; signal segments containing high amplitude artefacts were excluded from feature calculation.
33 Channels that were initially excluded were then interpolated by taking the local mean of 4 surrounding
34 channels. This signal was then passed to the source reconstruction algorithm.

35 During pre-processing, 297 out of 331 EEG recordings passed the first stage (artefact detection), and an
36 additional 6 files did not make it through the second stage (artefact removal) resulting in 291 files for
37 further processing. Of these 291 EEG recordings, the number of contaminated channels had a median of
38 3 (IQR: 2 to 5; $n = 291$) and the number of ICA components removed had a median of 2 (IQR: 1 to 4; $n =$

291). A further 18 recordings had outliers in at least two feature values within at least one parcel signal resulting in 277 recordings from 136 infants for analysis.

Training

1. Organize data into feature vector and age variable.
2. Exclude outliers from individual recording per channel (adjust feature for linear relationship with age, apply *rmoutliers* Matlab function with default setting to residual and remove outliers).
3. Augment data with synthetic data to homogenize the distribution of data with respect to age.
4. Train model (MDL) – support vector regression with hyperparameters as used in methods
 - a. Use 10-fold cross-validation to estimate out-of-sample performance on all data
5. Return MDL and apply to testing data.

Testing

1. Organize data into feature vector and age variable.
2. Exclude outliers from individual recording per channel (adjust feature for linear relationship with age, apply *rmoutliers* Matlab function with default setting to residual and remove outliers)
3. Apply MDL to feature vector to estimate FBA

The use of outlier removal based on feature value to exclude regions per recording during both training and testing stage, and the use of paired measurements to estimate Δ FBA between training and testing regions, means that not all of the 277 recordings were used for each combination of training and testing regions. Outlier removal resulted in a median number of 221 recordings (IQR: 217 to 227; minimum 201, maximum 232) for training the region-specific FBA and 218 recordings (IQR: 214 to 225; minimum 196; maximum 231) when paired with recordings from testing regions. In total only, 267 recordings from 135 infants were used at least once throughout the 1936 possible combinations.

Statistical Analysis: technical confounders

We tested to see if the process of source reconstruction introduced spatial gradients into the output sources. To do this, we used surrogate EEG recordings constructed by randomizing the original location of each EEG channel removing the spatial structure of EEG recording. Source reconstructed cortical signals were then extracted from this null EEG recording and our analysis pipeline applied to this data. If no gradients were detected, i.e. no significant correlation between axial distance and age, then we assume that gradients found with no reordering are legitimate. We repeated this process 10 times using different random EEG channel assignments, noting that it is possible for random channel re-ordering to align with similar axial distances.

Figure S1: Parcels that were excluded from analysis as being too far from the scalp (B). Several of these parcels correspond with parcels of high age prediction error (denoted with red circles in (A)).

Figure S2: Calculating measures of bursting from the cortical activity data. Top panel: EEG data is converted into amplitude using the Hilbert transform, this is then segmented into bursts of different duration (denoted by different colours in the barcode) using a simple thresholding procedure. The amplitude threshold is selected to maximize the number of bursts. Note, the amplitude is shown here for clarity, but it is the amplitude squared that is used. Middle left panel: Individual bursts are then extracted from the amplitude and ordered according to duration for the exemplar EEG segment. Middle right panel: bursts within a duration bin (in this case 0.125 to 0.25s) are then normalised in time and amplitude and averaged together to form a single burst template per duration. Symmetry and sharpness indices are calculated on the average burst amplitude and are analogous to skewness and kurtosis, respectively, assuming the average burst amplitude is a probability density function. Bottom left panel; a line of best fit to symmetry vs burst duration, calculated across the entire hour-long recording, results in the extraction of a slope and intercept value. Bottom right panel; a line of best fit to sharpness vs burst duration results in the extraction of a slope and intercept value – see Roberts et al. 2014 for more details.

Figure S3: Areal ranking of the sensorimotor-association axis based on Sydnor et al. 2021 projected onto our parcellation.

24 **Figure S4:** Changes in gradients along a frontal-occipital axis when calculating Δ FBA
25 using 1) all available data, 2) recordings less than 36 weeks PMA, or 3) recordings
26 greater than 36 weeks PMA.

27

28

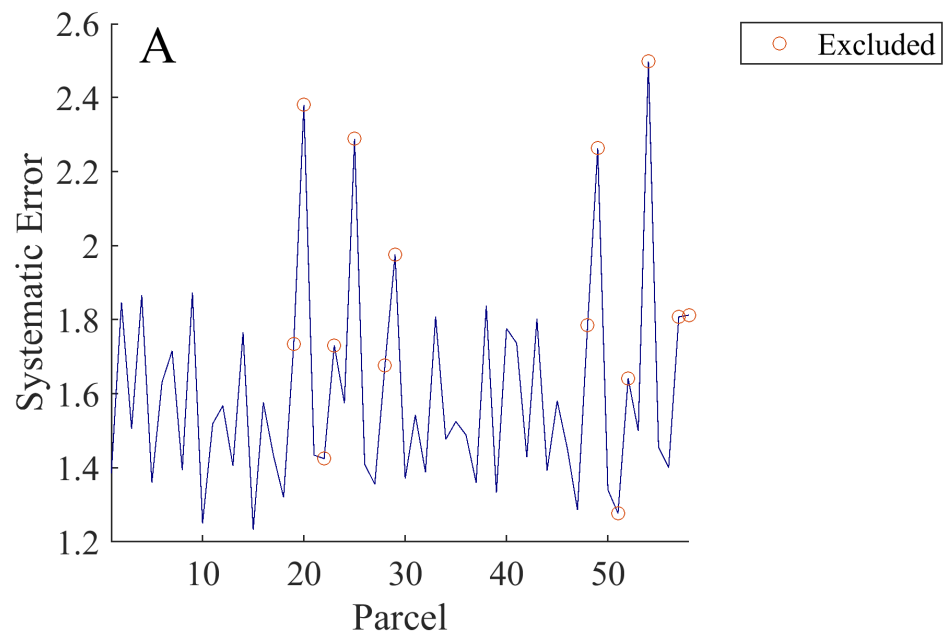
Table Legends

Table 1: The subset of EEG features used to predict age per parcel signal. Superscripts denote references for each feature and the signal domain/class of each feature.

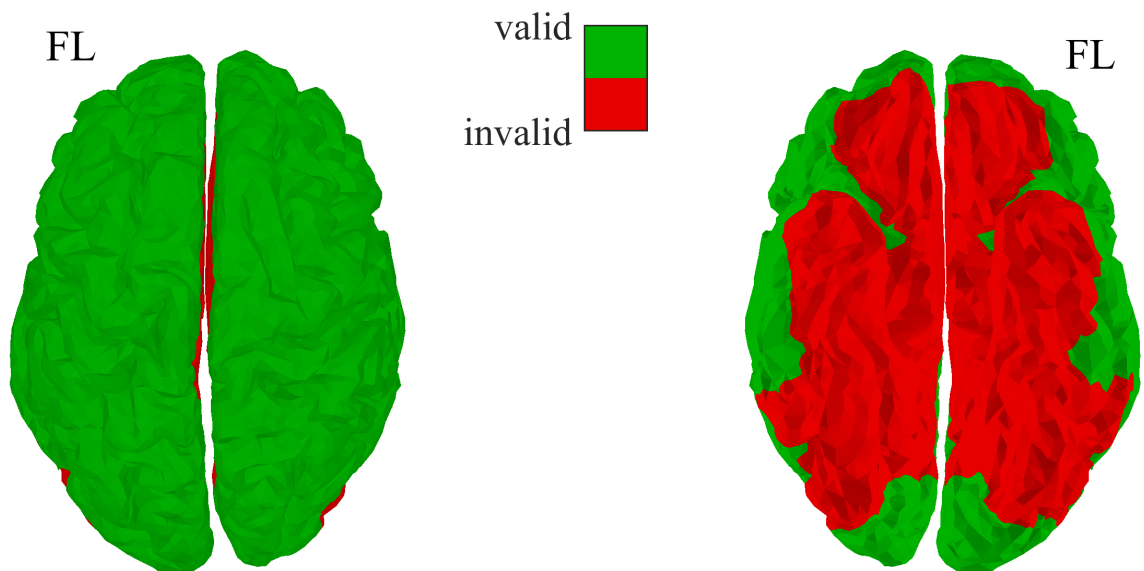
Table 2: Gradients estimated using Δ FBA calculated on all available data, only recordings from less than 36 weeks PMA, and only recordings after 36 weeks PMA. The values shown are the correlation coefficient (95%CI) between Δ FBA and axial distance for all axes tested and * denotes values with a 95%CI that does not span 0 and are, therefore, assumed to be significant.

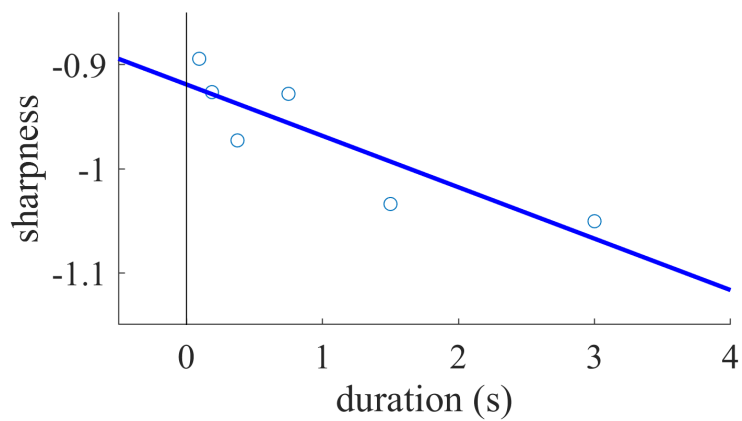
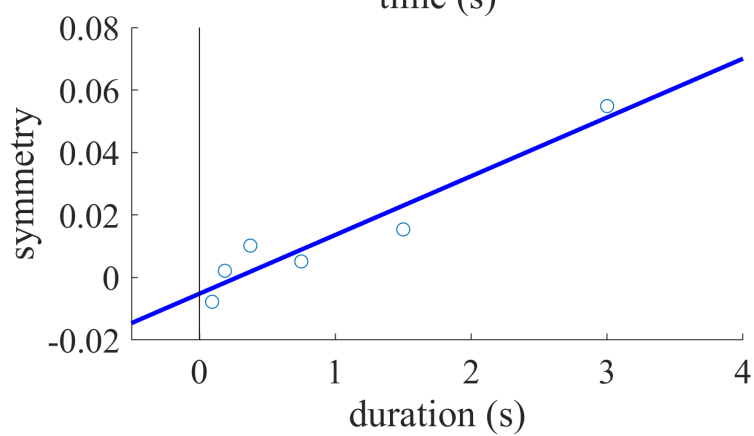
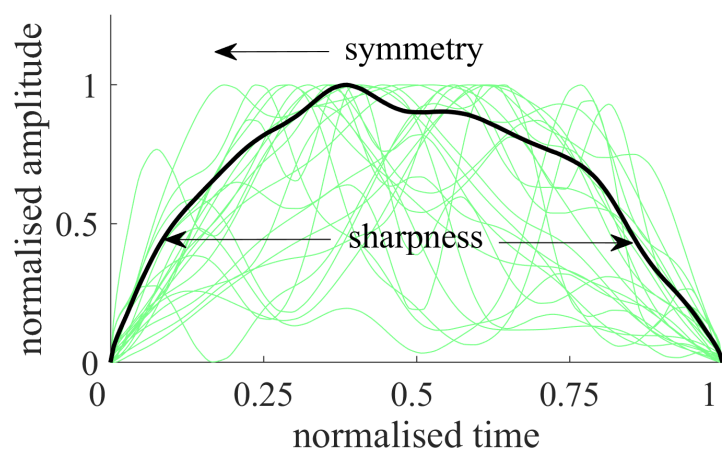
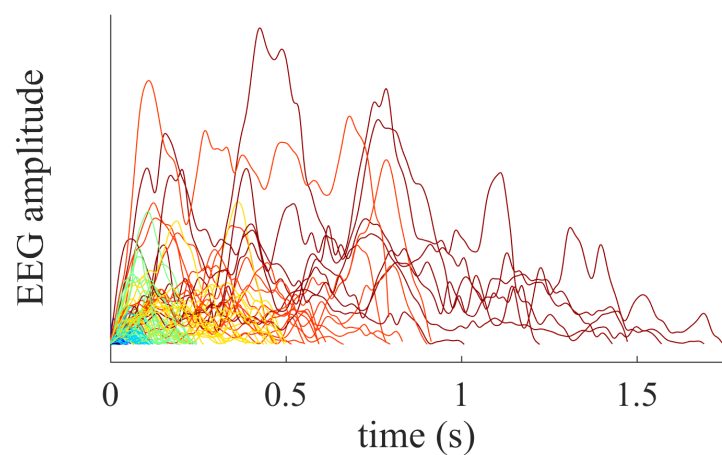
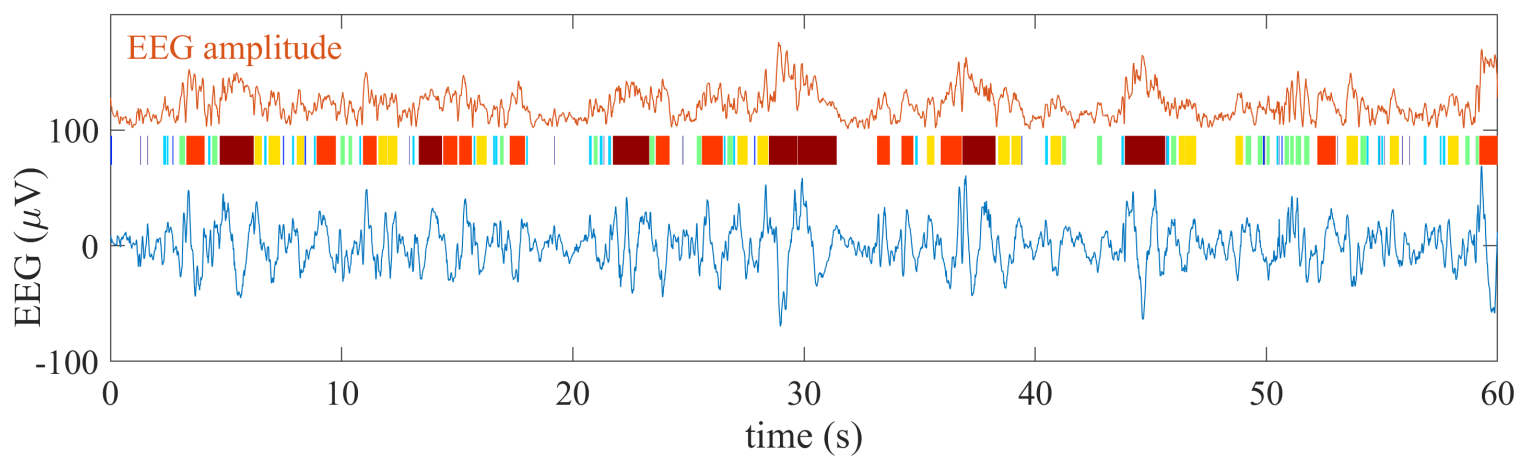
Table S1: Changes in correlations between FBA and spatial axis (FO, lateral, and S) due to confounders. Infants without any demographic information are also excluded. n refers to the number of infants excluded in addition to the 66 infants without demographic information. Significant correlations suggest negligible confounding and are denoted with *.

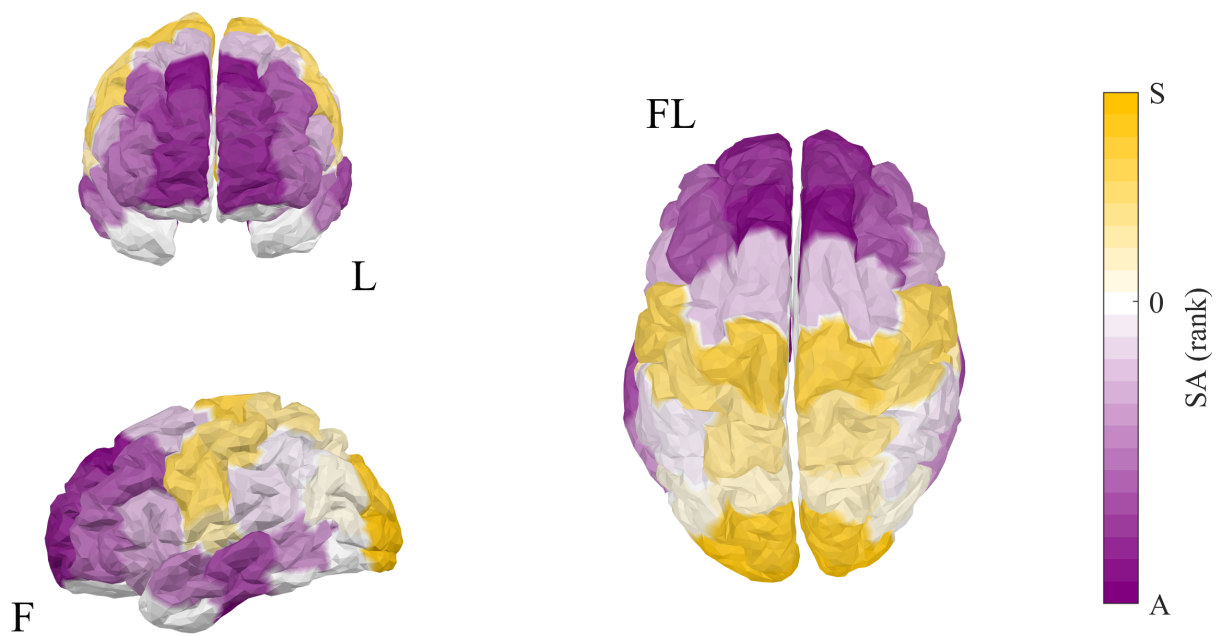
Table S2: Correlations between FBA and spatial axis using a null model of random channel assignment in the sensor space. Correlations that are not significant suggest negligible effects of the source space reconstruction process. Significant correlations found along an axis are denoted with *, r is the linear correlation coefficient between predicted age difference and rank/distance for the frontal-occipital (FO), lateral and sensorimotor-association (SA) axis, respectively.

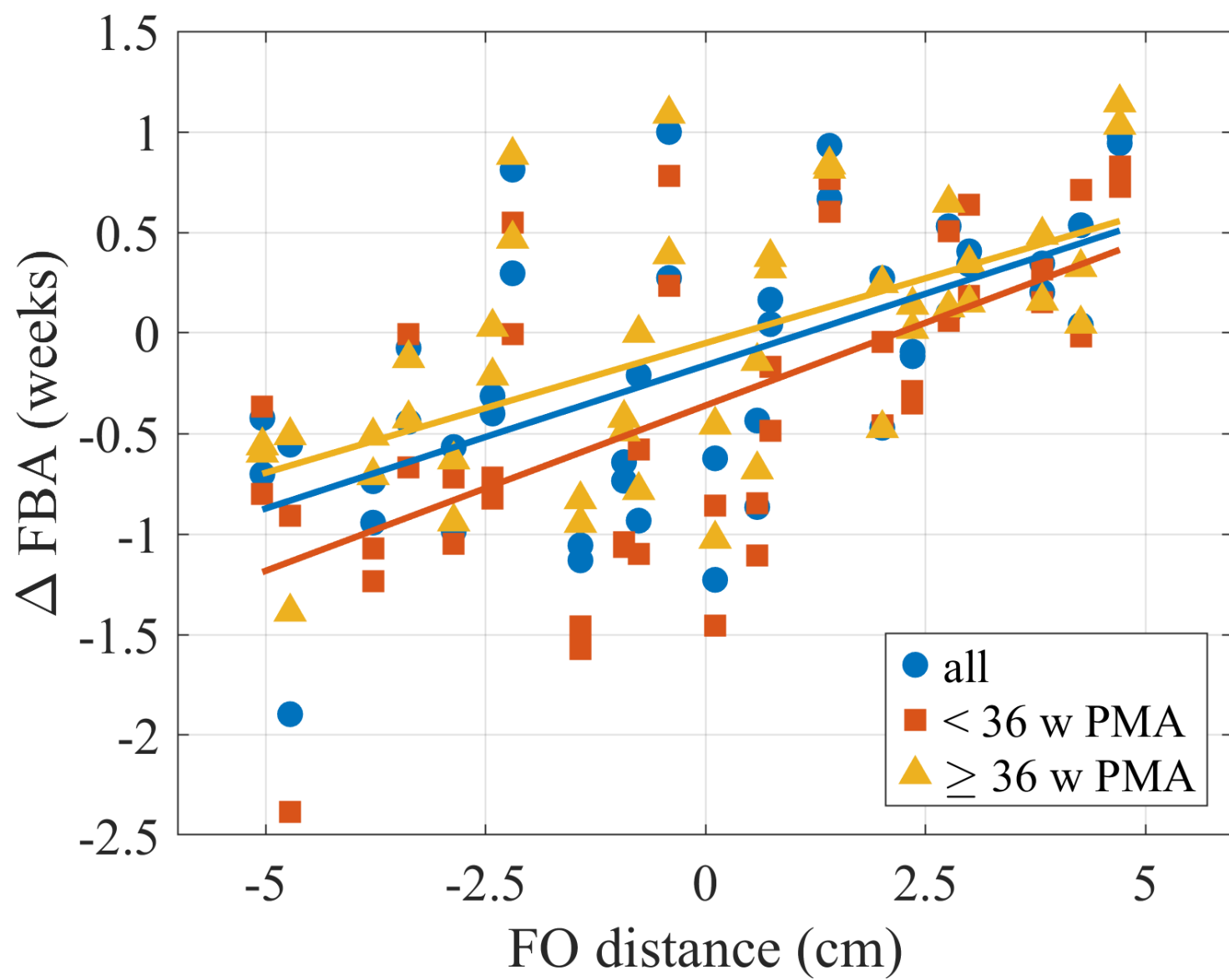


B









1 **Table S1:** Changes in correlations between FBA and spatial axis (FO, lateral, and S) due to confounders.
2 Infants without any demographic information are also excluded. n refers to the number of infants
3 excluded in addition to the 66 infants without demographic information. Significant correlations suggest
4 negligible confounding and are denoted with *.

Confounder	Exclude	Δ FBA vs FO, r [95% CI]	Δ FBA vs lateral distance, r [95% CI]	Δ FBA vs SA rank, r [95%CI]
Biological Sex	Male (n=36)	0.548 [0.299 to 0.726]*	-0.428 [-0.643 to -0.150]*	0.317 [0.022 to 0.561]*
	Female (n=35)	0.565 [0.322 to 0.738]*	-0.426 [-0.642 to -0.148]*	0.340 [0.048 to 0.578]*
Outcome	Normal (n=33)	0.529 [0.276 to 0.714]*	-0.429 [-0.644 to -0.151]*	0.308 [0.013 to 0.555]*
	Abnormal (n=38)	0.567 [0.324 to 0.739]*	-0.436 [-0.649 to -0.160]*	0.339 [0.047 to 0.578]*
Treatment	Treatment (n=44)	0.551 [0.304 to 0.729]*	-0.405 [-0.626 to -0.123]*	0.314 [0.019 to 0.559]*
	No Treatment (n=27)	0.549 [0.301 to 0.727]*	-0.463 [-0.668 to -0.192]*	0.323 [0.029 to 0.566]*

5

6

Table S2: Correlations between FBA and spatial axis using a null model of random channel assignment in the sensor space. Correlations that are not significant suggest negligible effects of the source space reconstruction process. Significant correlations found along an axis are denoted with *, r is the linear correlation coefficient between predicted age difference and rank/distance for the frontal-occipital (FO), lateral and sensorimotor-association (SA) axis, respectively.

Random Seed	Δ FBA vs FO distance, r [95%CI]	Δ FBA vs lateral distance r [95% CI]	Δ FBA vs SA Rank r [95%CI]
1	0.052 [-0.249 to 0.343]	0.267 [-0.032 to 0.523]	0.158 [-0.146 to 0.435]
2	0.054 [-0.247 to 0.345]	0.145 [-0.158 to 0.424]	0.233 [-0.068 to 0.496]
3	-0.026 [-0.321 to 0.273]	0.019 [-0.279 to 0.314]	0.148 [-0.156 to 0.426]
4	-0.030 [-0.324 to 0.269]	0.218 [-0.084 to 0.484]	0.014 [-0.284 to 0.310]
5	-0.057 [-0.348 to 0.244]	0.080 [-0.222 to 0.368]	0.210 [-0.092 to 0.477]
6	0.180 [-0.124 to 0.452]	0.094 [-0.208 to 0.381]	0.407 [0.126 to 0.628]*
7	0.234 [-0.067 to 0.497]	0.028 [-0.271 to 0.323]	0.295 [-0.002 to 0.544]
8	-0.138 [-0.418 to 0.165]	0.321 [0.026 to 0.564]*	0.068 [-0.233 to 0.358]
9	0.144 [-0.159 to 0.423]	0.194 [-0.109 to 0.464]	0.234 [-0.067 to 0.497]
10	-0.164 [-0.439 to 0.140]	0.263 [-0.037 to 0.519]	0.113 [-0.190 to 0.397]