Supplemental Information

eMethods

Source Time Course Calculation

Sources were localised for each subject using Cortical-Start Spatio-Temporal (CSST) multidipole analysis with integrated Multiple Signal Classification (MUSIC)[1]. CSST is an objective multidipole, multistart procedure in which initial dipole locations are randomly selected from a predefined cortical volume and a nonlinear simplex search is performed for each of these initial configurations. Initial dipole locations were selected from within a predefined head volume, which was defined by a subsample of points taken from within the cortical volume, as determined by co-registered structural MRI. The error is minimised using a reduced chi-square statistic to obtain a final set of dipole configurations, which most fully explain the data. CSST source localisation was calculated using 3, 4, 5, 6, and 7 dipole models, based on the averaged responses occurring between 0 ms and 320 ms after the onset of the stimulus for responses to somatosensory stimuli. A shorter time window after stimulus onset was used for dipole modelling of somatosensory responses to increase the power to detect somatosensory evoked sources. The Nelder-Meade minimisation procedure was carried out 1500 to 8000 times, depending on the number of dipoles in the model, to help to ensure that the procedure would reach a global minimum. The dipole model that best explained the data was selected for source time course analysis.

Following the selection of the optimal source model, the single-trial waveforms of each dipolar source were calculated within a realistic head model with the minimum norm estimate (MNE) software[2]. The inverse solution yielded estimates of continuous time series of cortical currents. For each patient, the realistic cortical surface and three layers (inner skull, outer skull and skin) were reconstructed from the anatomical MRI images using the Freesurfer software (Compumedics, Charlotte, NC). The boundary element model (BEM) was then constructed with the reconstructed surfaces. The co-registration of MEG and MRI images was achieved by matching the recorded positions of three fiducial points (nasion, left and right preauricular points) with the locations of these points from the MRI images. The lead field matrix relating MEG sensors to the cortical distributed dipoles was computed with the BEM model using MNE. The dipole model, cortical surface and lead field matrix were then used in the MNE software to extract the single-trial time courses of sources. MNE was applied before time-frequency decomposition here.

An evaluation was carried out by calculating the residual variance (RV) of the signal, i.e., the percentage of data that cannot be

explained by the fitted dipoles. Within the selected time interval from 0 to 320 ms, the data-based model explains the scalp distribution of the somatosensory evoked MEG for the HC subjects with a RV of 5.5% (10.8±4.2%) at group (single-subject) level, respectively. For the PAE group, the group level RV is 4.8% and the single-subject level RV is 9.6 ±3.2%.

Renormalized Partial Directed Coherence Combined with State Space Modeling

The time-dependent Granger causal connectivity analysis method was employed in this study[3,4]. A time continuous multivariate dynamical process $Z(t)$ can only be observed as a multivariate time discrete sampled signal[3]

$$Y(t_i) = g(Z(t_i), v) + \eta(t_i), \quad (1)$$

where $g(\cdot)$ denotes the observation function with parameter set v; $\eta(t)$ is a Gaussian distributed independent measurement noise with a given variance. Assuming a linear observation function, we obtain the following model:

$$Z(i) = A(i)Z(i\text{-}1) + \varepsilon(i), \varepsilon(i) \sim N(0, \tilde{\Sigma}), \quad (2)$$

$$Y(i) = CZ(i) + \rho(i), \rho(i) \sim N(0, \tilde{\Gamma}), \quad (3)$$

for some appropriately chosen variances $\tilde{\Sigma}$ and $\tilde{\Gamma}$ that are optimally determined in the estimation process and where C represents the linear observation matrix. A reasonable assumption is that the parameter matrix $A(i)$ should change more slowly than the (stochastic) dynamics itself. The model is then augmented to the over-arching state space model as follows:

$$a(i) = a(i-1) + \varsigma(i), \varsigma(i) \sim N(0, \tilde{\Omega}), \quad (4)$$

$$Z(i) = A(i)Z(i\text{-}1) + \varepsilon(i), \varepsilon(i) \sim N(0, \tilde{\Sigma}), \quad (5)$$

$$Y(i) = CZ(i) + \rho(i), \rho(i) \sim N(0, \tilde{\Gamma}), \quad (6)$$

The $a(i)$ are the matrix entries of $A(i)$ rearranged into a vector. The causal influences can be represented as directed edges in a network, in which the nodes represent the processes. Thus the matrix $A(i)$ contains the interactions between the components of the

original process $Z(i)$ and thus the information about the network structure. Since we do not make any assumption about the origin of $Z(i)$, it can model the sensor space as well as the source space equally well.

In networks, influences with a certain delay are typically relevant. This can be accounted for in state space modelling by including previous time steps,

$$Z(i) = \left[\sum_{i'=1}^{p} A_{i'}(i) Z(i-i')\right] + \xi(i), \xi(i) \sim N(0, \tilde{\Sigma}), \quad (7)$$

up to a maximum time lag p. This maximum p can be determined relying on *a priori* knowledge or based on model selection criteria, such as Akaike's Information Criterion (AIC) used in our study[3]. The higher order process can be rewritten as a first order process by introducing

$$\hat{Z}(i) = \left(Z(i), Z(i-1), \ldots, Z(i-p+1)\right)', \quad (8)$$

$$\hat{Z}(i) = \hat{A}(i)\hat{Z}(i-1) + \hat{\xi}(i), \hat{\xi}(i) \sim N(0, \tilde{\Sigma})' \quad (9)$$

The matrix $\hat{A}(i)$ assumes the specific structure

$$\hat{A}(i) = \begin{pmatrix} A_1(i) & A_2(i) & \ldots & A_p(i) \\ 1_n & 0_n & \ldots & 0_n \\ 0_n & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \quad (10)$$

In the above form, the trajectory of the $Z(i)$ and the $a(i)$ can be determined purely based on the observations $Y(i)$. The expectation maximisation (EM) algorithm is an iterative algorithm that converges in the sense of maximum likelihood to the best estimator of the underlying dynamical process $Z(i)$ and the parameters $a(i)$. In the expectation step of the EM algorithm, the dual Kalman filter is used[3,4].

rPDC is derived as follows[3,4]:

$$\lambda_{u \leftarrow v}(\omega) = X_{uv}(\omega)'(V_{uv}(\omega))^{-1} X_{uv}(\omega), \quad (11)$$

87 This is a frequency domain measure for Granger-causality[3,4] that quantifies the direction and the strengths of network connections.

88 $X_{uv}(\omega) = \left[R(FT(A_{i',uv})), I(FT(A_{i',uv})) \right]'$ with R(.) the real and I(.) the imaginary parts, and FT(.) the Fourier transform. The

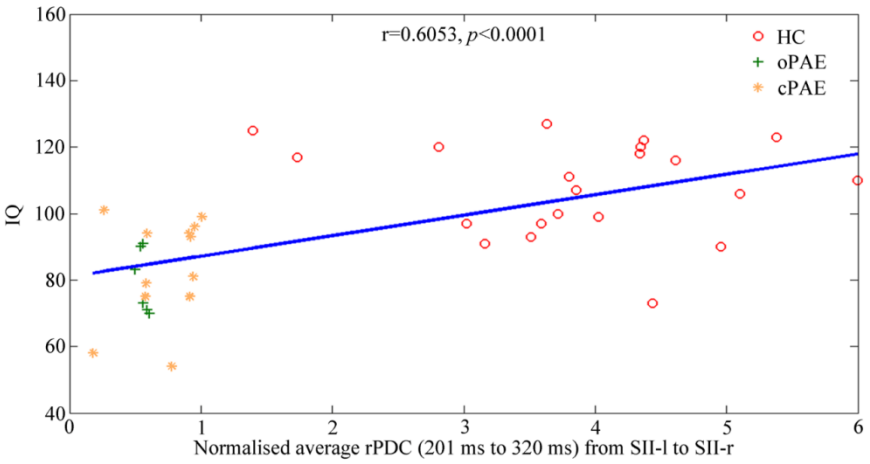
89 normalization by $\left(V_{uv}(\omega)\right)^{-1}$ is given by the inverse of

90 $$V_{uv}(\omega) = \sum_{i_1',i_2'=1}^{p} \mathrm{cov}(A_{i_1',uv}, A_{i_2',uv}) \cdot \begin{pmatrix} \cos\omega i_1' \cos\omega i_2' & \cos\omega i_1' \sin\omega i_2' \\ \sin\omega i_1' \cos\omega i_2' & \sin\omega i_1' \sin\omega i_2' \end{pmatrix}. \quad (12)$$

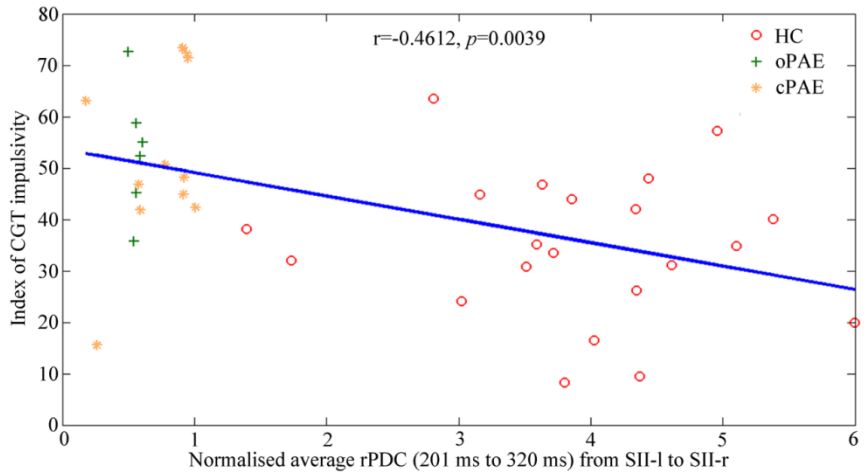
91 The covariance matrix of the estimated parameters $\mathrm{cov}(A_{i_1',uv}, A_{i_2',uv})$ is determined in the dual Kalman filter. The mathematical

92 details were demonstrated in our previous work[3,4].

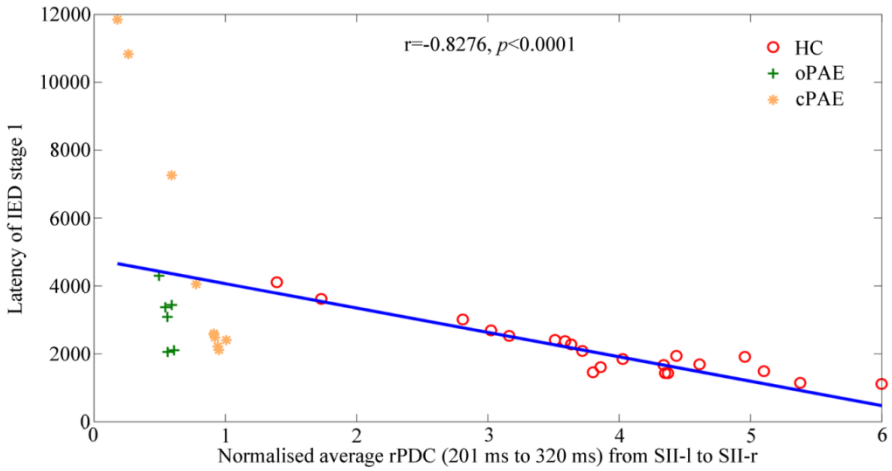
93



A



B



C

Figure s1: Linear regression analysis of normalized average rPDC values (201 to 320 ms) from SII-l to SII-r versus IQ (A), CGT impulsivity index (B), and latency of intra-extra dimensional IED shift task latency (C) respectively for the HC subjects, oPAE and cPAE subjects.

99 **Table s1. The dual state space model for the coefficients and process vectors of a time-dependent**

100 **VAR[p] process contaminated with observational noise**

Dual state space model	
Process state space	**Parameter state space**
$\vec{u}(t) = A(t)\vec{u}(t-1) + \vec{\varepsilon}_u(t)$	$\vec{a}(t) = \vec{a}(t-1) + \vec{\varepsilon}_a(t)$
$\vec{y}(t) = C_u\vec{u}(t) + \vec{\eta}(t)$	$\vec{y}(t) = C_a(t)\vec{a}(t) + \vec{\varepsilon}_x(t) + \vec{\eta}(t)$
Process transition matrix	**Parameter observation matrix**
$A^{\vec{a}(t-1)} = \begin{pmatrix} A_0(t) \\ A_1 \end{pmatrix}$	$C^{\vec{u}(t-1)} = \begin{pmatrix} \vec{u}^T(t-1) & \cdots & \vec{0}_{np}^T \\ \vdots & \ddots & \vdots \\ \vec{0}_{np}^T & \cdots & \vec{u}^T(t-1) \end{pmatrix}$
$A(t) := A^{\vec{a}(t-1)}$	$C_a(t) := C^{\vec{u}(t-1)}$
Expectation step	
Improved dual Kalman filter	
Process Kalman filter for t=1,…, N	Parameter Kalman filter for t=1,…, N
Process transition matrix	Parameter transition matrix
$A(t\|t-1) = \begin{pmatrix} A_0(t\|t-1) \\ A_1 \end{pmatrix}$	$C_a(t\|t-1) := C^{\vec{u}(t-1\|t-1)}$
Process prediction	**Parameter prediction**
$\vec{u}(t\|t-1) = A(t\|t-1)\vec{u}(t-1\|t-1)$	$\vec{a}(t\|t-1) = \vec{a}(t-1\|t-1)$
$\boldsymbol{P^u(t\|t-1) = A(t\|t-1)P^u(t-1\|t-1) + Q_u}$	$P^a(t\|t-1) = P^a(t-1\|t-1) + Q_a$
$P^x(t\|t-1) = P^x(t\|t-1) + C_a(t\|t-1)P^a(t-1\|t-1)C_a^T(t\|t-1)$	$P^{y_a}(t\|t-1) = P^{y_u}(t\|t-1) + C_a(t\|t-1)Q_aC_a^T(t\|t-1)$
$P^{y_u}(t\|t-1) = C_uP^u(t\|t-1)C_u^T + R$	
Process update	**Parameter update**
$K_u(t) = P^u(t\|t-1)C_u^T(P^{y_u}(t\|t-1))^{-1}$	$K_a(t) = P^a(t\|t-1)C_a^T(t\|t-1)C_x^T(P^{y_a}(t\|t-1))^{-1}$
$\vec{u}(t\|t) = \vec{u}(t-1\|t-1) + K_u(t)(\vec{y}(t) - C_u\vec{u}(t\|t-1))$	$\vec{a}(t\|t) = \vec{a}(t\|t-1) + K_a(t)(\vec{y}(t) - C_a(t)\vec{a}(t\|t-1))$

Dual state space model

$$P^u(t|t) = (I_{np} - K_u(t)C_u)P^u(t|t-1)$$

$$P^a(t|t) = (I_{n^2p} - K_a(t)C_a(t))P^a(t|t-1)$$

Dual smoothing filter

Process smoothing filter for $t=N,\ldots,1$

$$B_u(t-1) = P^u(t-1|t-1)A^T(t|t-1)(P^u(t|t-1))^{-1}$$

$$\vec{u}(t-1|N) = \vec{u}(t-1|t-1) + B_u(t-1)(\vec{u}(t|N) - \vec{u}(t|t-1))$$

$$P^u(t-1|N) = P^u(t-1|t-1) + B_u(t-1).\ .(P^u(t|N) - P^u(t|t-1))B_u^T(t-1)$$

Parameter smoothing filter for $t=N,\ldots,1$

$$B_a(t-1) = P^a(t-1|t-1)(P^a(t|t-1))^{-1}$$

$$\vec{a}(t-1|N) = \vec{a}(t-1|t-1) + B_a(t-1)(\vec{a}(t|N) - \vec{a}(t|t-1))$$

$$P^a(t-1|N) = P^a(t-1|t-1) + B_a(t-1).\ .(P^a(t|N) - P^a(t|t-1))B_a^T(t-1)$$

Maximization step

$$A(t|N) := A^{\vec{a}(t-1|N)}$$

$$Q_u^{(m+1)} = \frac{1}{N}\sum_{t=1}^{N}\{(\vec{u}(t|N) - A(t|N)\vec{u}(t-1|N)) \cdot (\vec{u}(t|N) - A(t|N)\vec{u}(t-1|N)^T + P^u(t|N) - A(t|N)B_u(t-1)P^u(t|N) - P(t|N)^u B_u^T(t-1)A^T(t|N) + A(t|N)P^u(t-1|N)A^T(t|N)\}$$

$$R^{(m+1)} = \frac{1}{N}\sum_{t=1}^{N}\{(\vec{y}(t) + C_u P^u(t|N)C_u^T - C_u\vec{u}(t|N)(\vec{y}(t) - C_u\vec{u}(t|N))^T\}$$

$$Q_a^{(m+1)} = \frac{1}{N}\sum_{t=1}^{N}\{(\vec{a}(t|N) - \vec{a}(t-1|N)) \cdot (\vec{a}(t|N) - \vec{a}(t-1|N)^T + P^a(t|N) - B_a(t-1)P^a(t|N) - P^a(t|N)B_a^T(t-1) + P^a(t-1|N)\}$$

Table s2. Source location of right somatosensory MEG for HC and PAE groups

Source location	HC group	PAE group
Left primary somatosensory cortex	21	19
Right secondary somatosensory cortex	17	17
Left secondary somatosensory cortex	15	17
Left secondary auditory cortex	5	4
Right primary somatosensory cortex	5	3
Anterior cingulate cortex	2	1
Left supramarginal gyrus	2	1
Right premotor cortex	1	0
Right supramarginal gyrus	1	0
Colliculus	1	0
Thalamus	1	0
Left intraparietal sulcus	1	0
Medial prefrontal cortex	1	1
Right secondary auditory cortex	1	0
Left primary auditory cortex	1	0
Posterior cingulate gyrus	1	1
Eye	0	3

Table s3. Wilcoxon test results for the averaged spectral coherence within 5-30 Hz between HC and PAE subjects among all three source locations (SI-l and SII-l, SI-l and SII-r, SII-l and SII-r) respectively

Coherence	*P*	*t*
SI-l and SII-l	0.4108	0.8341
SI-l and SII-r	0.3635	0.9218
SII-l and SII-r	0.0070*	2. 8933

*$P<0.05/3$

Table s4. Maternal drinking levels of PAE participants

Subjects No.	Maternal drinking levels	PAE group
1	all	cPAE
2	all	cPAE
3	all	cPAE
4	1st trimester	oPAE
5	all	cPAE
6	off and on	oPAE
7	all	cPAE
8	off and on	oPAE
9	all	cPAE
10	all	cPAE
11	all	cPAE
12	all	cPAE
13	off and on	oPAE
14	all	cPAE
15	all	cPAE
16	off and on	oPAE
17	all	cPAE
18	off and on	oPAE
19	all	cPAE

"all" – child was exposed to alcohol consistently throughout the whole pregnancy.

"off and on" – child was exposed to alcohol occasionally but throughout the whole pregnancy.

"1st trimester" – child was only exposed to alcohol during the first trimester of the pregnancy.

135 “oPAE” – occasional prenatal alcohol exposure group

136 “cPAE” – consistent prenatal alcohol exposure group

137

Table s5. T-test results of the normalized mean rPDC values, i.e. the strength of the connections, between HC, oPAE and cPAE groups with significant differences after Bonferroni correction

rPDC	*P*	*t*
from SII-l to SI-l (HC vs oPAE, 0-200 ms)	<0.0001**	5.1220
from SII-l to SI-l (HC vs cPAE, 0-200 ms)	<0.0001**	6.7820
from SII-l to SII-r (HC vs oPAE, 201-320 ms).	<0.0001**	6.2215
from SII-l to SII-r (HC vs cPAE, 201-320 ms)	<0.0001**	10.5880

**P<0.01/18

142 Table s6. Cognitive measurement of HC participants

Subjects No.	IQ	SWM Choice Latency	CGT Impulsivity Index	IED Stage1 Latency
1	91	793.97	44.91	2528.29
2	118	762.15	42.01	1672.67
3	116	677.90	31.10	1691.50
4	122	612.70	9.44	1427.29
5	106	1074.49	34.86	1491.00
6	100	774.05	33.57	2085.50
7	97	950.71	24.17	2691.00
8	73	879.53	47.92	1941.14
9	90	646.83	57.22	1911.38
10	120	884.67	63.47	3011.57
11	107	841.90	43.91	1606.33
12	123	944.25	40.14	1145.57
13	125	724.70	38.06	4108.67
14	110	821.31	19.86	1114.57
15	99	809.38	16.53	1847.88
16	117	1292.92	32.08	3616.57
17	111	890.17	8.31	1457.17
18	93	795.43	30.75	2411.50
19	120	802.51	26.25	1433.50
20	97	617.47	35.15	2374.57
21	127	410.79	46.75	2279.56

Subjects No.	IQ	SWM Choice Latency	CGT Impulsivity Index	IED Stage1 Latency
Average	107.71	809.90	34.59	2087.96

The neuropsychological measurements include “IQ” (Intelligence Quotient), “CGT impulsivity Index” (the impulsivity index of Cambridge Gambling Task), “SWM Choice Latency” (the latency of spatial working memory task) and “IED Stage1 Latency” (the latency of intra-extra dimensional shift task).

148 Table s7. Cognitive measurement of PAE participants

Subjects No.	IQ	SWM Choice Latency	CGT Impulsivity Index	IED Stage1 Latency
1	94	872.61	73.33	2595.91
2	96	919.34	71.37	2113.11
3	94	1613.59	41.81	7260.43
4	70	1018.04	55.05	2107.30
5	79			
6	91	748.57	58.75	2055.17
7	54	3174.42	50.67	4051.33
8	83		72.64	4300.25
9	75	1676.94	44.86	2562.38
10	58	750.73	63.13	11848.17
11	99	1057.25	42.33	2404.75
12	101	1088.35	15.56	10830.50
13	73	1199.12	45.17	3086.50
14	81		72.27	2218.22
15	93	605.44	48.14	2483.57
16	71	839.31	52.34	3439.67
17				
18	90	687.73	35.71	3371.625
19	75	1010.69	46.81	
Average	82.06	1150.81	52.35	4170.55

149 The neuropsychological measurements include “IQ”, “CGT impulsivity Index”, “SWM Choice Latency” and “IED Stage1 Latency”.

Table s8. Pearson correlation analysis between the rPDC values (201 to 320 ms, SII-l to SII-r) and the cognitive measures for the pooled subjects (HC & PAE group), HC group and PAE group, respectively

Cognitive measures	Group	*r*	*P*
IQ	HC & PAE	0.5912	<0.0001**
	HC	-0.1285	0.5789
	PAE	0.2639	0.2899
SWM Choice Latency	HC & PAE	-0.3387	0.0433
	HC	-0.1666	0.4705
	PAE	0.1692	0.5467
CGT Impulsivity Index	HC & PAE	-0.4862	0.0020*
	HC	-0.1310	0.5713
	PAE	0.3150	0.2181
IED Stage 1 Latency	HC & PAE	-0.5631	0.0003**
	HC	-0.9242	<0.0001**
	PAE	-0.7743	0.0004**

*P<0.05/8

**P<0.01/8

REFERENCES

[1] Ranken DM, Stephen JM, George JS. MUSIC seeded multi-dipole MEG modeling using the constrained start spatio-temporal modeling procedure. *Neurology & Clinical Neurophysiology.* 2004;2004:80.

[2] Hämäläinen MS, Ilmoniemi RJ. Interpreting magnetic-fields of the brain-minimum norm estimates. *Medical & Biological Engineering & Computing,* 1994;32(1):35-42.

[3] Schelter B, Mader M, Mader W, et al. Overarching framework for data-based modelling. *European Phys. Lett.* 2014;83:30004.

[4] Gao L, Sommerlade L, Coffman B, et al. Granger causal time-dependent source connectivity in the somatosensory network. *Scientific Reports*. 2015;5:10399.