Supplementary Material for the following paper

Multi-Subject Analysis for Brain Developmental Patterns Discovery via Tensor Decomposition of MEG Data

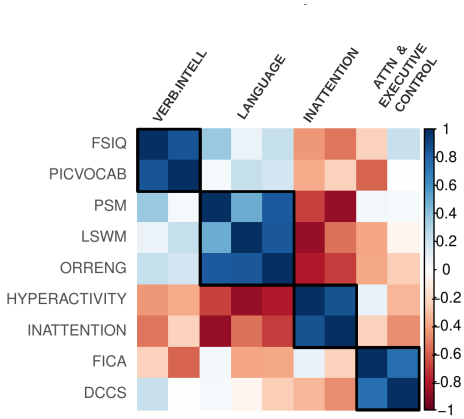
Irina Belyaeva[1*], Ben Gabrielson[1], Yu-Ping Wang[2], Tony W. Wilson[3], Vince D. Calhoun[4], Julia M. Stephen[5] and Tülay Adali[1]

[1*]Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, Maryland, USA.
[2]Department of Biomedical Engineering, Tulane University, New Orleans,LA, USA.
[3]The Institute for Human Neuroscience, Boys Town National Research Hospital, Boys Town, NE, 68010, USA.
[4]Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Atlanta, GA, USA.
[5]The Mind Research Network, Lovelace Biomedical Research Institute, Albuquerque, NM, USA.

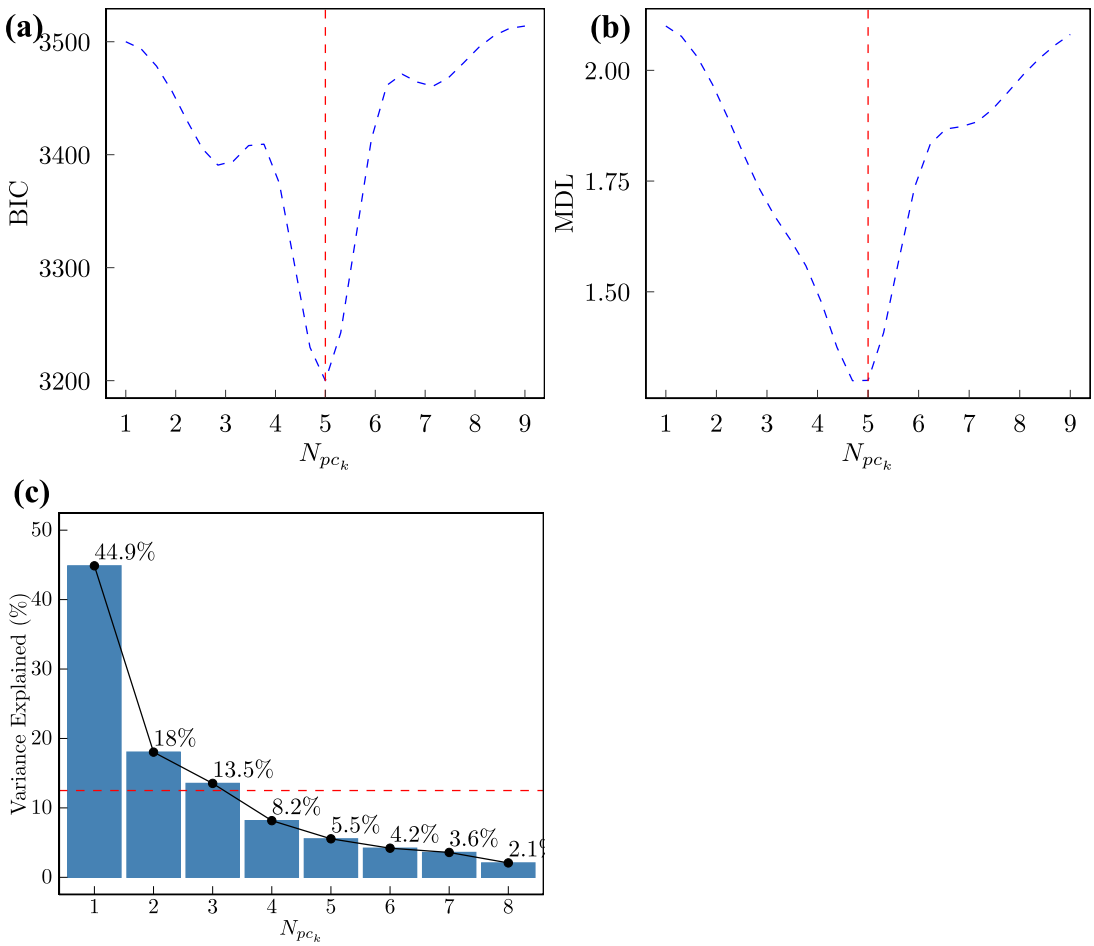
*Corresponding author(s). E-mail(s): irinbel1@umbc.edu;
Contributing authors: bengabr1@umbc.edu; wyp@tulane.edu; tony.wilson@boystown.org; vcalhoun@gsu.edu; jstephen@mrn.org; adali@umbc.edu;

Table of Content

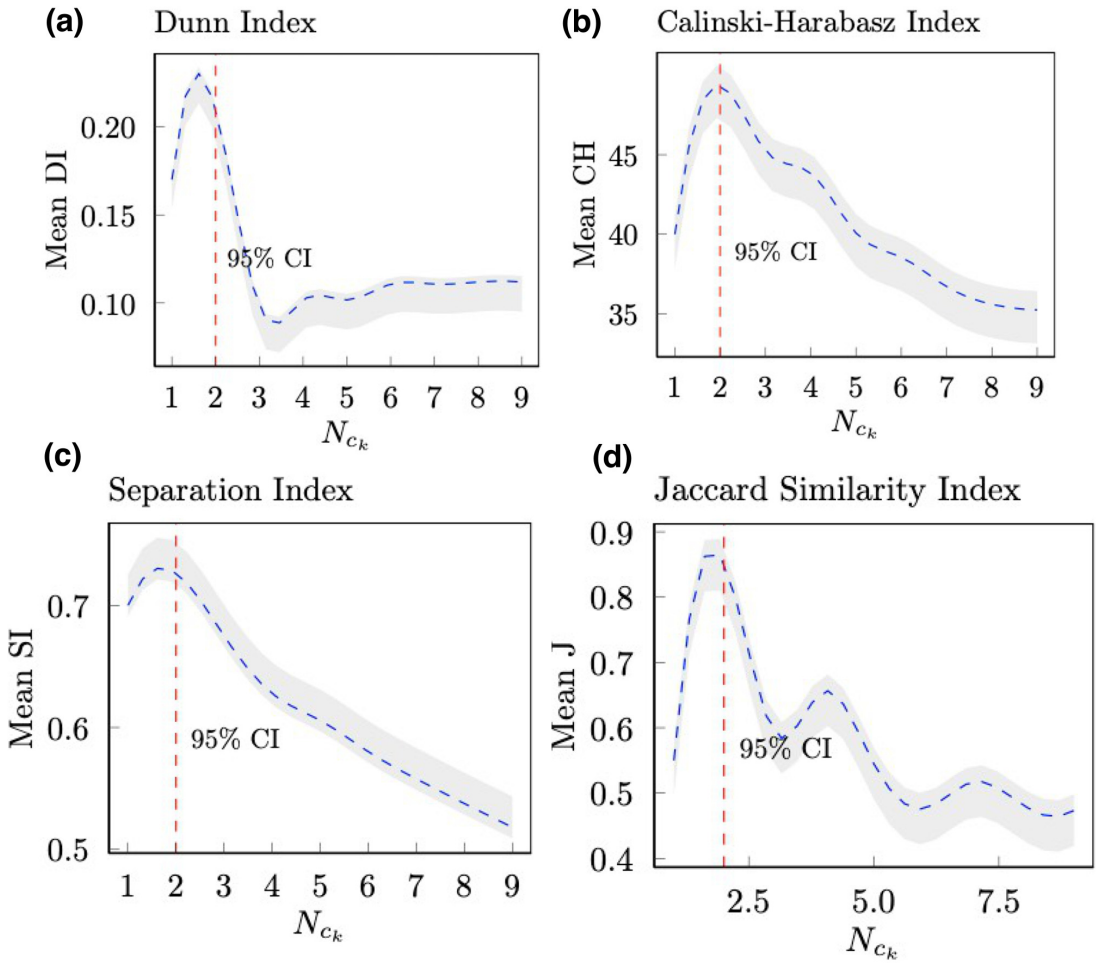
1. Supplementary Figures



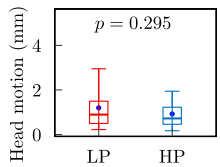
Supplementary Fig S.1. Partial correlations of neurophysiological (T) scores controlled for age, gender and parental socioeconomic status.



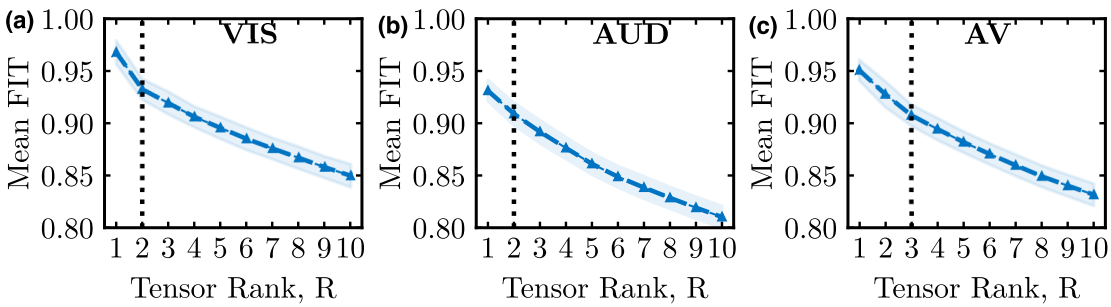
Supplementary Fig S.2. Evaluating the optimal number of principal components used in the HCPC algorithm on the neuropsychological measures dataset. Minimum BIC and MDL values were used as model selection criteria. The optimal model was selected based on the minimum BIC and MDL. Top: BIC and MDL a function of the number of principal components. Five principal components was the most appropriate number of principal components, as indicated by the BIC and MDL values. The selected number of principal components explained a large part of the dataset, as supported by the value of the cumulative variance explained, 89 (%), when the number of PCs $N_{pck} = 5$. (A) BIC. (B) MDL. (C) Percentage of variance explained for a chosen number of principal components.



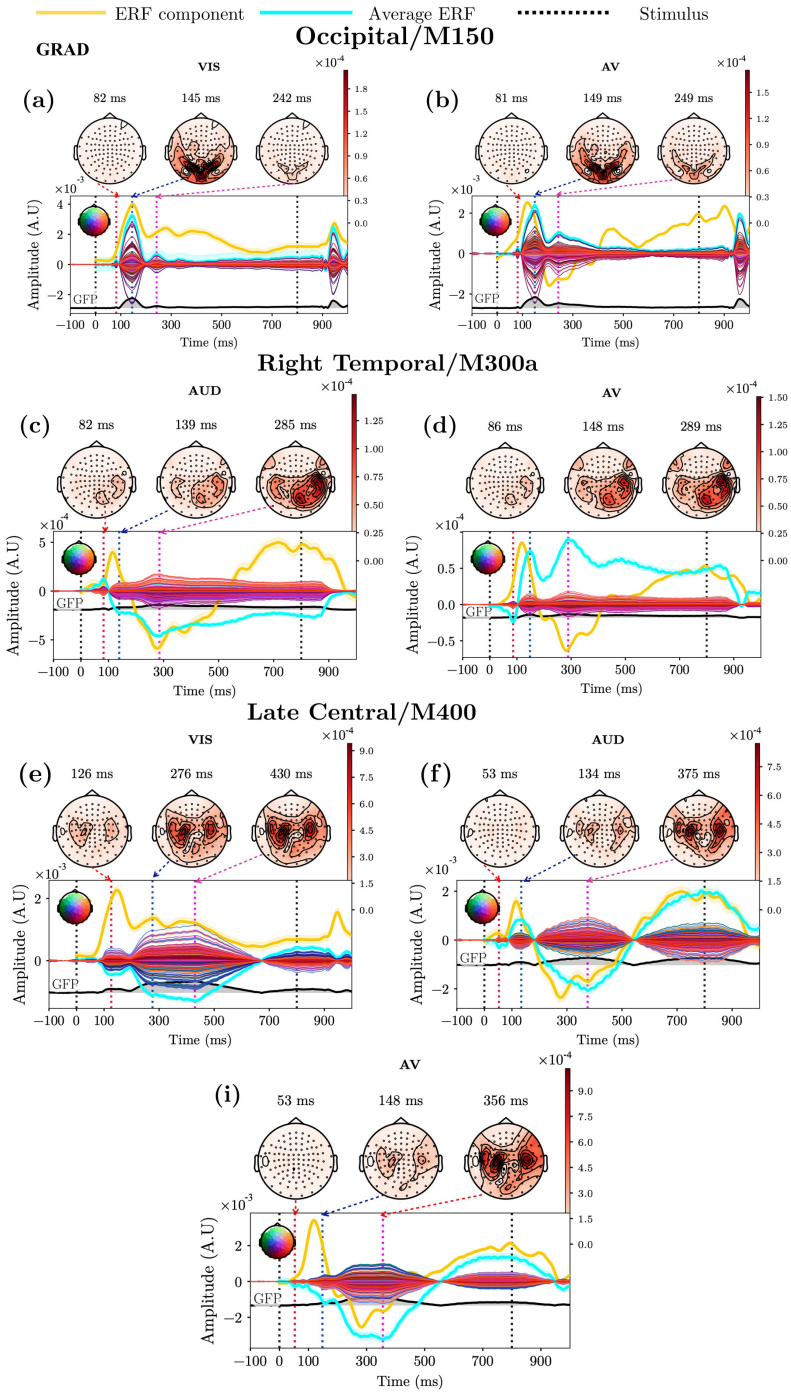
Supplementary Fig S.3. Evaluation of the goodness of the clustering solution as a function of the number of clusters. The goodness of fit is accessed using the elbow criterion of the cluster validity indexes that maximize between-cluster distances and minimize within-cluster distances. Mean values of cluster validity indices are shown along with a 95% confidence interval (CI). Three cluster quality metrics show an inflection point when the number of clusters $N_{c_k} = 2$. Cluster stability is evaluated using the Jaccard similarity index via a nonparametric bootstrap resampling technique of subject neuropsychological measures with $n = 1000$ repetitions. The reproducibility of the clusters was established with a high degree of cluster similarity for the Jaccard index $J = 0.9$ at $N_{c_k} = 2$. Thus, the empirically obtained cluster similarity measure suggests that the clusters derived from the original data reflect the properties of the subject neuropsychological measures dataset. The cluster validity indexes and cluster stability metrics provide strong evidence that the optimal number of clusters for a given neurophysiological dataset is two. (a) Dunn index. (b) Calinski-Harabasz index. (c) Separation index. (d) Jaccard similarity index.



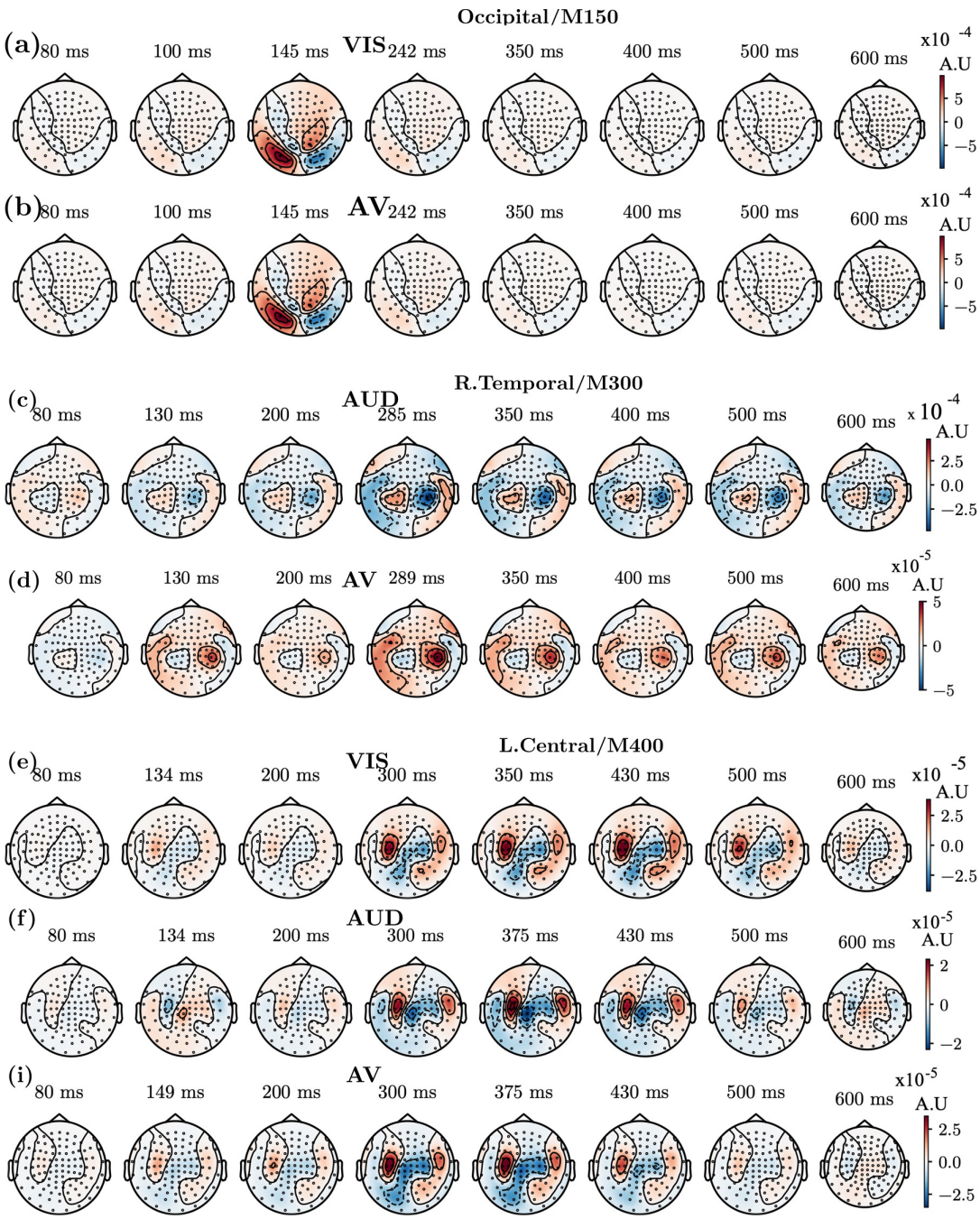
Supplementary Fig S.4. Main effect of the subject group (HP vs. LP, $N = 170$) on the subject head motion is summarized in the boxplots. The boxplots summarize the distribution of the mean values of the head motion during the scan. The tops of the whisker lines indicate the 25th and 75th percentile values. The mean values are plotted in blue. The error bars represent the standard error of the mean. Post hoc analyses with two-tailed t-tests (FDR corrected, $p < 0.05$) indicate that the mean value of of the head motion of the HP group was not significantly different than for LP group with $p > 0.05$.



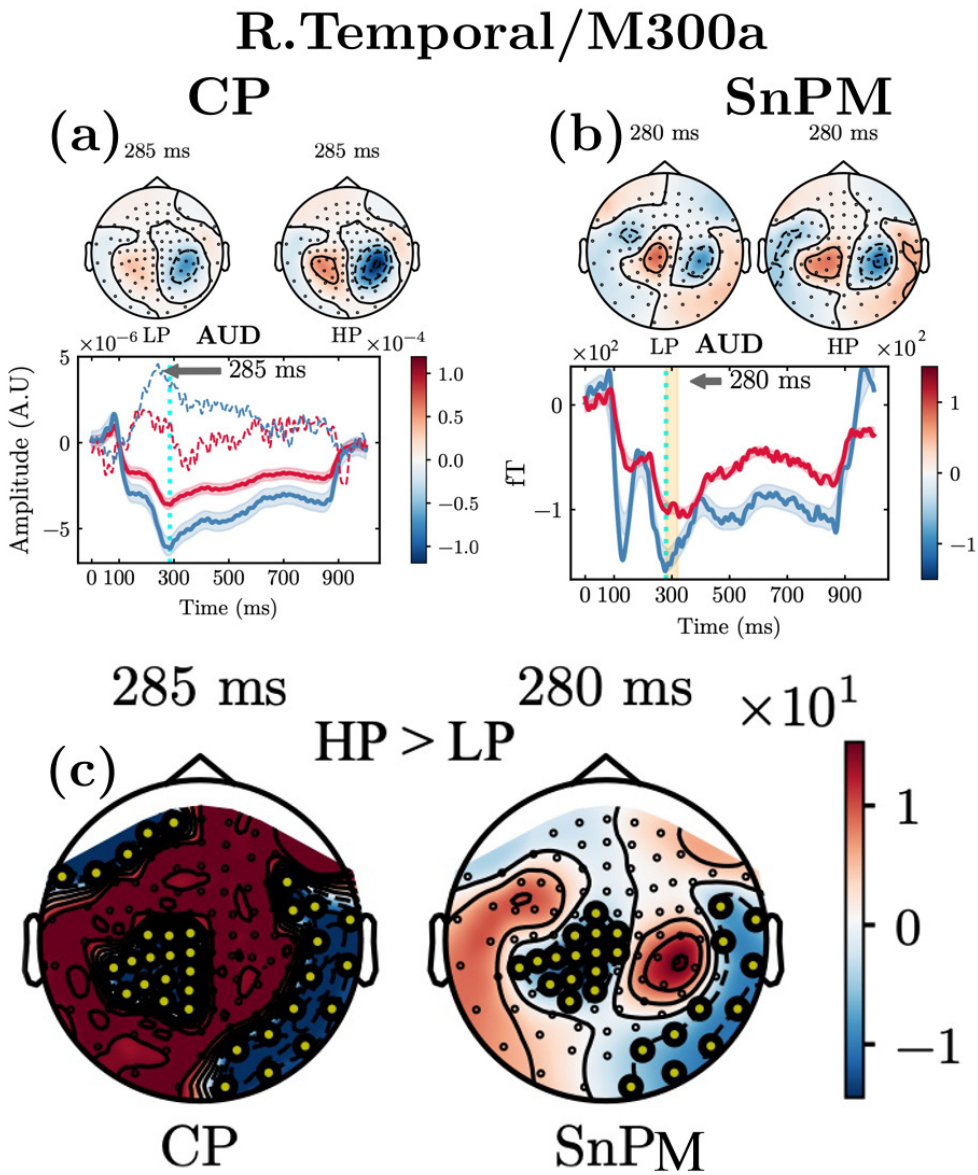
Supplementary Fig S.5. Distribution of mean FIT values of the CP decomposition for MEG data by stimulus condition as a function of tensor rank. Mean and standard error are shown. Shaded areas around each line depict the standard error of the mean. Dotted line denotes the selected tensor rank used in the CP model. Details are listed in Supplementary Table S.2.



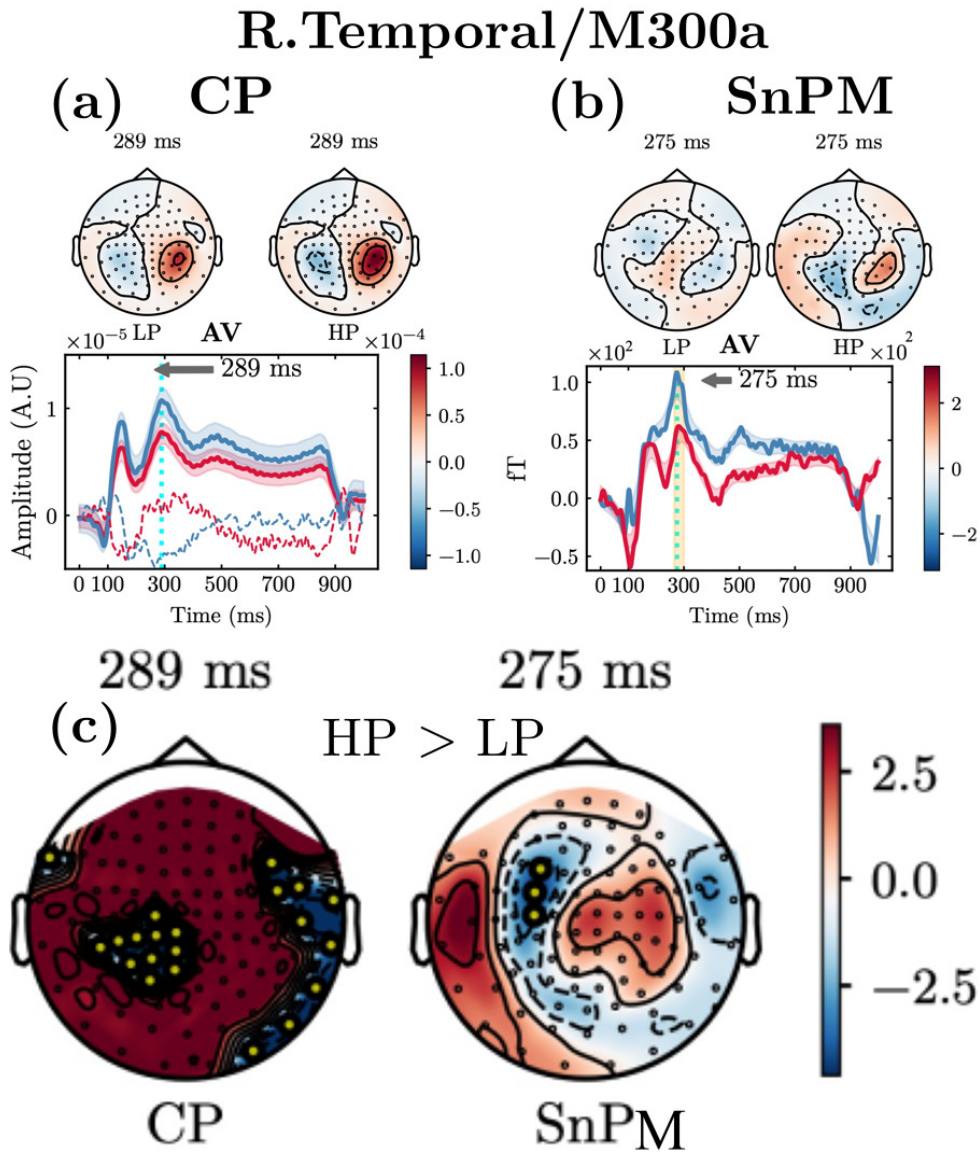
Supplementary Fig S.6. Tensor decomposition results of sensor-level MEG data for target stimuli. Temporal and spatial patterns from the components of the CP model. Top: The topographic maps (gradiometers view) show the density of spatial patterns that correspond to prominent time peaks denoted with red and blue arrows. Bottom: ERF component with signal traces from all individual MEG sensors averaged across subject ERF components. The shaded areas around each line depict the standard error of the mean. The average stimulus-related ERF timecourse is shown in yellow, and the average ERF (average across sensors) component is plotted in cyan. (a)-(b) Occipital 130-150 ms component in the VIS and AV conditions. (c) Right temporal 280-300 ms component in the AUD and AUD conditions. (e)-(i) Late central 350-430 ms component in the VIS, AUD, and AV conditions.



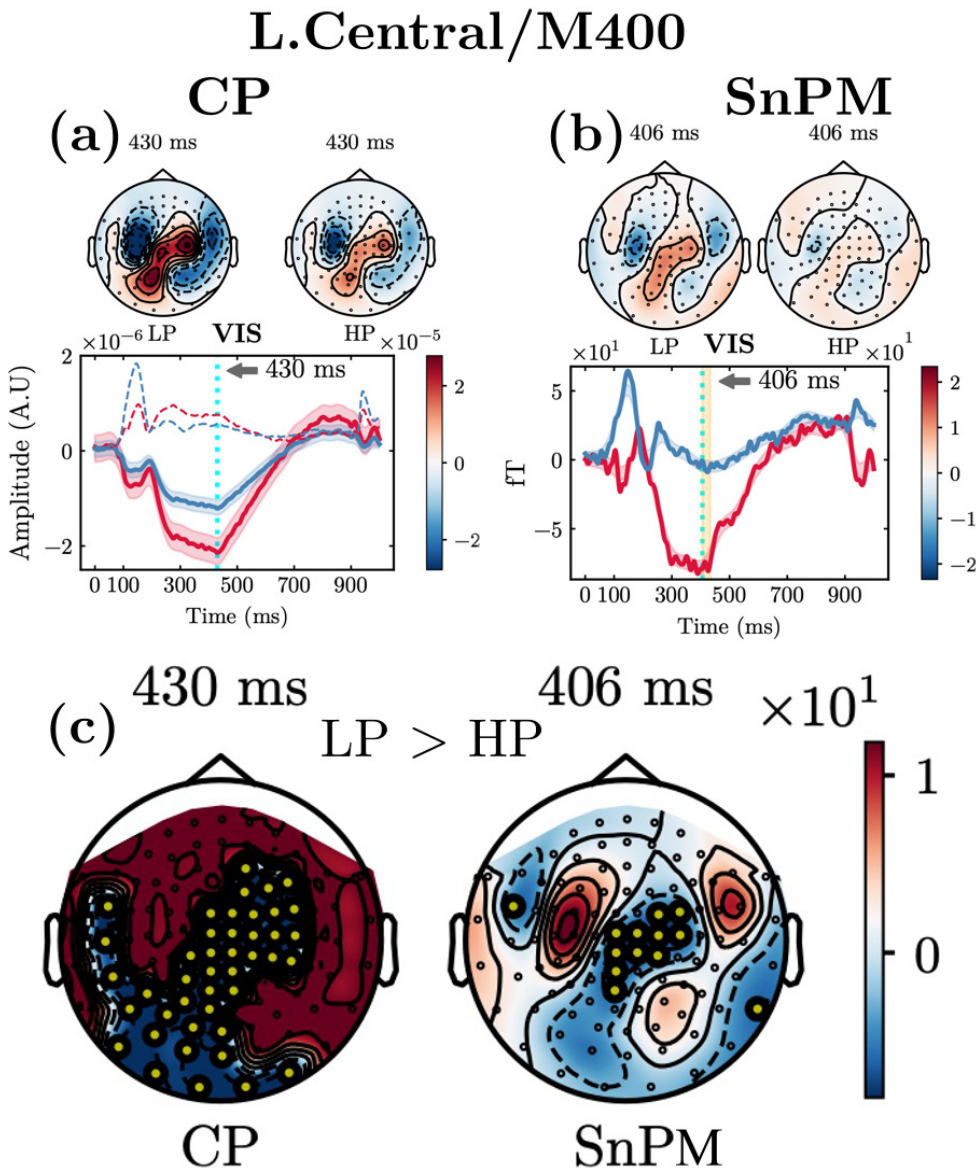
Supplementary Fig S.7. ERF components topographic maps over time are shown. (a)-(b) Occipital/M150 component. (c)-(d) R.Temporal/M300a component. (e)-(i) L.Central/M400 component.



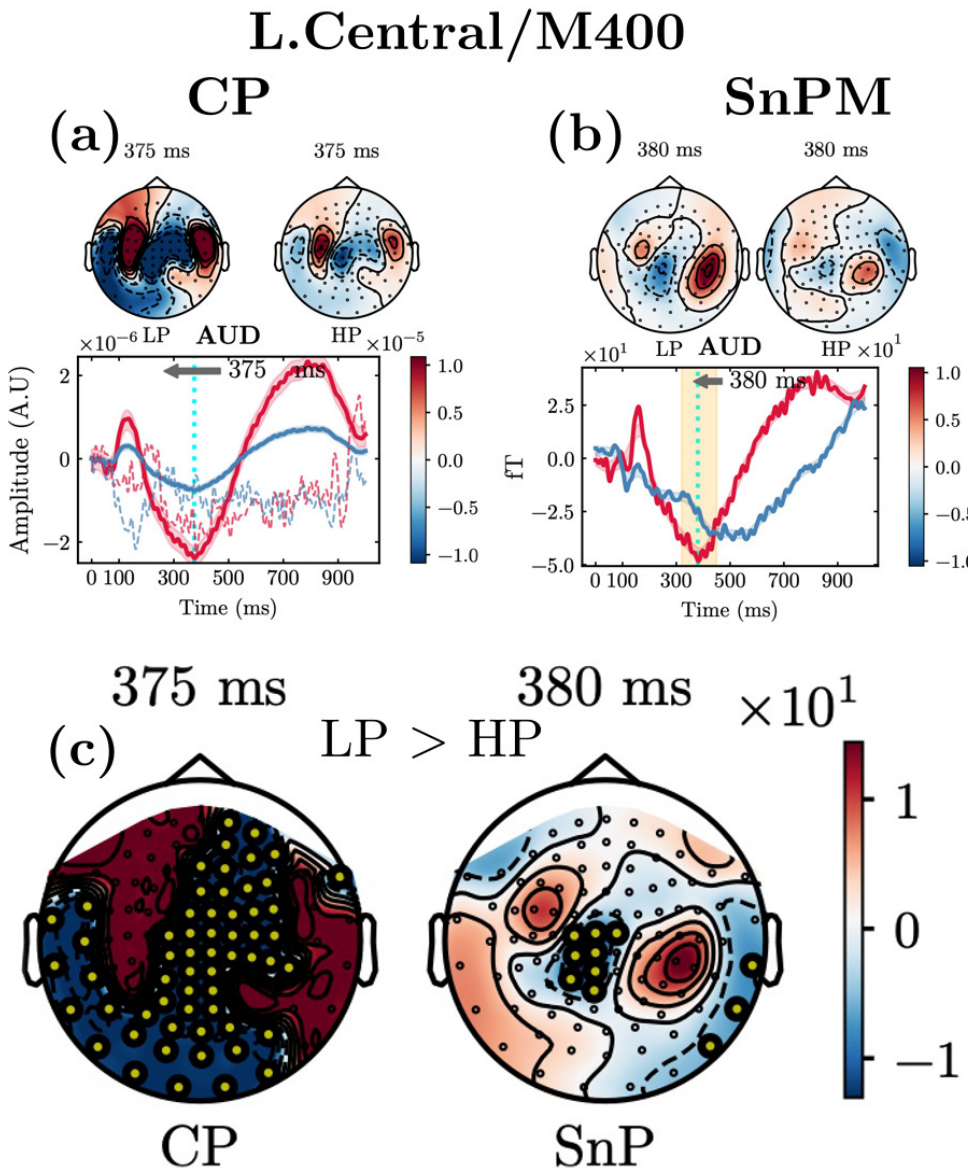
Supplementary Fig S.8. Sensitivity analysis of the ERF components generated with different group imaging methods. Estimation of the R.Temporal/M300a component for the AUD condition by the CP decomposition and SnPM methods.The group-level T-maps between HP and LP groups for the CP and SnPM methods are shown. The T-maps (nonparametric permutation two-tailed t-test with a maximum t-statistics) are thresholded at $p < 0.05$. Siginificant sensors are overlayed on T-map are shown in yellow. (a) CP AUD M300 component. (b) SnPM AUD M300 component. The significant time interval of group differences (275-320 ms) is depicted in the shaded area. (e) Left: CP AUD M300a T-map. Right: SnPM AUD M300a T-map.



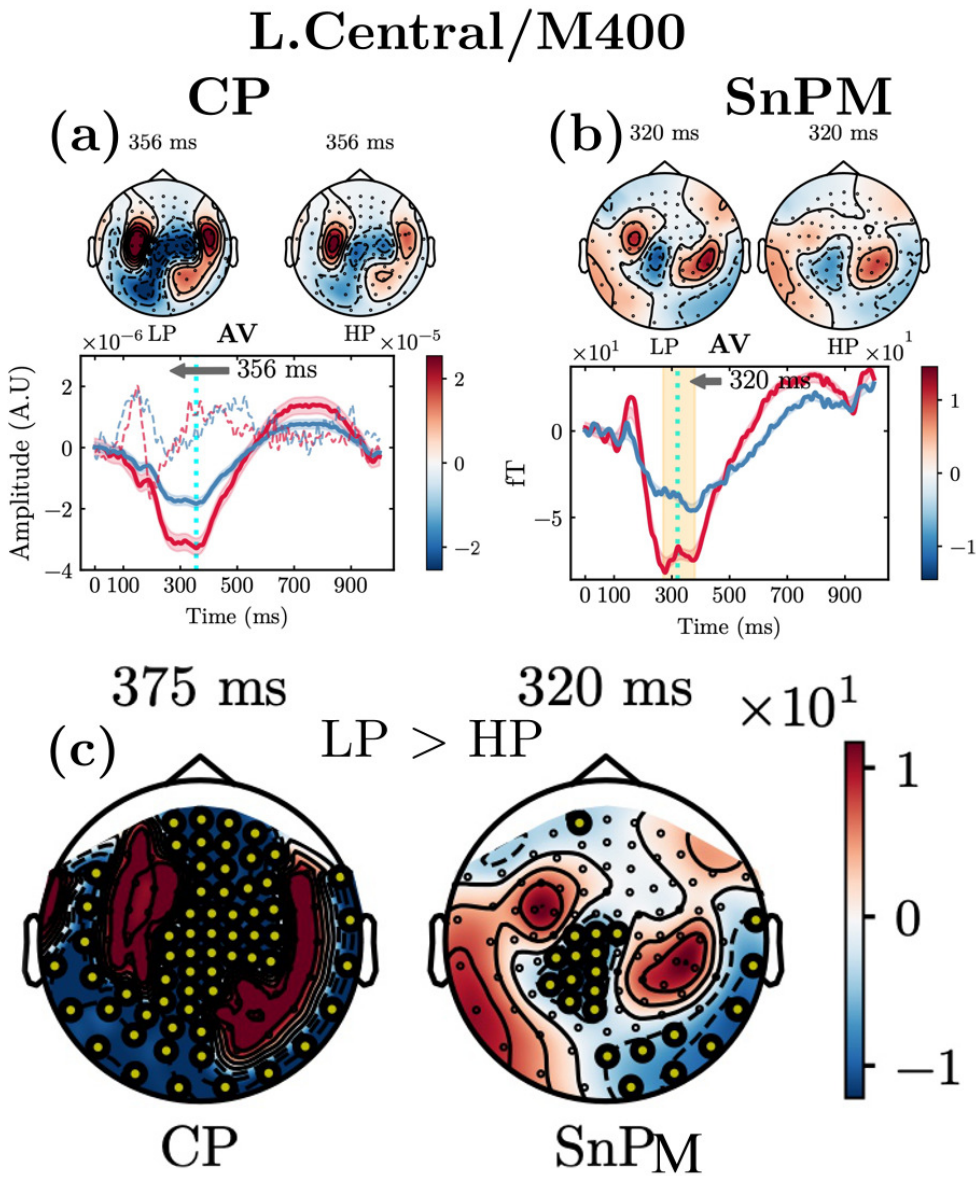
Supplementary Fig S.9. Sensitivity analysis of the ERF components generated with different group imaging methods. Top row: Estimation of the R.Temporal/M300a component for the AV condition by the CP decomposition and SnPM methods. Bottom row: The group-level T-maps between HP and LP groups for the CP and SnPM methods are shown. The T-maps (nonparametric permutation two-tailed t-test with a maximum t-statistics) are thresholded at $p < 0.05$. The yellow circles on scalp maps show the location of significant sensors. (a) CP AV M300 component. (b) SnPM AV M300 component. The significant time interval of group differences (275-320 ms) is depicted in the shaded area. (e) Left: CP AV M300a T-map. Right: SnPM AV M300a T-map.



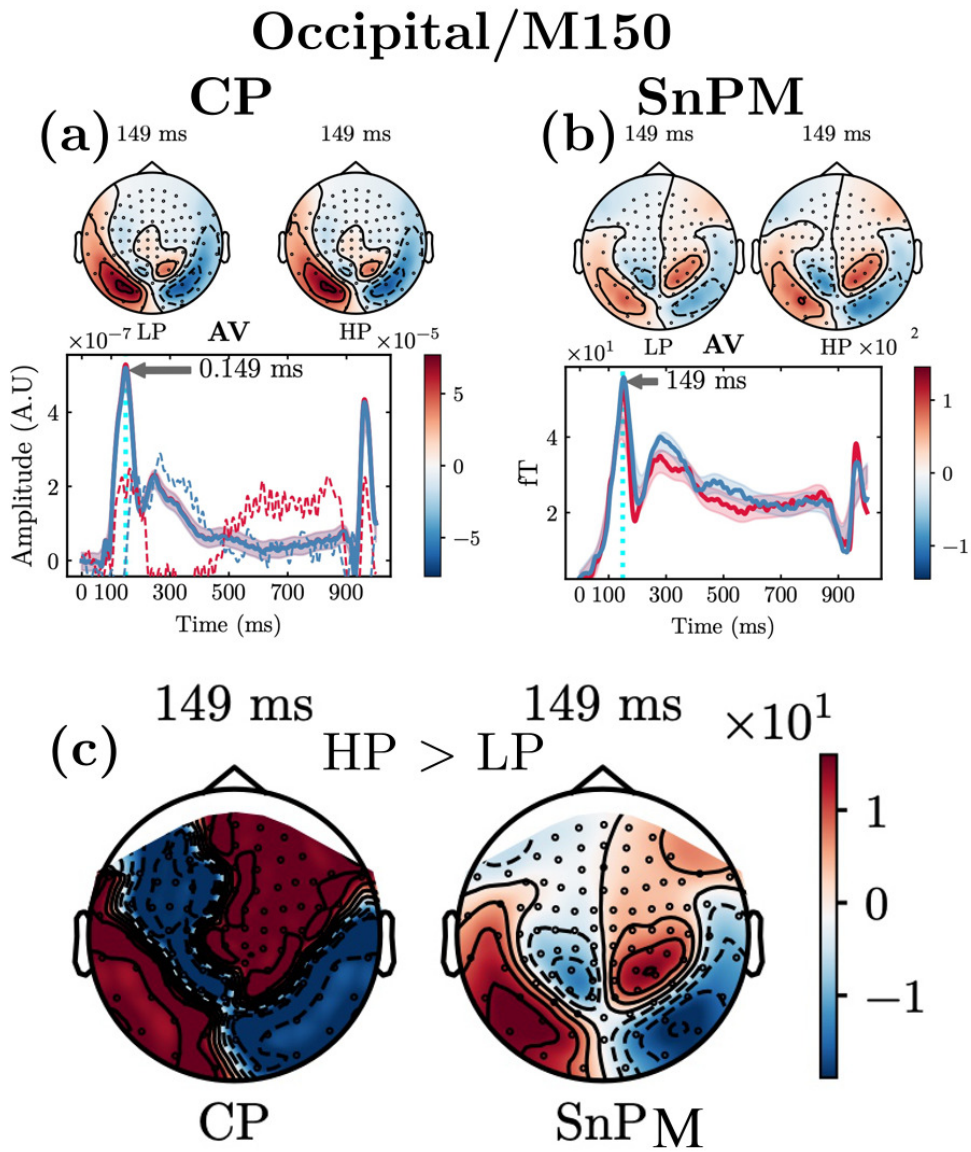
Supplementary Fig S.10. Sensitivity analysis of the ERF components generated with different group imaging methods. Estimation of the L.Central/M400 component for the VIS condition by the CP decomposition and SnPM methods. The group-level T-maps between HP and LP groups for the CP and SnPM methods are shown. The T-maps (nonparametric permutation two-tailed t-test with a maximum t-statistics) are thresholded at $p < 0.05$. The yellow circles on scalp maps show the location of significant sensors. (a) CP VIS M400 component. (b) SnPM VIS M400 component. The significant time interval of group differences (400-433 ms) is depicted in the shaded area. (e) Left: CP VIS M400 T-map. Right: SnPM VIS M400 T-map.



Supplementary Fig S.11. Sensitivity analysis of the ERF components generated with different group imaging methods. Estimation of the L.Central/M400 component for the AUD condition by the CP decomposition and SnPM methods. The group-level T-maps between HP and LP groups for the CP and SnPM methods are shown. The T-maps (nonparametric permutation two-tailed t-test with a maximum t-statistics) are thresholded at $p < 0.05$. The yellow circles on scalp maps show the location of significant sensors. (a) CP AUD M400 component. (b) SnPM AUD M400 component. The significant time interval of group differences (320-450 ms) is depicted in the shaded area. (e) Left: CP AUD M400 T-map. Right: SnPM AUD M400 T-map.



Supplementary Fig S.12. Sensitivity analysis of the ERF components generated with different group imaging methods. Estimation of the L.Central/M400 component for the AV condition by the CP decomposition and SnPM methods. The group-level T-maps between HP and LP groups for the CP and SnPM methods are shown. The T-maps (nonparametric permutation two-tailed t-test with a maximum t-statistics) are thresholded at $p < 0.05$. The yellow circles on scalp maps show the location of significant sensors. (a) CP AV M400 component. (b) SnPM AV M400 component. The significant time interval of group differences (275-330 ms) is depicted in the shaded area. (e) Left: CP AV M400 T-map. Right: SnPM AV M400 T-map.



Supplementary Fig S.13. Sensitivity analysis of the ERF components generated with different group imaging methods. Top row: Estimation of the Occipital/M150 component for the AV condition by the CP decomposition and SnPM methods. Bottom row: The group-level T-maps between HP and LP groups for the CP and SnPM methods are shown. The T-maps (nonparametric permutation two-tailed t-test with a maximum t-statistics) are thresholded at $p < 0.05$. The AV M150 component did not discriminate betwee subject subgroups. No significant spatiotemporal regions were identified. (a) CP AV M150 component. (b) SnPM AV M150 component. (c) Left: CP AV M150 T-map. Right: SnPM AV M150 T-map.

15

2. Supplementary Tables

Supplementary Table S.1. Comparison of participants demographics and neuropsychological (T) scores by group

	HP ($n = 89$)	LP ($n = 81$)	$t(167)/\tilde{\chi}^2$	p value
Gender (M/F) ratio[b]	43M (48%), 46F	42M (51%), 39F	0.00	0.99[d]
Age (years)[c]	11.92 (1.72)	11.92 (1.84)	0.03	0.97[d]
Parental SES[c]	44.69 (11.53)	45.56 (9.85)	0.517	0.30[d]
FSIQ[c]	120.58 (11.27)	102.16 (11.98)	9.16	$<$ **0.0001**[d]
PICVOCAB[c]	60.28 (12.06)	50.03 (7.17)	5.99	$<$ **0.0001**[d]
ORRENG[c]	60.46 (11.11)	48.63 (8.92)	6.79	$<$ **0.0001**[d]
LSWM[c]	57.46 (10.16)	47.71 (7.64)	6.27	$<$ **0.0001**[d]
PSM[c]	57.54 (11.86)	49.58 (9.05)	4.36	$<$ **0.0001**[d]
DCCS[c]	56.79 (12.53)	47.3 (9.31)	4.96	$<$ **0.0001**[d]
FICA[c]	49.06 (10.57)	42.23 (6.24)	4.55	$<$ **0.0001**[d]
HYPERACTIVITY[c]	47.24 (6.79)	51.63 (11.35)	-2.17	**0.031**[d]
INATTENTION[c]	51.03 (9.68)	51.9 (9.43)	0.52	0.59[d]

[a] Summaries are mean (SD) for continuous variables, $n(\%)$ for categorical variables.
[b] Pearson chi-square $\tilde{\chi}^2$ test.
[c] Independent samples two-tailed t-test.
[d] p values (FDR corrected, $p < 0.05$).

Supplementary Table S.2. Mean and standard deviation of the FIT for the CP model by stimulus condition and tensor rank, R

	FIT[a]		
	Stimuli condition		
R	VIS	AUD	AV
	M (SD)	M (SD)	M (SD)
1	0.967 (0.01)	0.931 (0.01)	0.095 (0.01)
2	**0.932**[b](0.01)	**0.909**[b](0.01)	0.927 (0.02)
3	0.919 (0.01)	0.892 (0.01)	**0.907**[b](0.01)
4	0.906 (0.01)	0.876 (0.01)	0.894 (0.01)
5	0.895 (0.01)	0.861 (0.01)	0.882 (0.01)
6	0.885 (0.01)	0.849 (0.01)	0.870 (0.01)
7	0.876 (0.01)	0.838 (0.01)	0.861 (0.01)
8	0.867 (0.01)	0.829 (0.02)	0.849 (0.03)
9	0.858 (0.02)	0.810 (0.03)	0.840 (0.05)
10	0.850 (0.02)	0.810 (0.03)	0.832 (0.05)

[a] The plots corresponding to this table can be seen in Supplementary Fig. S.5.
[b] The FIT value corresponds to the chosen tensor rank R of the final CP model fit for each stimulus condition (VIS, AUD or AV).

Supplementary Table S.3. Comparison of the extracted MEG ERF components with the original ERF datasets using repeated measures correlation analysis

Component	VIS				AUD				AV			
	$r(224399)$	R^2	95% CI	p value	$r(224399)$	R^2	95%	p value	$r(224399)$	R^2	95%	p value
Occipital/M150	0.588	0.34	[0.58, 0.58]	<0.001[a]	-		-	-	0.419	0.22	[0.41, 0.42]	<0.001[a]
Right Temporal/M300a	-	-	-	-	0.677	0.49	[0.66, 0.68]	<0.0001	0.395	0.28	[0.39, 0.40]	<0.001[a]
Late Central/M400	0.625	0.38	[0.62, 0.63]	<0.001[a]	0.519	0.27	[0.51, 0.52]	<0.001[a]	0.492	0.24	[0.49, 0.51]	<0.001[a]

[a] p values (FDR corrected, $p < 0.05$).

Supplementary Table S.4. Comparison of ERF timecourse group amplitudes by subject subgroup extracted by statistical nonparametric mapping (SnPM) method

Component	AUD				VIS				AV			
	HP ($n = 89$)	LP ($n = 81$)	$t(165)$	p value	HP ($n = 89$)	LP ($n = 81$)	$t(165)$	p value	HP ($n = 89$)	LP ($n = 81$)	$t(165)$	p value
	$M(SD)$				$M(SD)$				$M(SD)$			
Occipital/M150	-	-	-	-	95.5(64.4))	35.0(44.9)	6.32	**< 0.001**[a]	1150(1745)	−780(1835)	0.175	0.851[a]
Right Temporal/M300a	115(80)	85.3(64.1)	2.31	0.022[a]	-	-	-	-	74.2(98.6)	39.0(95.9)	2.09	0.038[a]
Late Central/M400	-55.2(45.8)	-24.7(47.6)	-3.76	0.002[a]	-7.06(47.5)	-38.3 (58)	-3.41	0.001[a]	-32.7(51.6)	49.2(55.0)	-1.81	0.07[a]

[a] p values (FDR corrected, $p < 0.05$) from post hoc two-tailed t-tests adjusted for age, gender, and parental SES.

Supplementary Table S.5. CP ERF components orthogonality analysis

Component 1	Component 2	Condition		
		VIS	AUD	AV
		$\langle$Component 1, Component 2$\rangle$[a]	$\langle$Component 1, Component 2$\rangle$[a]	$\langle$Component 1, Component 2$\rangle$[a]
M150	M300a	-	-	0.095
M300a	M400	-	0.068	0.101
M150	M400	0.079	-	0.112

[a] $\langle \cdot, \cdot \rangle$ inner product of the arguments

Supplementary Table S.6. CP ERF components correlations

Component 1	Component 2	Condition					
		VIS		AUD		AV	
		r	p	r	p	r	p
M150	M300a	-		-		0.004	0.958[a]
M300a	M400	-		0.008	0.917[a]	0.001	0.989[a]
M150	M400	0.007	0.927[a]	-		0.006	0.938[a]

[a] p values (FDR corrected, $p < 0.05$).

3. Supplementary Methods

3.1. Clustering Analysis of Neurophysiological Dataset

To assess the potential interrelatedness of the observed event-related-field (ERF) patterns with the neurophysiological (T) scores across subjects, we performed a clustering analysis using selected variables that are representative of major cognitive domains. Before conducting clustering analysis, covariates of no interest, namely, age, parental socioeconomic (SES), and gender, were regressed out from the original neurophysiological dataset. The neurophysiological dataset contains multiple continuous variables that are highly correlated (see Supplementary Fig. S.1). Principal component analysis (PCA) can reduce the dimensions of data into a few continuous variables, thereby maximizing the variance explained. Thus, we employed a hierarchical clustering on principal components (HCPC) approach (Argüelles, Benavides, & Fernández, 2014; Husson, Josse, & Pagès, 2010), combining three standard techniques (PCA, hierarchical clustering, and K-means algorithm) to achieve an improved clustering solution. Fig. 5 of the main text shows a schematic view of the HCPC method. All neurophysiological (T) scores were standardized prior to the application of the HCPC algorithm using z-score formula to account for different scales. First, we applied the PCA algorithm to the neurophysiological dataset, represented as a subject score matrix $\mathbf{P} \in \mathbb{R}^{K \times L}, K = 170, L = 9$, to reduce the dataset into fewer dimensions, i.e., principal components (PCs), which were uncorrelated. We used Bayesian information criterion (BIC) (Schwarz, 1978) and minimum description length (MDL) (Rissanen, 1978) as model selection criteria for the number of PCs retained ($N_{p_{ck}}$). We selected five PCs based on the minimum BIC and MDL values as a function of the number of PCs. The selected number of PCs was supported by the value of the cumulative variance explained 89(%) when $N_{p_{ck}} = 5$ (see Fig. S.3). We computed a distance matrix $\mathbf{D} \in \mathbb{R}^{K \times K}$ of these PCs, which uses a dissimilarity measure such as distance correlation (Székely, Rizzo, & Bakirov, 2007). Next, we applied hierarchical clustering using Ward's D2 Murtagh and Legendre (2014) method on the distance matrix $\mathbf{D}$ to select the clusters based on the height of the hierarchical tree and the cluster significance. Significant clusters were selected based on the approximately unbiased (AU) probability (Efron, Halloran, & Holmes, 1996) p-values with $p < 0.05$. As a choice of the pairwise distance between the nodes of the tree, we used Ward's D2 criterion (Murtagh & Legendre, 2014) because, like PCA, it is based on multidimensional variance minimization. We evaluated the goodness of fit of the clustering solution using the elbow criterion of the cluster validity indexes that maximize between-cluster distances and minimize within-cluster distances, such as the Dunn Index (DI) Dunn (1974), Calinski-Harabasz index (CH) (Caliński & Harabasz, 1974), and separation index (SI) Qiu and Joe (2006). As shown in Supplementary Fig. S.3a–c, three cluster validation metrics (DI, CH and SI) have an inflection point when the number of clusters $N_{c_k} = 2$. Cluster stability was evaluated using the Jaccard similarity index (J) (Jaccard, 1912) and a nonparametric bootstrap resampling technique with nonoverlapping split-half samples of subject scores and $n = 1000$ repetitions. The reproducibility of the clusters was established with a high degree of cluster similarity for the Jaccard index $J = 0.9$ at $N_{c_k} = 2$ (see Supplementary Fig. S.3d). Thus, the cluster compactness metrics and cluster stability measure provided strong evidence that for a given neurophysiological dataset, the optimal number of clusters is two (see Supplementary Fig. S.3). We obtained the final clustering solution by applying the K-means algorithm to the hierarchical clustering output generated using Ward's D2 method.

3.2. Repeated Measures Correlation Analysis between ERF components and ERF time courses

We conducted repeated measures correlation analysis (Bakdash & Marusich, 2017) between the original subject's ERF waveforms and the extracted subject's ERF components for each stimuli condition to determine the common associations between the original data and the extracted components shared among the subject's datasets. As described by (Bakdash & Marusich, 2017), repeated measures correlation takes into account

nonindependence among observations by statistically adjusting for intersubject variability and thereby provides a higher statistical power than that of the simple average correlation. We used two paired measures (original amplitude as the dependent variable and component amplitude as the independent variable) in the analysis, and the nonparametric bootstrap resampling technique was used to obtain 95% confidence intervals with $n = 1000$ repetitions. We observed significant correlations between all ERF components and the original data from each condition (Supplementary Table S.3; $p < 0.001$). Supplementary Table S.3 shows the common associations according to repeated measures correlation analyses.

The ERF components accounted for 72%, 76%, and 74% of the variance in the VIS, AUD and AV conditions, respectively.

References

Argüelles, M., Benavides, C., Fernández, I. (2014). A new approach to the identification of regional clusters: hierarchical clustering on principal components. *Applied Economics*, *46*(21), 2511–2519.

Bakdash, J.Z., & Marusich, L.R. (2017). Repeated measures correlation. *Frontiers in psychology*, *8*, 456.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, *3*(1), 1–27.

Dunn, J.C. (1974). Well-separated clusters and optimal fuzzy partitions. *Journal of cybernetics*, *4*(1), 95–104.

Efron, B., Halloran, E., Holmes, S. (1996). Bootstrap confidence levels for phylogenetic trees. *Proceedings of the National Academy of Sciences*, *93*(14), 7085–7090.

Husson, F., Josse, J., Pagès, J. (2010). Principal component methods–hierarchical clustering. partitional clustering. *Why would we need to choose for visualizing data*.

Jaccard, P. (1912). The distribution of the flora in the alpine zone. 1. *New phytologist*, *11*(2), 37–50.

Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? *Journal of classification*, *31*(3), 274–295.

Qiu, W., & Joe, H. (2006). Separation index and partial membership for clustering. *Computational statistics & data analysis*, *50*(3), 585–603.

Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, *14*(5), 465–471.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461–464.

Székely, G.J., Rizzo, M.L., Bakirov, N.K. (2007). Measuring and testing dependence by correlation of distances. *The annals of statistics*, *35*(6), 2769–2794.