

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

Irina Belyaeva^{1*}, Ben Gabrielson¹, Yu-Ping Wang², Tony W. Wilson³, Vince D. Calhoun⁴, Julia M. Stephen⁵ and Tülay Adalı¹

^{1*}Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, Maryland, USA.

²Department of Biomedical Engineering, Tulane University, New Orleans, LA, USA.

³The Institute for Human Neuroscience, Boys Town National Research Hospital, Boys Town, NE, 68010, USA.

⁴Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Atlanta, GA, USA.

⁵The Mind Research Network, Lovelace Biomedical Research Institute, Albuquerque, NM, USA.

*Corresponding author(s). E-mail(s): irinbell@umbc.edu;

Contributing authors: bengabr1@umbc.edu; wyp@tulane.edu; tony.wilson@boystown.org; vcalhoun@gsu.edu; jstephen@mrn.org; adali@umbc.edu;

Abstract

Objective: Identification of informative signatures from electrophysiological signals is important for understanding brain developmental patterns, where techniques such as magnetoencephalography (MEG) are particularly useful. However, less attention has been given to fully utilizing the multidimensional nature of MEG data for extracting components that describe these patterns. **Methods:** Tensor factorizations of MEG yield components that encapsulate the data’s multidimensional nature, providing parsimonious models identifying latent brain patterns for meaningful summarization of neural processes. To address the need for meaningful MEG signatures for studies of pediatric cohorts, we propose a tensor-based approach for extracting developmental signatures of multi-subject MEG data. We employ the canonical polyadic (CP) decomposition for estimating latent spatiotemporal components of the data, and use these components for group level statistical inference. **Results:** Using CP decomposition along with hierarchical clustering, we were able to extract typical early and late latency event-related field (ERF) components that were discriminative of high and low performance groups ($p < 0.05$) and significantly correlated with major cognitive domains such as attention, episodic memory, executive function, and language comprehension. **Conclusion:** We demonstrate that tensor-based group level statistical inference of MEG can produce signatures descriptive of the multidimensional MEG data. Furthermore, these features can be used to study group differences in brain patterns and cognitive function of healthy children. **Significance:** We provide an effective tool that may be useful for assessing child developmental status and brain function directly from electrophysiological measurements and facilitate the prospective assessment of cognitive processes.

Keywords: tensor decomposition, canonical polyadic decomposition, MEG, multi-subject analysis, cognitive function, developmental neuroscience

1. Introduction

The characterization and identification of typical brain developmental patterns can provide important insights into brain organization and function. Brain function can be described by cognitive processes, which may include, but are not limited to, sensory, grammatical, semantic processing, memory retrieval, or motor events (Hernández, Puupponen, & Jantunen, 2022). One method of studying the timing properties of cognitive processes is focusing on their underlying brain mechanisms. This can be achieved using functional neuroimaging techniques such as magnetoencephalography (MEG) and electroencephalography (EEG) (Hernández et al., 2022) by extracting neural sources (latent components) describing these processes. MEG is a powerful neuroimaging technique that measures the magnetic fields generated by neuronal activity, which arises collectively within the brain from population neuronal responses to target stimuli (Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993). The direct measurement of neuronal currents by MEG, its high spatial resolution, and its excellent temporal resolution makes it an especially useful noninvasive technique for studying brain function. Thus, MEG is the preferred method for studying the rapid spatiotemporal dynamics of brain activity (He & Liu, 2008). To this end, event-related fields (ERFs) and event-related potentials (ERPs) have been identified as important research tools for understanding brain developmental patterns of pediatric cohorts. ERFs and ERPs represent a time-locked MEG/EEG activity that measures brain responses elicited by stimuli.

Multi-subject MEG studies that focus on large pediatric cohorts have high potential to provide important insights into brain organization and brain development in children and adolescents. One approach to this is by using multi-subject latent component analysis, through which latent components within multiple MEG datasets are learned jointly by exploiting dependence across the datasets (Akhonda, Levin-Schwartz, Bhinge, Calhoun, & Adali, 2018; Gabrielson et al., 2020). However, as we discuss next, many existing multi-subject latent component analysis techniques in neuroimaging are based on matrix factorization methods. Matrix representations cannot account for the multiple dimensions of the data, such

as participant, stimulus condition, variations in time and space, and the relationships across these dimensions.

Over the past decades, substantial efforts have been made in finding ways to model and extract common hemodynamic or electrophysiological components from multi-subject task and resting state neuroimaging (NI) data such as fMRI and MEG/EEG. The common components detected during the task may be indicative of a typical or atypical patient state and can be further used to discover the prognostic imaging biomarkers. Data fusion and joint analysis methods based on matrix decompositions such as joint independent component analysis (jICA) (Calhoun, Adali, Pearlson, & Kiehl, 2006), group ICA (GICA) (Calhoun & Adali, 2012; Calhoun, Adali, Pearlson, & Pekar, 2001; Labounek et al., 2018; Salman et al., 2019), dictionary learning (Akhavan, Baghestani, Kazemi, Karami, & Soltanian-Zadeh, 2022; Jin, Dontaraju, Kim, Akhonda, & Adali, 2020), independent vector analysis and its transposed variant (tIVA) (Adali, Levin-Schwartz, & Calhoun, 2015) have been used for analysis and fusion of NI data. A reason for the popularity of these methods is the convenience of presenting the time-varying NI data as a matrix of time \times space. It was shown in many studies that these matrix-based approaches are powerful for extracting meaningful components (Calhoun & Adali, 2012). The group-level ICA methods exploit high-order statistics (Hyvärinen & Oja, 2000) of the data and enable assessment of complex spatiotemporal relationships (Calhoun & Adali, 2012). However, a primary problem of two-way techniques is that components are defined only by two signatures, which are not determined uniquely without further constraints on the model. The uniqueness is achieved by imposing constraints such as independence or sparsity (Acar, Rasmussen, Savorani, Næs, & Bro, 2013; Adali, Anderson, & Fu, 2014; Akhavan et al., 2022; Jin et al., 2020; Lahat, Adali, & Jutten, 2015). To use matrix-based methods for the higher dimensional data, the unfolding and the dimension reduction into a matrix is required, which is done by concatenation or stacking of the data (Calhoun, Liu, & Adali, 2009; Delorme & Makeig, 2004). Such unfolding inevitably discards the inherent multilinear structure of brain imaging data, and therefore, may ignore complex

important interactions between/among the folded modes (Cong et al., 2015).

Given that most NI data can be conveniently expressed as a high order array, tensor decomposition techniques are preferred to represent the original data as a mixture of the latent components with corresponding signatures from each dimension. In addition, certain tensor decompositions provide uniqueness under mild constraints. The uniqueness property is critical for an unambiguous interpretation of the components, finding matches with neural processes and/or component signatures. Moreover, tensors provide a natural representation of the inherently multidimensional NI data and preserve the structural information among the tensor modes, thus effectively exploiting the multilinear correlation structure and enabling robust group-level statistical analyses for multiple datasets.

Among the tensor decomposition techniques, the canonical polyadic (CP) decomposition and the Tucker decomposition (Carroll & Chang, 1970; Kolda & Bader, 2009) are particularly useful in fMRI, MEG/EEG processing of real (Cong et al., 2015) and complex-valued data (Kuang et al., 2019). The main advantage of the CP model (Carroll & Chang, 1970) is that it is essentially unique up to scaling and permutations (Sidiropoulos & Bro, 2000). However, it is worth noting that the CP model cannot effectively take into account higher-order statistical information like ICA-based methods (Kroonenberg & De Leeuw, 1980). The disadvantage of the Tucker decomposition compared with the CP model is limited model interpretability without imposing the orthogonality constraint, which is unrealistic for the brain components. Thus, in this paper we chose the CP tensor format as our primary model of interest.

Tensor-based analysis of MEG/EEG has received increased attention during the last decade (Chatzichristos et al., 2022; Cong et al., 2012; Liu et al., 2021; D. Wang, Zhu, Ristaniemi, & Cong, 2018; X. Wang, Liu, Toivainen, Ristaniemi, & Cong, 2020; Zhu et al., 2020). The CP model has been extensively used for high-order decompositions of multi-subject EEG data (organized as a channel \times time \times subject third-order tensor) or wavelet EEG (organized as a channel \times time \times frequency \times subject) (Cong et al., 2012; Vanderperren et al., 2013; D. Wang et al.,

2018; X. Wang et al., 2020). The specialized multiway algorithms have been proposed for ERP analysis of EEG to deal with noisy and nonstationary signals using the Bayesian CP model (Wu et al., 2014) and fifth-order ERP feature extraction (D. Wang et al., 2018). In (Kinney-Lang, Ebied, Auyeung, Chin, & Escudero, 2019; Kinney-Lang, Spyrou, Ebied, Chin, & Escudero, 2018), authors employed the CP decomposition for developmental feature extraction from EEG pediatric datasets. In (Liu et al., 2021; Zhu et al., 2020), the CP model was used to study the functional connectivity patterns of MEG data (organized as time \times frequency \times connectivity third-order tensor).

Despite a substantial number of studies dedicated to high-order ERP analysis, the multidimensional nature of MEG has not been fully exploited for the data-driven extraction of sensor-level ERF components. MEG ERF components can better inform about the rapid spatiotemporal dynamics of brain information processing compared with EEG due to higher spatial resolution of MEG. Provided that ERFs are collected using the same stimuli, the assumption is that activity elicited by the same stimuli is highly correlated among subjects, which can be seen as a prerequisite for applying the CP decomposition. Thus, multi-subject MEG studies generate ERFs that can be naturally represented using CP tensor format.

Several works (Ablin, Cardoso, & Gramfort, 2021; Boonyakitanont et al., 2022; Ikeda & Toyama, 2000; Jung et al., 2001; Stephen et al., 2013) focus on the characterization and identification of sensor-level MEG ERFs using matrix-based approaches such as ICA/jICA algorithms. The algorithms that have been used in (Ablin et al., 2021; Boonyakitanont et al., 2022; Ikeda & Toyama, 2000; Jung et al., 2001; Pinner, Coffman, & Stephen, 2020; Stephen et al., 2013) inherently transform three-dimensional (3D) multi-subject MEG data into two-dimensional (2D) matrix representation. For multi-subject MEG data, such a 2D transformation loses the multidimensional low-rank structure that may provide an intrinsic description of the spatiotemporal interactions. On the hand, the low-rank structure of MEG data can be fully captured by the CP tensor format as we propose. Hence, we model the multi-subject MEG data as a 3D tensor with dimensions of subject \times time \times channel. This high-order representation of the multi-subject MEG dataset maximizes the

simultaneous use of spatiotemporal modes and multilinear interactions across modes within the data.

Our goal is the identification of typical brain developmental patterns that could be used as descriptive imaging signatures in a healthy population of children and adolescents. Using CP decomposition, we propose a group-level tensor analysis method to characterize and identify sensor-level ERF components in task-related multi-subject MEG data. The proposed model enables the analysis of a multi-subject MEG dataset as a third-order tensor and, thus, exploits the multidimensional nature of the group-level data. We use hierarchical clustering on principal components (HCPC) (Argüelles, Benavides, & Fernández, 2014; Husson, Josse, & Pagès, 2010) approach to find subject groups using a supplementary cognitive measures dataset.

We summarize our contributions as follows: The paper presents a CP analysis framework to robustly identify common brain developmental patterns for multi-subject sensor-level MEG data. The proposed formulation of the CP model shown in Fig. 3 was capable of extracting typical early (M150), and late latency (M300a and M400) ERF components representative of visual spatial attention, associative memory and semantic processing, similar to the results reported in existing ERF/ERP studies. We develop a group-level inference approach that allows robust statistical inferences directly using CP component matrices. We demonstrate the statistical significance of tensor group-level analyses by identifying the discriminative ERF components that can differentiate between high performance and low performance groups. We show that the discriminative ERF components were significantly correlated with major cognitive domains such as attention, episodic memory, executive function, and language comprehension.

A preliminary work using the same MEG data with similar clustering of subjects used ICA model and is presented as a conference contribution (Boonyakitanont et al., 2022). The current paper presents a novel formulation for group-level analysis using the CP model, a detailed description of the clustering approach (see Section 5.2), and novel experimental results.

This paper is organized as follows. The description and notations for CP decomposition are

introduced in Section 2.1. We describe multidimensional generative data model and CP tensor decomposition for multi-subject MEG data in Section 3. The experimental setup is described in Section 4.2. In Section 5, the typical ERF components extracted from the CP model and the group-level tensor-based statistical inference results are presented. The experimental results are discussed in Section 6. The conclusions and future work are presented in Section 8.

2. Materials and Methods

2.1. Notations and Definitions

In this paper, the mathematical notations and definitions are adopted from (Kolda & Bader, 2009) and (Cichocki et al., 2016). We denote scalar with the lower case letter x , vectors with boldface lowercase letters ($\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$), matrices with boldface capital letters ($\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \dots$), and tensors with bold calligraphic uppercase letters ($\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \dots$). The number of dimensions is called the *order*, and each dimension is referred to as a *mode*. $\|\cdot\|_F$ denotes the Frobenius norm, $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker, $\mathbf{A} \odot \mathbf{B}$ denotes the Khatri-Rao product and $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \mathbf{b}$ denotes the inner product of two vectors. A rank-1 tensor is expressed as the outer product of vectors, i.e., $\mathcal{X} = \mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$, where \circ represents the vector outer product. The mode- n matricization of a given tensor along dimension n is denoted by $\mathbf{X}_{(n)} \in \mathbb{R}^{I_n \times I_1 I_2 \dots I_{n-1} I_{n+1} \dots I_N}$ (Kolda & Bader, 2009). The n -mode product of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \dots \times I_n \times I_N}$ and a matrix $\mathbf{A} \in \mathbb{R}^{J_n \times I_n}$ along the n th mode, denoted as $\mathcal{X} \times_n \mathbf{A}$, is a tensor of size $I_1 \times \dots \times J_n \times \dots \times I_N$.

2.2. Participants

The participants included 170 healthy children (89 male, 81 female) and adolescents between the ages of 9 and 15 ($M = 11.92$ years, $SD = 1.18$) with no reported clinical diagnoses from the Mind Research Network (MRN) in Albuquerque, New Mexico (90) and the University of Nebraska Medical Center (UNMC) in Omaha, Nebraska (80) as part of the Developmental Chronnecto-Genomics (Dev-CoG) study (Stephen et al., 2021). The participants and parents signed consent forms approved by each institutional review board (IRB)

prior to joining the study. All procedures were approved by the MRN and UNMC IRBs prior to the start of the experiment.

2.3. Neuropsychological Testing

All the participants completed the Wechsler abbreviated scale of intelligence (Second Edition; WASI-II; (Wechsler, 2011)) to assess full-scale IQ (FSIQ) and NIH-Toolbox Cognitive Battery (Weintraub et al., 2013) (NIHTB-CB) tests assessing age-adjusted neuropsychological (T) scores in six cognitive domains: attention, episodic memory, executive function, language, processing speed, and working memory. The data collection also included the Conners 3 Inattention/Hyperactivity scores (Conners, 2008) for assessing attention-deficit hyperactivity disorder (ADHD), and children with diagnosed ADHD were excluded from the study. There were no significant differences ($p > 0.05$) in terms of age or gender with respect to the neuropsychological measures in participants from the MRN and UNMC.

2.4. MEG Experimental Paradigm

Participants completed a multisensory task while MEG data were recorded (see Fig. 1). The visual stimulus was a full-screen, black and white vertical grating (0.25 cycles/degree). The auditory stimulus was a 40 Hz modulated 1000 Hz tone. For multisensory stimulus, the auditory and visual stimuli were presented simultaneously. The baseline fixation was a red box in the center of the screen. Subjects were instructed to press their index finger when they saw anything, heard anything, or both. Each MEG trial began with a fixation for an inter-trial interval (ITI) that pseudo-randomly changed between 2400 and 2600 milliseconds (ms) in 10 ms increments. Following fixation, a sensory stimuli (auditory (AUD), visual (VIS), or audio-visual (AV)) was presented for 800 ms (Stephen et al., 2021). The total task duration was approximately 18 min.

2.5. MEG Data Acquisition and Image Preprocessing

The MEG data acquisition and preprocessing details were previously published in (Stephen et al., 2021). MEG recordings were acquired with

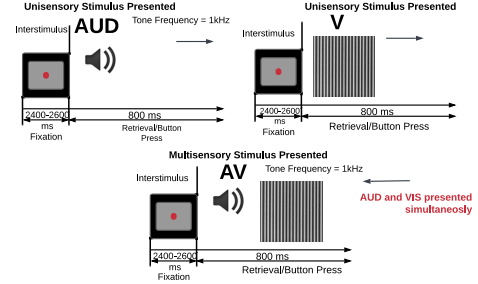


Fig. 1 Multisensory task paradigm. The presentation started with a baseline fixation screen, followed by the appearance of auditory (AUD), visual (VIS) or multisensory stimulus (AV).

the Elekta/MEGIN MEG system with 306 magnetic sensors (204 gradiometers and 102 magnetometers) in a magnetically shielded room. The MEG data were continuously sampled at 1000 kHz with a passband between 0.1 and 330 Hz. We used preprocessing technique such as signal-space separation (SSS) (Taulu & Kajola, 2005) for MEG data denoising and to ensure comparability between magnetometer and gradiometer source reconstructions (Garcés, López-Sanz, Maestú, & Pereda, 2017). The MEG sensor-level artifacts were removed during preprocessing at both the MRN and UNMC sites. MEG epochs between -100 to 1000 ms (1100 time points) around the stimulus onset were averaged across 300 trials within respective stimuli and formed sensor-level ERFs time-locked to the stimulus condition (AUD, VIS, or AV).

Prior to MEG recording 3D digitization was performed to collect positioning data for four head-position indicator (HPI) coils, and the scalp surface. The HPI coils data were collected throughout the recordings, which allowed offline head movement correction (Stephen et al., 2021). The Maxfilter program was used to adjust the location of the head to a common head location within the dewar. The movement compensation extension of the Maxfilter program allows one to correct for head movement throughout the scan (effectively correcting small changes in head position through re-mapping the MEG data to a constant head) (Taulu & Simola, 2006). Another use of this capability is to map each subject's MEG data to a common head position within the dewar. Prior work has shown that too much adjustment of the head position can lead to noise amplification. Therefore, we chose a head position

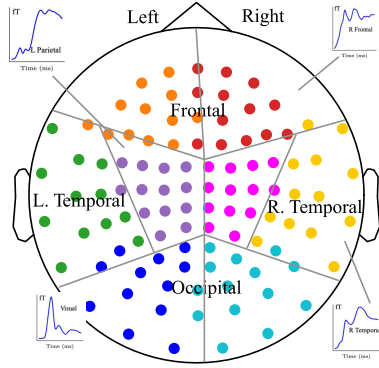


Fig. 2 Top down view of MEG magnetometers. MEG regional division based on the sensor's spatial adjacency matrix. Inserts show global field power of MEG. The approximate sensor groups were used to describe the ERF components spatial patterns on the scalp topographic map. R - right, L - left. Adopted from (Stephen et al., 2013).

that was the closest to the average participant head location within the dewar and mapped all data to this common head position. Once the data were mapped to a common head position, the results were compared across participants as is often done with EEG sensor data.

We did not perform source reconstruction and worked in sensor space when we applied tensor decomposition. Using the sensor's spatial adjacency matrix, we associated each sensor with a sensor spatial region (Occipital, Frontal, Parietal, Temporal, and left/right hemisphere). These sensor regions shown in Fig. 2 do not correspond to the brain regions or anatomical labels. Throughout the paper, the approximate sensor groups were used to describe ERF component spatial patterns on the scalp topographic map (topomap).

3. Tensor Analysis of MEG Data for Brain Pattern Extraction

3.1. Multidimensional Model for Multi-Subject MEG Data

The MEG experimental paradigm shown in Fig. 1 results in simultaneously recorded neural measurements elicited in C common sensors at T timepoints across K subjects. As a result, the observed MEG recordings are modeled as a mixture of the underlying neural sources of interest

synchronized in time across subjects within a specific task. To identify the common brain developmental patterns elicited by sensory stimuli across subjects, we applied the CP tensor decomposition to extract the latent brain activity patterns. The proposed approach has two important advantages. The CP representation of MEG data allows us to take into account the higher-order structure of the multi-subject data to extract common patterns across subgroups. By virtue of the CP decomposition, the MEG factorization provides a unique solution under mild constraints (Kruskal, 1977; Sidiropoulos & Bro, 2000). The importance of the uniqueness condition cannot be overstated since it allows finding meaningful components unambiguously and matching them to the true brain processes.

To preserve the intrinsic multidimensional nature of multichannel MEG data, the data are tensorized as third-order tensor $\mathcal{X} \in \mathbb{R}^{K \times T \times C}$ by stacking subject ERF matrices $\mathbf{S}_k \in \mathbb{R}^{C \times T}$ in subject mode. Fig. 3a shows the generative model for multidimensional representation of multi-subject MEG data.

3.2. Multi-Subject MEG Tensor Decomposition

By adopting the tensorization strategy shown in Fig. 3a, we present the CP model of the multi-subject MEG data as a third-order tensor $\mathcal{X} \in \mathbb{R}^{K \times T \times C}$ ($\mathbb{R}^{\text{subject} \times \text{time} \times \text{channel}}$). The CP decomposition approximates tensor $\mathcal{X} \in \mathbb{R}^{K \times T \times C}$ as a sum of rank-1 tensors:

$$\mathcal{X} \approx \sum_{r=1}^R \lambda_r \circ \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r = \mathbf{A} \times_1 \mathbf{B} \times_2 \mathbf{C} \quad (1)$$

where $\mathbf{a}_r \in \mathbb{R}^K$, $\mathbf{b}_r \in \mathbb{R}^T$, $\mathbf{c}_r \in \mathbb{R}^C$ are the factor vectors normalized to the 2-unit norm; λ_r represents the scale factor for each component, and the norms absorbed into diagonal matrix \mathbf{A} ; $\mathbf{A} \in \mathbb{R}^{K \times R}$, $\mathbf{B} \in \mathbb{R}^{T \times R}$, $\mathbf{C} \in \mathbb{R}^{C \times R}$ are the factor matrices and R is the rank or number of components. Each rank-1 tensor obtained from the $\lambda_r \circ \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$ decomposition, can be interpreted as a distinct spatiotemporal brain pattern, where \mathbf{a}_r , \mathbf{b}_r and \mathbf{c}_r are the subject weights of the time-varying spatial patterns, timecourses, and spatial maps, respectively, as illustrated in Fig. 3b. The

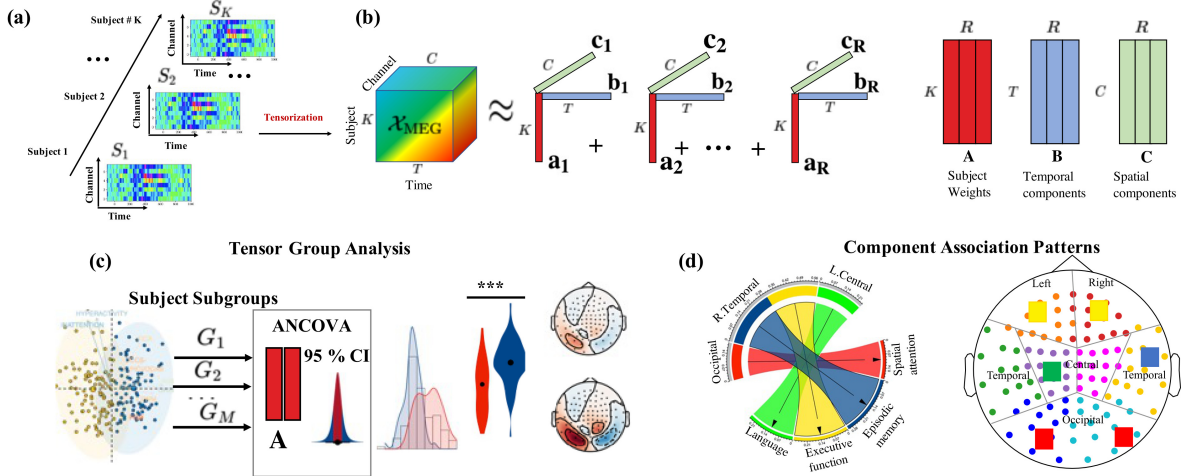


Fig. 3 Brain developmental patterns discovery via tensor decomposition. (a)-(b) Illustration of tensor analysis for multi-subject MEG data. (a) Tensor formation by arranging the subjects along the first dimension. (b) MEG tensor decomposition into R rank-1 components. Each rank-1 component represents a distinct spatiotemporal brain activity pattern with subject weight (\mathbf{a}_r) and temporal (\mathbf{b}_r) and spatial signatures (\mathbf{c}_r). (c) Tensor group-level analysis, which includes subgroup identification and group-level statistical inference. (d) Left: Component association patterns with cognitive domains. Right: Sensor spatial locations associated with the components.

CP model optimizes a least-squares fit of the following cost function (Kolda & Bader, 2009):

$$\begin{aligned}
 f(\Lambda, \mathbf{A}, \mathbf{B}, \mathbf{C}) = & \\
 \min_{\Lambda, \mathbf{A}, \mathbf{B}, \mathbf{C}} & \frac{1}{2} \|\mathcal{X} - \Lambda \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}\|_F^2, \\
 \text{s.t. } & \|\mathbf{a}_r\|_2 = \|\mathbf{b}_r\|_2 = \|\mathbf{c}_r\|_2 = 1, \forall r = 1, \dots, R.
 \end{aligned} \quad (2)$$

We apply alternating least squares (ALS) to estimate the factor matrices (Cichocki et al., 2016; Kolda & Bader, 2009). The minimization problem is solved by fixing two matrices and optimizing over the third-one. Each least squares subproblem is convex and has a closed-form solution (Kolda & Bader, 2009).

3.3. Component Number Estimation

As in many dimensionality reduction methods, a critical step is the selection of the optimal number of components. We use three methods to make this choice for the CP decomposition: the core consistency diagnostic (CORCONDIA/CCD) (Bro & Kiers, 2003), the average congruence product (ACP) (Tomasi & Bro, 2005) and the Bayesian information criterion (BIC) (Schwarz, 1978) as a function of tensor rank R . The CCD measures the similarity between the estimated core and the superdiagonal ideal core, in the absence of noise

(Bro & Kiers, 2003). According to (1), the CP core can be estimated as

$$\mathcal{G} = \mathcal{X} \times_1 \times_2 \mathbf{A}^\dagger \times_3 \mathbf{B}^\dagger \times_3 \mathbf{C}^\dagger. \quad (3)$$

The CCD in (%) is defined as in (Bro & Kiers, 2003)

$$\text{CCD}(\%) = 100 \times \left(1 - \frac{\|\mathcal{G}_R - \mathcal{I}_R\|_F^2}{R}\right), \quad (4)$$

where $\mathcal{G} \times \mathbb{R}^{R \times R \times \dots \times R}$ and $\mathcal{I} \times \mathbb{R}^{R \times R \times \dots \times R}$ are the estimated and ideal CP cores, respectively. We choose the model with the highest number of components such that

$$\hat{R}_{\text{CCD}} = \arg \max_r (\text{CCD}) \quad \text{s.t. } \text{CCD}(r) \geq \eta, \quad (5)$$

where $0 < \eta < 100\%$ is the threshold coefficient, with $r = 1, \dots, R$. Typically, $80\% < \eta < 90\%$ is used.

Furthermore, to assess the number of components, we computed the ACP measure of all fitted models for a given tensor rank R . The ACP metric measures the correlation between components extracted from different models for a given tensor rank R :

$$\text{ACP} = \max_{\mathbf{P}} \text{tr} \left((\mathbf{A}_r^{(1)T} \mathbf{A}_r^{(2)}) \right)$$

$$(\mathbf{B}_r^{(1)T} \mathbf{B}_r^{(2)})(\mathbf{C}_r^{(1)T} \mathbf{C}_r^{(2)})\mathbf{P}), \quad (6)$$

where $\mathbf{A}_r^{(k)}, \mathbf{B}_r^{(k)}, \mathbf{C}_r^{(k)}$ represents the r th component of the i th solution $i = 1, \dots, I$, $r = 1, \dots, R$, \mathbf{P} is the permutation matrix that accounts for the ambiguity (Harshman et al., 1970) of ordering the solutions, and $\text{tr}(\cdot)$ is the trace of the matrix. We select the model that produces the highest ACP value such that

$$\hat{R}_{\text{ACP}} = \arg \max_r (\text{ACP}), \quad r = 1, \dots, R. \quad (7)$$

We used the BIC metric further to assess the number of components as an information-theoretical criterion. The BIC measure is based on the negative log-likelihood and the maximum a posteriori (MAP) approximation (Stoica & Selen, 2004). The BIC metric is defined in terms of the sum squared error ($\text{SSE} = \|\mathcal{X} - \hat{\mathcal{X}}\|_F^2$) (Mørup & Hansen, 2009), where \mathcal{X} stands for the original data tensor, and $\hat{\mathcal{X}}$ denotes the fitted model

$$\text{BIC} = S \log \frac{\text{SSE}}{S} + F \log S, \quad (8)$$

F is the degree of freedom, and $S = \prod_{n=1}^N I_n$ is the number of tensor data elements. We chose the model that produces the lowest BIC value such that

$$\hat{R}_{\text{BIC}} = \arg \min_r (\text{BIC}), \quad r = 1, \dots, R. \quad (9)$$

3.4. Clustering Analysis for Subject Subgroup Identification

In this section, we present the clustering analysis methodology for identifying subgroups. A preliminary version of the clustering approach presented here using cognitive measures collected during the Dev-Cog study presented as a conference contribution (Boonyakitanont et al., 2022). The detailed clustering protocol is described in the Supplementary Methods Section 3.1. We partitioned the subject cohort ($N = 170$) into distinct subgroups using a neuropsychological dataset. We performed HCPC clustering using nine cognitive variables from six cognitive domains, including the Connors 3 inattention/hyperactivity scores: WASI-II FSIQ, ORRENG, PICVOCAB, PSM, LSWM, DCSS, FICA, INATTENTION and HYPERACTIVITY.

The HCPC method (Argüelles et al., 2014; Husson et al., 2010) combines three standard techniques (principal component analysis (PCA), hierarchical clustering, and the K -means algorithm) to obtain a higher quality clustering solution. A schematic view of the subgroup identification using the HCPC algorithm is presented in Fig. 4.

First, the PCA algorithm is applied to the neuropsychological 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 called principal components (PCs), which are uncorrelated with each other. We compute a distance matrix $\mathbf{D} \in \mathbb{R}^{K \times K}$ of these PCs, which uses the dissimilarity measure such as distance correlation (Székely, Rizzo, & Bakirov, 2007). The distance correlation measure allows the detection of nonlinear correlations (Székely et al., 2007) that might not be identified by the Pearson correlation (Székely et al., 2007), which may result in suboptimal performance of the downstream tasks. Next, we apply hierarchical clustering using Ward's D2 (Murtagh & Legendre, 2014) method on the distance matrix \mathbf{D} to select the clusters based on the height of the hierarchical tree. The significant clusters are selected on the basis of the approximately unbiased (AU) probability (Efron, Halloran, & Holmes, 1996) p -values with $p < 0.05$. The quality of clustering is assessed according to the compactness metrics (Halkidi, Batistakis, & Vazirgiannis, 2002a, 2002b) (see Supplementary Methods Section 3.1 and Supplementary Fig. S.3). The cluster stability is evaluated as a function of the number of clusters using the Jaccard similarity index (J) (Jaccard, 1912) via a nonparametric bootstrap technique with a number of repetitions $n = 1000$ (Supplementary Methods Section 3.1). The final clustering solution is obtained by applying the K -means algorithm to the hierarchical clustering output.

4. Numerical Experiments

4.1. Data Preprocessing

The MEG multi-subject dataset consists of 170 subjects taken from the Dev-CoG study (Stephen et al., 2021). Before the tensor analysis, we normalized the data by centering the third-order MEG tensor across the time mode, and scaling within the subject mode by its standard deviation

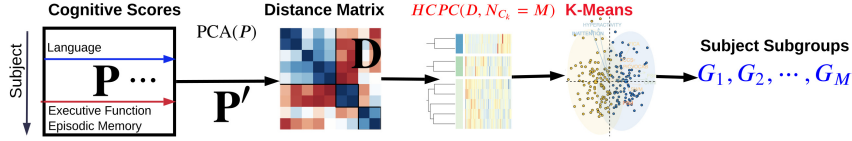


Fig. 4 Hierarchical clustering on principal components of neuropsychological (T) scores for subject group identification using Ward’s D2 criterion (Murtagh & Legendre, 2014). PCA is used on the subject cognitive matrix $\mathbf{P} \in \mathbb{R}^{K \times L}$ to remove highly correlated continuous variables. Next, we apply hierarchical clustering using Ward’s D2 method on the distance matrix \mathbf{D} to select the clusters based on the height of the hierarchical tree. The distance matrix $\mathbf{D} \in \mathbb{R}^{K \times K}$ is computed using the dissimilarity measure such as the distance correlation (Székely et al., 2007) of the PCs. The initial number of clusters N_{C_k} is assessed according to the compactness metrics (Halkidi et al., 2002a, 2002b), and the cluster stability is evaluated using the Jaccard similarity index (Jaccard, 1912) via a nonparametric bootstrap technique with a number of repetitions $n = 1000$ (see detailed protocol in Supplementary Methods Section 3.1). We select significant clusters based on the approximately unbiased probability p -values (Efron et al., 1996), as shown in Fig. 8a. We provide the final clustering solution by applying the K -means algorithm to the hierarchical clustering output.

(Bro & Smilde, 2003). We used 204 planar gradiometers and 102 magnetometers out of available 306 sensors after SSS preprocessing (see Section 2.5). Thus, the data preprocessing resulted in $170 C \times T$ ERF subject datasets with $C = 306$, and $T = 1100$. Furthermore, tensor analyses were performed separately for three stimulus conditions (AUD, VIS, and AV). We selected nine age-adjusted cognitive (T) scores from available neuropsychological measures (see Section 2.3) in the data analyses: the WASI-II FSIQ, Picture Sequence Memory (PSM) (T) score (Weintraub et al., 2013), Picture Vocabulary (PICVOCAB) (T) score (Weintraub et al., 2013), Oral Reading Recognition (ORRENG) (T) score (Weintraub et al., 2013), List Sorting Working Memory (LSWM) (T) score (Weintraub et al., 2013), Flanker Inhibitory Control and Attention (FICA) (T) score (Weintraub et al., 2013), Dimensional Card Sorting (DCCS) (T) score (Weintraub et al., 2013), and the Conners 3 Inattention/Hyperactivity scores. The neuropsychological (T) scores were aggregated to construct a cognitive score matrix $\mathbf{P} \in \mathbb{R}^{K \times L}$, where K is the number of subjects and L is the number of cognitive tests. Prior to clustering, the matrix was standardized by the z -score to account for scale differences. In addition to the neuropsychological (T) scores, we used parental socioeconomic status (SES), age and gender as model covariates (see the detailed protocol in Section 4.5.1).

4.2. Experimental Design

The goal of this study was to estimate common imaging patterns representing typical brain development in healthy children and adolescents.

Three MEG data tensors were constructed (\mathcal{X}_{VIS} , \mathcal{X}_{AUD} , and \mathcal{X}_{AV}) for each stimulus condition according to the generative model shown in Fig. 3a. The tensor rank R was estimated for each original data tensor as described in Section 3.3. Three separate CP decompositions were conducted for each stimulus condition with the chosen tensor rank. The fitted CP models resulted in three estimated tensors, consisting of R -component factor matrices $\mathbf{A} \in \mathbb{R}^{K \times R}$, $\mathbf{B} \in \mathbb{R}^{T \times R}$, $\mathbf{C} \in \mathbb{R}^{C \times R}$ that described the latent ERF spatiotemporal brain patterns.

To associate brain function with the cognitive performance observed in the neuropsychological tests, we partitioned the subject dataset into two distinct subgroups, high performance (HP) and low performance (LP), using the HCPC method (Argüelles et al., 2014; Husson et al., 2010). Thus, we could perform group-level statistical analyses using the extracted ERF components to identify group-level discriminative brain developmental signatures. We identify associations between extracted latent ERF components and cognitive processes by correlating these latent components with children’s scores in the cognitive domains. We hypothesize that statistically significant latent ERF components can differentiate between children’s brain patterns in those with low vs. high performance and could indicate brain developmental trajectory or cognitive development status. Fig. 3 illustrates the application of tensor decomposition to identify brain developmental patterns using MEG data.

4.3. Execution Details

The CP model (2) was fit using CP-ALS (Kolda & Bader, 2009) from the TensorLy toolbox (Kossaifi, Panagakis, Anandkumar, & Pantic, 2019), MNE-Python (Gramfort et al., 2013) was used to generate topographic maps, and R software (Team et al., 2013) version 3.6.0 (R Foundation for Statistical Computing, Vienna, Austria) was used for statistical analyses. All experiments were performed on a Linux workstation with 4 Quad-Core Intel Xeon 3.1 processors and 16 GB memory.

4.4. Model Selection and Evaluation

The model performance was assessed with qualitative and quantitative metrics. The qualitative assessment used interpretations of the extracted components and comparisons with existing literature findings on adolescent cohorts. We computed the reconstruction error of the CP model as $RMSE = \|\mathbf{X} - \hat{\mathbf{X}}\| / \prod_{n=1}^N I_n$, and the model fit $FIT = (1 - \|\mathbf{X} - \hat{\mathbf{X}}\|_F^2 / \|\hat{\mathbf{X}}\|_F^2)$. The CP-ALS stopping criteria included reaching 1000 iterations or achieving a convergence tolerance of $\epsilon \leq 10^{-8}$. We investigated the model order and stability by running the CP-ALS algorithm 100 times for each stimulus condition and R values of one to ten, with each run randomly initialized. This procedure allowed us to determine whether some runs converged to local minima with high reconstruction error. The error plot in Fig. 5c reveals that all runs at fixed R yielded the same RMSE with a standard error of the mean (SEM) < 0.0001 . These results suggest that all CP-ALS local minima are similar and presumably also similar to the global minimum.

We assessed the number of components for the CP model (2) by generating average CCD (4) plots, average APC (6) plots and average BIC (8) plots as a function of tensor rank R for $R = 1, \dots, 10$. Figs. 5a–b and Fig. 5d show boxplots of the mean CCD (4), mean APC (6) and mean BIC (8) metrics for each stimulus, demonstrating the sensitivity of the solution to the selection of R and initialization parameters of the CP-ALS algorithm.

According to (Bro & Kiers, 2003), the tensor rank of the CP model should be chosen such that the CCD value is greater than 90%. Fig. 5a reveals that $R = 2$ should be chosen for the AUD ($M =$

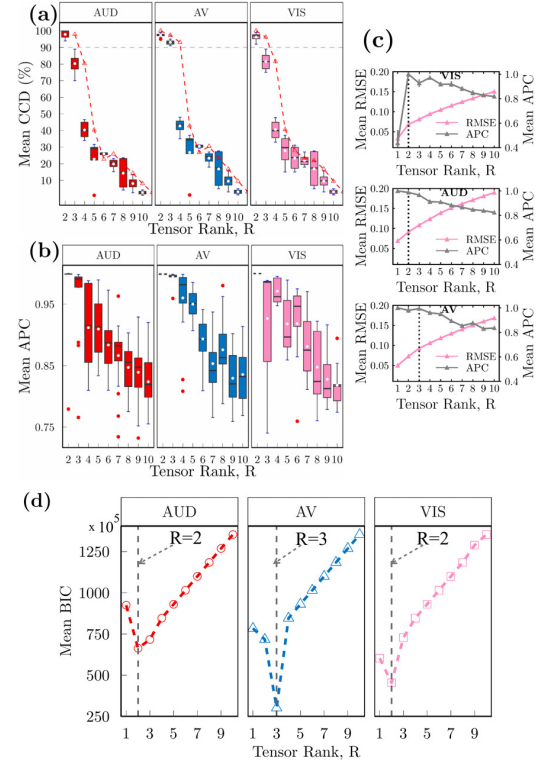


Fig. 5 Estimation of the number of components for the CP model, showing the mean values ($N = 100$) of the CCD, APC, BIC and RMSE metrics as a function of tensor rank R for 100 random initializations for each stimulus condition (VIS, AUD, and AV). (a)–(b) Boxplots summarize the distribution of the mean CCD and mean APC as a function of tensor rank. Median values are represented by black lines inside the boxplot, with the top of the whisker lines indicating the 25th and 75th percentile values. Mean values are plotted in white circles, and red circles represent outliers. Error bars represent the standard error of the mean. The plot of mean values of RSME suggests that all runs at fixed R yielded the same RMSE with a standard error of the mean < 0.0001 . These results suggest that all CP-ALS local minima are similar and presumably also similar to the global minimum.

(a) CCD boxplot. (b) APC boxplot. (c) Mean and standard error of the RMSE and APC as a function of tensor rank R . (d) Mean BIC as function of tensor rank R .

97.9, $SD = 2.28$) and VIS ($M = 96.2$, $SD = 2.25$) conditions, while $R = 3$ should be chosen for the AV ($M = 93.1$, $SD = 1.37$) condition.

The APC values for different R was another method for assessing the number of components. Fig. 5b shows that adding more components resulted in lower mean and higher SEM values for the APC metric. Similar to the CCD boxplot, the APC boxplot confirms that $R = 2$ is the correct number of components for the AUD ($M =$

0.988, $SD = 0.05$) and VIS ($M = 0.998$, $SD = 0.02$) conditions, while $R = 3$ is the best for the AV ($M = 0.988$, $SD = 0.05$) condition.

The BIC (8) method was used as model-driven measure to complement the CCD and ACP metrics for assessing the number of components for different R . Fig. 5d shows that for the AUD ($M = 4.53 \times 10^5$, $SD = 1015$) and VIS ($M = 6.63 \times 10^5$, $SD = 930$) conditions $R = 2$ and for the AV ($M = 3.01 \times 10^5$, $SD = 845$) $R = 3$ should be chosen based on the minimum BIC value. As shown in Fig. 5d, the BIC criterion demonstrates the agreement in terms of the number of components with the CCD and ACP measures.

The final solution was selected based on the chosen R , which produced the minimum RMSE value, maximum CCD and ACP values, and minimum BIC value.

4.5. Statistical Analysis

We quantified the CP model performance to produce latent factors for differentiating subject subgroups using mixed measures analysis of covariance (ANCOVA). We performed post hoc analyses with two-tailed parametric t -tests and corrections for multiple comparisons using the false discovery rate (FDR) (Benjamini, Drai, Elmer, Kafkafi, & Golani, 2001) with the significance level of $\alpha = 0.05$ to determine statistical significance. The ANCOVA and post hoc analyses results were accompanied with F -statistics, t -statistics, p -values and effect size. The effect size was evaluated by generalized eta squared (η_G^2) (Olejnik & Algina, 2003), Cohen's d values and characterized as small (< 0.06), medium (0.06–0.14), or large (> 0.14), according to (Cohen, 2013). Additionally, we reported the mean (M), and standard deviation (SD) of the measures of interest.

4.5.1. Group-Level Statistical Inference of CP Component Matrices

The columns of the factor matrices $\mathbf{A} \in \mathbb{R}^{K \times R}$ in subject mode contain the component loading factors (coefficients), with the column index corresponding to the loading factors of the given R th component. The loading factors of each component indicate how much of the component is required to reconstruct the subject's source data (Acar, Schenker, Levin-Schwartz, Calhoun,

& Adali, 2019). A higher subject loading factor signifies an increased contribution of that component (Stephen et al., 2013). Therefore, group-discriminative components can be obtained by statistically comparing the component loading factors of subgroups to determine significant between-group differences.

Group differences in the component loading factors were assessed with 2×2 mixed measure ANCOVAs with the stimuli condition (AUD, VIS, or AV) as a within-subject factor and subgroup (HP vs. LP) as a between-subject factor. The ANCOVAs were calculated for each component and condition while controlling for age, gender and parental SES. In addition to the ANCOVA tests, planned direct comparisons between HP and LP groups were made for each component and condition separately to determine if subgroups differed significantly in the component loading factors of any specific stimulus condition while controlling for the same covariates. We applied a two-tailed level of significance ($p < 0.05$) and an FDR correction for the number of tests performed for each condition.

Prior to performing group-level statistical analyses, we examined group differences in the subject head motion between subgroups to determine if it may cause differences in component loading factors. We assessed group differences in the head motion values with one-way ANCOVA with the subgroup (HP vs. LP) as a between-subject factor while controlling for age. There was no significant difference in the mean values of the head motion ($F_{1,167} = -1.051$, $p > 0.05$) between subgroups. Post hoc independent samples two-tailed t -test (FDR corrected, $p < 0.05$) with unequal variances correction confirmed there was a no significant difference in the head motion for HP ($M = 0.922$, $SD = 0.807$) and LP ($M = 1.12$, $SD = 1.036$) groups ($t_{167} = -1.051$, $p = 0.295$). These results suggest that the head motion would not impact results of the group-level statistical analyses. The summary of the head motion statistical analysis is presented in Supplementary Fig. S.4.

4.5.2. Correlation Analysis between Component Loading Factors and Neuropsychological (T) Scores

To identify the specific neuropsychological scores associated with the ERF components, we separately correlated the ERF components with the neuropsychological (T) scores in the HP group, LP group and full sample. The relationships were evaluated with Pearson’s correlation tests. Partial correlation analyses (controlling for age, gender and parental SES) were performed between the component loading factors in subject mode (columns of matrix **A**) and neuropsychological age-adjusted (T) scores. Specifically, we computed the two-tailed Pearson’s partial correlation coefficient (r) between the component loading factors and the nine cognitive variables, namely, WASI-II FSIQ, PSM, PICVOCAB, ORRENG, LSWM, FICA, DCCS, and the Conners 3 inattention/hyperactivity scores. Partial correlations were considered significant below the FDR-corrected threshold ($p < 0.05$, $N = 170$).

5. Results

The proposed model was used to extract MEG ERF components using CP decomposition, followed by statistical group-level analyses (see Sections 4.5.1 and 4.5.2). In this section, we describe the results of the multi-subject tensor analyses for extracting typical brain developmental patterns from the original MEG data.

5.1. Multi-Subject MEG Tensor Analysis Using the CP Model

We employed CP decomposition to determine the component structure of MEG ERF responses to multisensory task. The MEG tensor $\mathcal{X} \in \mathbb{R}^{K \times T \times C}$ was decomposed using the CP factorization model (2) shown in Fig. 3b. The data tensor was fitted with the number of components R as determined in Section 4.4. The average model FIT indicates (VIS: $R = 2$, $M = 0.93$, $SD = 0.01$; AUD: $R = 2$, $M = 0.91$, $SD = 0.01$; AV: $R = 3$, $M = 0.91$, $SD = 0.01$; see experimental setup in Section 4.4) that the extracted factors account for a large part of the explained variance of the original datasets. Supplementary Fig. S.5 and Supplementary Table S.2 show the mean

Table 1 Summary of MEG ERF components

Stimulus condition	ERF Component		
	Occipital/M150	Right Temporal/M300a	Late Central/M400
VIS	✓	-	✓
AUD	-	✓	✓
AV	✓	✓	✓

values of FIT metric of the fitted CP decomposition for each stimulus condition (VIS, AUD, and AV). To quantify the common associations between the original MEG ERF subject’s datasets and the extracted ERF components, we performed repeated measures correlation analyses (Bakdash & Marusich, 2017) between these paired datasets (see Supplementary Methods Section 3.2).

The tensor analysis yielded seven ERF components (see Table 1), which describe patterns of temporal variance (temporal factors), spatial variance (spatial factors), and the subject factors. The spatial loadings provide the measure of activity in the MEG ERF as a function of time for each spatial factor. The subject loadings modulate the magnitude of these spatiotemporal patterns, representing the pattern’s activation strength for the specific subject. We categorized the components as functional MEG ERF components that correspond to prominent spatiotemporal peaks (Stephen et al., 2013), and the spatiotemporal variance explained (R^2), which was determined by the repeated measures correlation analyses. Fig. 6 depicts the extracted temporal and spatial components time-locked to the target stimuli after CP tensor decomposition on the sensor-level MEG data (magnetometer view), and the ERF components gradiometer view is presented in Supplementary Fig. S.6. The temporal ERF components generated from individual sensor data averaged across subject ERF components are shown. The MEG topographic maps show the density of spatial patterns that correspond to prominent time peaks. The average ERF component (average across sensors, in cyan) and average ERF time-course (in yellow) for each stimulus condition are plotted. The ERF components are well-matched to distinct peaks present on the average ERF time-courses. The temporal evolution of the MEG ERF topographic maps is shown in Supplementary Fig. S7.

The repeated measures correlation analyses (see Supplementary Methods Section 3.2) found significant correlations ($p < 0.001$) between ERF

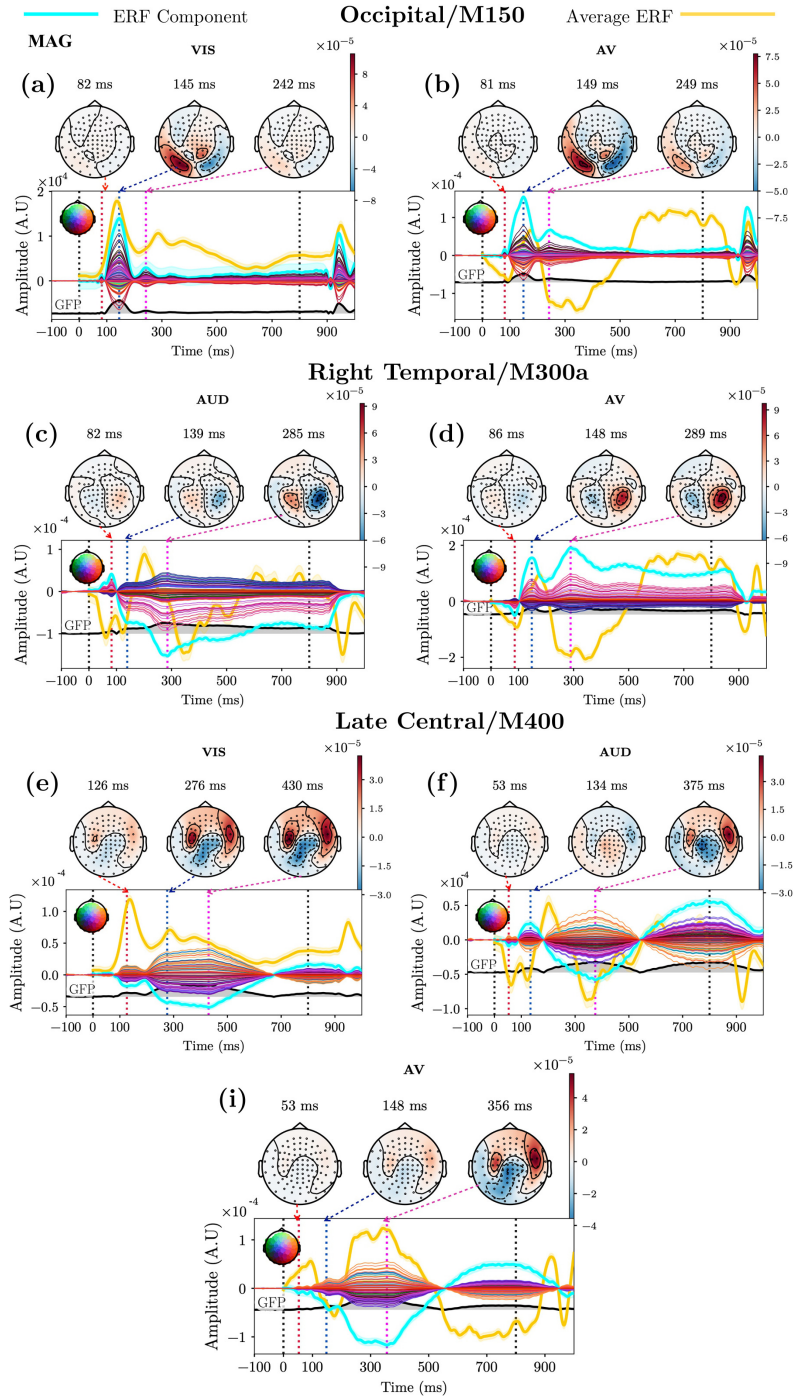


Fig. 6 Tensor decomposition results of sensor-level MEG data for target stimuli. Temporal and spatial patterns from the components of the CP model. Top: The topographic maps (magnetometers view) show the density of spatial patterns that correspond to prominent time peaks denoted with red and blue arrows. 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 AV conditions. (e)–(i) Late central 350–430 ms component in the VIS, AUD, and AV conditions.

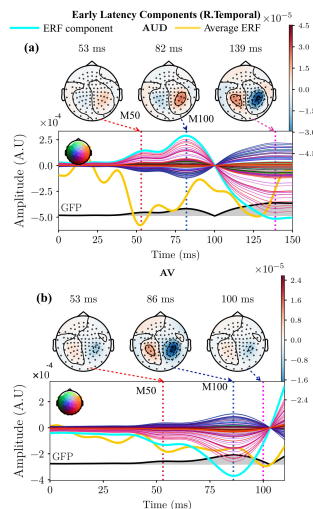


Fig. 7 Early latency M50 and M100 subcomponents within 0–150 ms time window of the right temporal 280–300 ms ERF component for the AUD (Fig. 6c) and AV (Fig. 6d) conditions are shown. (a) AUD subcomponents. (b) AV subcomponents.

components and the original data (VIS, AUD, and AV conditions) for the overall common slope (Supplementary Table S.3). The ERF components (Supplementary Table S.3) accounted for 72%, 76%, and 74% of the spatiotemporal variance (R^2) in the VIS, AUD and AV conditions, respectively.

5.1.1. Occipital Component/M150

The occipital component was found in the VIS and AV conditions, as shown in Figs. 6a–b. This component was associated with the first prominent visual peak at a latency of 130–150 ms. The spatial distribution map at 145–149 ms (see Figs. 6a–b) shows that the positive deflection reflects MEG activity in the bilateral occipital sensors. The positive deflection resembles the visual P100 wave described in previous MEG/EEG studies which could reflect the allocation of attentional resources (Boehler, Schoenfeld, Heinze, & Hopf, 2008; Zhang & Luck, 2009).

5.1.2. Right Temporal Component/M300a

The right temporal component with the peak 280–300 ms was consistently found in the AUD and AV conditions in the right temporal and inferior left/right frontal sensors (see Figs. 6c–d). The positive deflection at a latency of 280–300 ms corresponds to the early phase of the P300a

component, which has been linked to different processes, such as detecting and evaluating novel and orienting responses (Pfefferbaum, Ford, Weller, & Kopell, 1985; Polich, 2007; Vogel, Luck, & Shapiro, 1998).

The M300a component revealed two separate early latency subcomponents at about 53–56 ms (M50) and 82–86 ms (M100). These early subcomponents were found in the right temporal sensors in the AUD and AV conditions. In addition, the M100a component followed the M100 component in the AV condition and peaked around 148 ms. We show a zoomed version of the M300a component within 0–150 ms time window in Fig. 7. Fig. 7 depicts the peak latencies of the M50 and M100 subcomponents after the onset of a stimulus and the topographic scalp distribution of these early latency components. The M50 and M100 components exhibit much smaller amplitudes compared to the later M300a amplitude. It was noted that the amplitude of the M100 component was more robust and more evident than the M50 amplitude.

5.1.3. Late Central Component/M400

The late central component was extracted for all stimuli conditions (AUD, VIS, and AV), as shown in Figs. 6e–i. This component consists of a sequence of negative (VIS and AV) and positive peaks (AUD) at approximately 126–134 ms and a prominent peak at 350–430 ms. Figs. 6e–i show that this component is primarily distributed in the left temporal-parietal and right prefrontal sensors. The prominent negative deflection resembles the late phase of the parietally distributed N400 component (Halgren et al., 2002; Marinković, 2004).

Additionally, the M400 component explained early latency M100a subcomponent peaked around 126–148 ms and was identified in the VIS, AUD and AV conditions.

5.2. Subject Subgroup Identification

In this section, we present the clustering analysis results for identifying subgroups described in Section 3.4. The hierarchical clustering results using Ward’s D2 (Murtagh & Legendre, 2014) distance are shown in the dendrogram in Fig. 8a, with the height of the branches indicating the distance or dissimilarity between clusters. As depicted in Fig. 8a, two significant clusters were selected according to the approximately unbiased

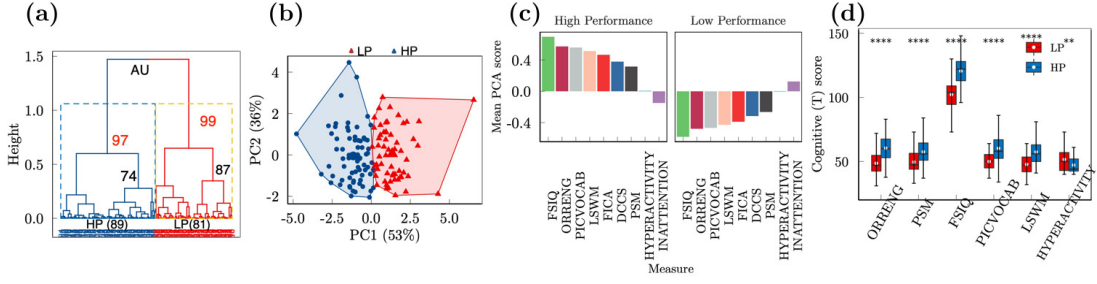


Fig. 8 Results of the hierarchical clustering on the principal components. (a) Dendrogram of hierarchical clustering based on Ward’s D2 criterion. The height of the branches indicates the dissimilarity between clusters. The number of retained clusters was chosen using an approximately unbiased probability measure (AU) (Efron et al., 1996). The significant clusters were selected based on unbiased probability p -values with $p < 0.05$ corrected for multiple comparisons using FDR. The final clustering solution was obtained with the K -means algorithm. (b) Clustering solution projected on the principal components. (c)–(d) Subgroup associations with neuropsychological (T) scores. (c) Distribution of mean PCA scores averaged across subject subgroups. (d) Main effect of subject subgroup on neuropsychological (T) score. Independent samples two-tailed t -test (FDR corrected, $p < 0.05$) showed statistically significant differences in the WASI-II FSIQ, language, memory ($p < 0.0001$) and inattention scores ($p < 0.01$) between the LP and HP groups. Details can be found in Supplementary Table S.1. **** $p < 0.0001$, ** $p < 0.01$.

probability (AU) (Efron et al., 1996) p -values with $p < 0.05$. The clustering solution projected on the PCs is shown in Fig. 8b. We used the HCPC clustering output to identify two subgroups with distinct distributions of mean PCA scores and categorized them as high ($n = 89$) or low ($n = 81$) performance. The distribution of mean PCA scores shown in Fig. 8c indicates that subjects in the HP group have higher PCA loading factors than the subjects in the LP group in all six cognitive domains except for the Conners 3 inattention and hyperactivity scores where a higher score implies greater inattention and hyperactivity.

We evaluated the effect of the subject group (HP vs. LP) on the cognitive assessments using independent samples two-tailed t -tests with unequal variances corrected for multiple comparisons using FDR ($p < 0.05$). We present the summary statistics of the neuropsychological (T) score distribution according to subject subgroup in Supplementary Table S.1. The groups did not differ significantly in terms of gender ($\chi^2 = 0.00, p = 0.99$), age ($t_{167} = 0.03, p = 0.97$) or parental SES ($t_{167} = 0.517, p = 0.61$). However, the WASI-II FSIQs differed significantly ($t_{167} = 9.16, p < 0.0001$), with higher FSIQ scores in the HP group than in the LP group. Similarly, the language (PICVOCAB, ORRENG), memory (PSM, LSWM), and executive function (DCCS, FICA) (T) scores differed significantly by group (Supplementary Table S.1; Fig. 8d; $p < 0.0001$), with cognitive (T) scores higher in the HP group

than in the LP group. The Conners 3 hyperactivity score differed significantly by group ($t_{167} = -2.17, p = 0.031$), with lower scores in the HP group than in the LP group. The Conners 3 inattention score did not differ significantly between the HP and LP groups ($t_{167} = 0.52, p = 0.61$). Fig. 8d shows the subject subgroup distribution of standard age-adjusted cognitive (T) scores.

5.3. Statistical Group-Level Analysis

In this section, we present the group-level analysis results described in Sections 4.5.1 and 4.5.2. This section has two subsections. The first subsection evaluates the statistical significance of the ERF components (see Section 5.1) to differentiate between the subgroups identified in Section 5.2. The second subsection assesses the covariant relationships between ERF components and neuropsychological measures to correlate brain responses with cognitive performance. The component loading factors in subject mode (columns of matrix **A**) and neuropsychological (T) scores were evaluated for normality. All analyses were corrected for multiple comparisons using FDR with a significance level of $\alpha = 0.05$ unless stated otherwise.

Prior to performing the group-level statistical (see Section 5.3.1) and component-cognitive scores correlation analyses (see Section 5.5), we analyzed pairwise component correlations (corrected for multiple comparisons using FDR ($p < 0.05$) for the VIS, AUD, and AV conditions). There were

Table 2 Summary of ERF components loading factor ANCOVA results

Effect	Occipital/M150		R.Temporal/M300a		L.Central/M400	
	$F(1, 336)$	p value	$F(1, 336)$	p value	$F(2, 494)$	p value
Condition	0.05	0.996 ^a	0.01	0.999 ^b	0.22	0.969 ^a
Group	33.96	< 0.0001 ^a	101.35	< 0.0001 ^a	176.73	< 0.0001 ^a
Condition \times Group	28.73	< 0.0001 ^a	6.82	0.031 ^a	0.79	0.982 ^a

^a p values from ANCOVAs adjusted for age, gender and parental SES (FDR corrected, $p < 0.05$).

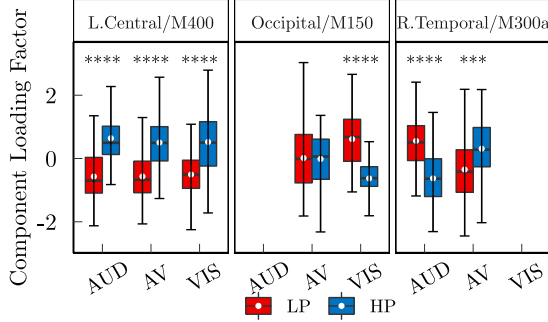


Fig. 9 Main effect of the subject group ($N = 170$, HP vs. LP) on component loading factors in subject mode for each stimulus condition. Boxplots summarize the distribution of the mean values of the component loading factors in subject mode for the HP and LP groups. Median values are represented by black lines inside the boxplot, with the top of the whisker lines indicating the 25th and 75th percentile values. Mean values are plotted in white circles. Error bars represent the standard error of the mean. Six components were statistically significant in the post hoc two-tailed t -test results (FDR corrected, $p < 0.05$), except for the occipital component/M150 in the AV condition, which did not differ between the HP and LP groups ($p = 0.864$). The post hoc t -tests results are shown in Table 3. **** $p < 0.0001$, *** $p < 0.001$ (post hoc and FDR corrected $p < 0.05$), which means significantly different.

no significant correlations between the CP components ($p > 0.05$ for all tests; see Supplementary Table S.6). These findings suggest that there is no need to adjust planned group-level statistical (see Section 4.5.1) and partial correlation analyses (see Section 4.5.2) for the presence of other CP components as model covariates. It should be noted that the CP model produces unique components so that the specific component or its factors are not associated with any other factors or other components (Kolda & Bader, 2009; Kruskal, 1977).

5.3.1. Group-Level Discriminative Components

We applied mixed measures two-way ANCOVA (see Section 4.5.1) on the component loading

factors in subject mode of each ERF component and stimulus condition to determine significant effects after controlling for the covariates. The mixed measures two-way ANCOVA comparison of the component loading factors showed a statistically significant stimulus condition \times group interaction (see Table 2) for the Occipital/M150 ($F_{1,336} = 28.73, p < 0.0001, \eta_G^2 = 0.101$) and R.Temporal/M300a components ($F_{1,336} = 6.82, p = 0.03, \eta_G^2 = 0.098$). There was no significant stimulus condition \times group interaction for the L.Central/M400 component ($F_{2,494} = 0.79, p = 0.982, \eta_G^2 = 0.004$). The main effect of the subject subgroup was statistically significant for each component (Table 2; Occipital/M150: ($F_{1,336} = 33.96, p < 0.0001, \eta_G^2 = 0.113$; R.Temporal/M300a: ($F_{1,336} = 101.35, p < 0.0001, \eta_G^2 = 0.281$; L.Central/M400: ($F_{1,494} = 176.73, p < 0.0001, \eta_G^2 = 0.311$). Post hoc analyses with two-tailed t -tests corrected for multiple comparisons using FDR ($p < 0.05$) revealed six components with significant group differences (HP vs. LP) in the component loading factors. The details are shown in Table 3 and Fig. 9. Fig. 10 depicts the group ERF components as solid lines (blue for HP and red for LP). The group ERF components peaked at the same time as the average group ERF timecourses, drawn in dashed lines (blue for HP and red for LP).

Occipital Component/M150:

The Occipital/M150 (130–150 ms) group component is shown in Fig. 10a–b. The activity was concentrated in the left and right occipital sensors. The HP group demonstrated a higher activation strength than LP group in the VIS condition (see Fig. 10a). The main effect of the group ($N = 170$, HP vs. LP) on the component loading factors was statistically significant for the VIS condition (Table 3, Fig. 9, Fig. 10a; VIS: $t_{165} = 7.86, p < 0.0001$, HP > LP, post hoc two-tailed t -test). However, there were no significant differences in the component loading factors of the Occipital/M150 component between the HP and LP groups for the AV condition (Table 3, Fig. 9, Fig. 10b; AV: $t_{165} = 0.166, p = 0.864$, post hoc two-tailed t -test).

Right Temporal Component/M300a:

The R.Temporal/M300a group component is shown in Figs. 11a–b. The component was associated with the peak at 280–300 ms and accounted for the activity in the right temporal and inferior

Table 3 Comparison of ERF component loading factors by subject subgroup

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
	<i>M(SD)</i>				<i>M(SD)</i>				<i>M(SD)</i>			
Occipital/M150	-	-	-	-	0.61(0.91)	-0.63(0.62)	7.86	< 0.0001 ^a	0.015(1.07)	0.013(0.93)	0.166	0.864 ^a
Right Temporal/M300a	-0.62(0.83)	0.56(0.79)	7.31	< 0.0001 ^a	-	-	-	-	0.54(0.85)	-0.31(1.05)	5.65	< 0.001 ^a
Late Central/M400	-0.57(0.81)	0.64(0.78)	-8.5	< 0.0001 ^a	-0.51(0.75)	0.52(0.96)	-7.22	< 0.0001 ^a	-0.56(0.83)	0.51(0.85)	-7.20	< 0.0001 ^a

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

^b the corresponding data are shown in Fig. 9.

left/right frontal sensors. Post hoc two-tailed *t*-tests found a significant main effect of group ($N = 170$, HP vs. LP) on the component loading factors for both the AUD and AV conditions (Table 3, Fig. 9, Figs. 11a–b; AUD: $t_{165} = 7.31, p < 0.0001$, HP > LP; AV: $t_{165} = 5.65, p < 0.001$, HP > LP).

Late Central Component/M400:

The L.Central/M400 group component is shown in Figs. 11c–e. This component was associated with activity in the left temporal-parietal and right prefrontal sensors. Two-tailed *t*-tests identified a significant main effect of group ($N = 170$, HP vs. LP) on the component loading factors for the VIS, AUD and AV conditions (Table 3, Fig. 9, Figs. 11c–d; AUD: $t_{165} = -8.5, p < 0.0001$, HP < LP; VIS: $t_{165} = -7.22, p < 0.0001$, HP < LP; AV: $t_{165} = -7.20, p < 0.0001$, HP < LP).

5.4. Group-Level Sensitivity Analyses

In this section, we evaluate the discriminative performance of the CP decomposition and the nonparametric method based on the permutation statistics of the ERF in sensor space. Previous studies have shown that statistical nonparametric mapping (SnPM) is a robust approach that can reliably detect ERF activity in sensor space (Nichols & Holmes, 2002; Pantazis, Nichols, Baillet, & Leahy, 2003).

To quantify differences between subject subgroups, group amplitudes were compared by running time-point by timepoint nonparametric permutation two-tailed *t*-tests Maris and Oostenveld (2007); Nichols and Holmes (2002) assessed at each sensor from 0 to 800 ms poststimulus. The nonparametric statistical threshold t_{\max} from the pseudo *t*-distribution was calculated to establish timepoint/sensor significance at the $p < 0.05$ level. The sensors and timepoints identified by these *t*-tests denoted spatiotemporal regions where statistically significant differences in group amplitudes occurred.

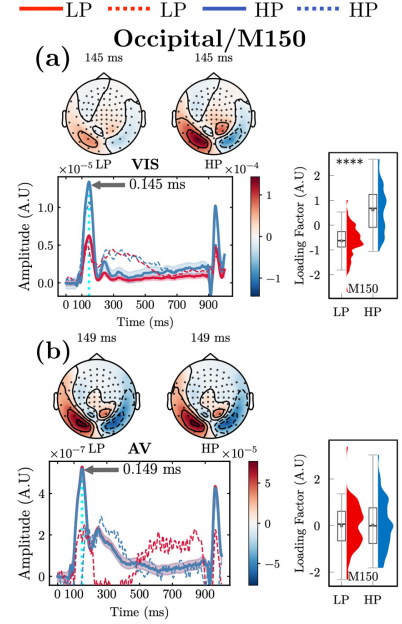


Fig. 10 Discriminative group MEG ERF components from the CP decomposition for the VIS and AV conditions are shown. The ERF components are indicated by solid lines (HP: high performance – blue, LP: low performance – red). The average group ERF timecourses are indicated by dashed lines. Main effect of the subject group (HP vs. LP, $N = 170$) on the component loading factors in subject mode is summarized in the boxplots. The boxplots summarize the distribution of the mean values of component loading factors in the subject mode for the HP and LP groups. The median values are represented by black lines inside the boxplot, and the tops of the whisker lines indicate the 25th and 75th percentile values. The mean values are plotted in dark grey. 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 the component loading factors of the HP group was significantly different than for LP group with $p < 0.001$ for six ERF components except for the occipital/M150 component in the AV condition (AV/M150: $p = 0.864$). (a)–(b) Occipital M150 component. **** $p < 0.0001$, *** $p < 0.001$ indicate significant differences (FDR corrected, $p < 0.05$). The post hoc *t*-tests results are shown in Table 3.

To compare the group-level sensitivity of the CP decomposition and SnPM, we computed timepoint/sensor-wise *t*-statistics from a

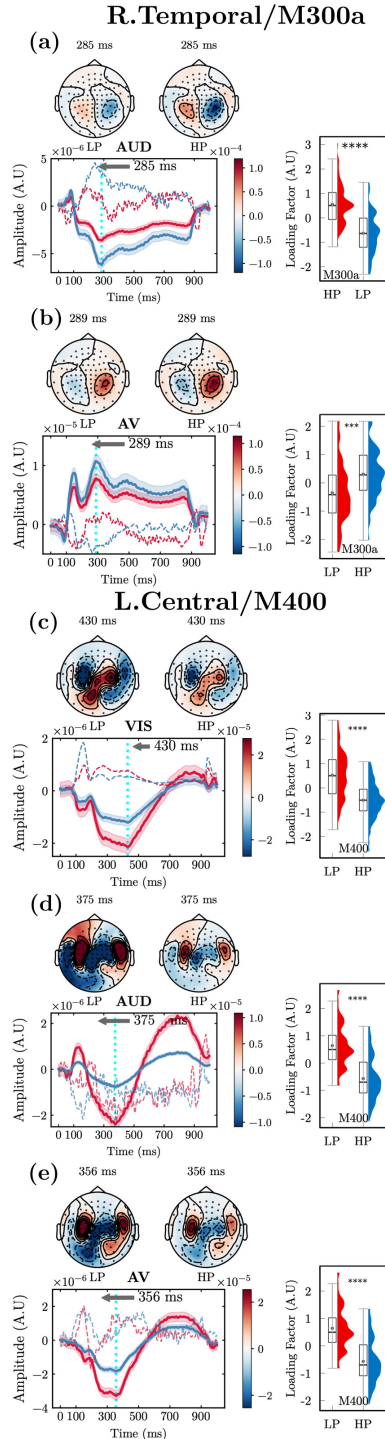


Fig. 11 Discriminative group MEG ERF M300a and M400 components from the CP decomposition for the VIS, AUD, and AV conditions are shown. (a)-(b) Right temporal M300a component. (c)-(e) Late central M400 component. See full caption and legend in Fig. 10.

two-tailed nonparametric permutation t -test and determined the significance of the group-level mean amplitudes between subject subgroups, taking into account the covariates. We evaluated the group-level sensitivity by investigating the SnPM method's ability to discriminate between subject subgroups. We present group-level ERF components and the statistical images (T-maps) after the CP and SnPM in Fig. 12 and Supplementary Figs. S.8-S.13. The SnPM identified five significant components: Occipital/M150 in the VIS condition (Supplementary Table S.4; VIS: $p < 0.001$; $t_{165} = 6.32$), R.Temporal/M300a component in the AUD and AV conditions (Supplementary Table S.4; AUD: $p = 0.022$; $t_{165} = 2.31$, AV: $p = 0.038$; $t_{165} = 2.09$), and L.Central/M400 in the VIS, AUD and AV conditions (Supplementary Table S.4; VIS: $p = 0.001$; $t_{165} = -3.76$; AUD: $p = 0.002$; $t_{165} = -3.41$).

The ERF components and T-maps generated after the CP and SnPM methods for the Occipital/M150 component in the VIS condition at 144-145 ms are shown in Fig. 12. The results illustrate that the CP decomposition provided a higher number of significant sensors than the SnPM method. We note a similar observation for the other ERF components presented in Supplementary Figs. S.8-S.13. It is evident from Fig. 12 and Supplementary Figs. S.8-S.13 that the number of adjacent sensors is smaller for the ERF components produced by the SnPM. We quantify the performance of each component estimation method by using Cohen's d effect size (the standardized magnitude difference between groups (Sullivan & Feinn, 2012)) and p -values. Fig. 13 and Table 4 show that the CP decomposition resulted in a higher magnitude of the group differences and higher p -values compared with the SnPM method. In summary, the results presented in Table 4, Figs. 12, Supplementary Figs. S.8-S.13 and Fig. 13 demonstrate that the t -statistics and the magnitude of the effect are higher for the CP decomposition, suggesting better sensitivity over the nonparametric statistical approach.

5.5. Analysis of ERF Component Association with Cognitive Domains

To correlate neuropsychological scores with ERF components, we performed two-tailed partial

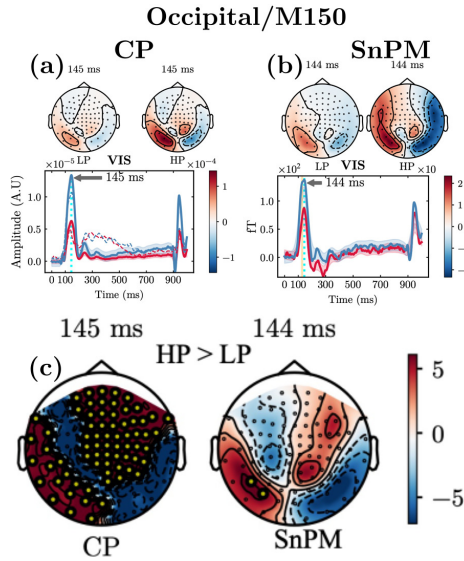


Fig. 12 Sensitivity analysis of the ERF components generated with different group imaging methods. Top row: Estimation of the Occipital/M150 component for the VIS condition by the CP decomposition and SnPM methods. Bottom row: The group-level T-maps between HP and LP groups for the CP and SnPM methods 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 VIS M150 component. (b) SnPM VIS M150 component. The significant time interval of group differences (140–150 ms) is depicted in the shaded area. (c) Left: CP VIS M150 T-map at 145 ms. Right: SnPM VIS M150 T-map at 144 ms.

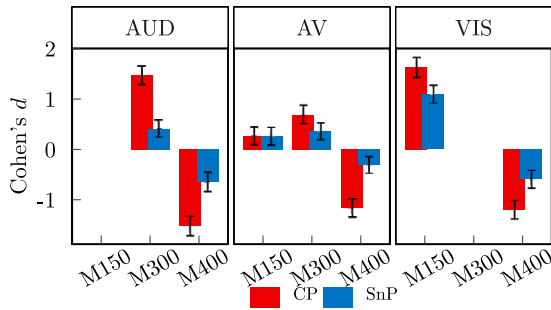


Fig. 13 Effect size comparison of the ERF components generated with different group imaging methods. The error bars represent the standard error of the mean. Cohen's d effect size of group-level discriminative (HP vs. LP) components using the CP decomposition and SnPM method for the AUD, AV, and VIS conditions. The effect size results are listed in Table 4.

Pearson's correlation tests between component loading factors in subject mode (columns of matrix **A**) for the nine cognitive age-adjusted neuropsychological (T) scores in each subject group

Table 4 Effect size comparison of group-level discriminative components after the CP and SnPM methods

Component	Method			
	SnPM		CP	
	Cohen's d (SEM)	p value	Cohen's d (SEM)	p value
Occipital/M150				
VIS	1.09(0.18)	$p < 0.001^a$	1.63(0.20)	$< 0.0001^a$
AV	0.11 (0.17)	0.851 ^a	0.10 (0.17)	0.864 ^a
R.Temporal/M300				
AUD	0.41(0.17)	0.022 ^a	1.47(0.19)	$< 0.0001^a$
AV	0.36(0.17)	0.038 ^a	0.69(0.18)	$< 0.001^a$
L.Central/M400				
VIS	-0.59(0.17)	0.001 ^a	-1.2(0.18)	$< 0.0001^a$
AUD	-0.65(0.17)	0.002 ^a	-1.52(0.19)	$< 0.0001^a$
AV	-0.31(0.17)	0.07 ^a	-1.27 (0.18)	$< 0.0001^a$

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

^b the corresponding data are shown in Fig. 13.

Table 5 Component loading factor and cognitive (T) score associations

Component	Condition	Cognitive (T) score					
		PSM					
		HP ($n = 89$)			LP ($n = 81$)		
		r	$t(86)$	p value	r	$t(78)$	p value
Right Temporal	AUD	0.342	3.33	0.003^a	0.291	2.65	0.039^a
	AV	0.364	3.59	0.001^a	0.297	2.71	0.016^a
PICVOCAB							
Late Central	VIS	-0.242	-2.29	0.013^a	-0.12	-1.051	0.295 ^a
DCCS							
Right Temporal	AUD	0.267	2.42	0.018^a	-0.017	-0.161	0.873 ^a
	AV	0.277	2.51	0.014^a	-0.071	-0.564	0.561 ^a
HYPERACTIVITY							
Occipital	VIS	-0.233	-2.21	0.030^a	-0.086	-0.685	0.992 ^a

^a p values (FDR corrected, $p < 0.05$) from two-tailed partial Pearson's correlation tests adjusted for age, gender and parental SES.

(HP: $n = 89$; LP: $n = 81$) and the full sample ($N = 170$) (see detailed protocol in Section 4.5.2). The partial correlation analyses were controlled for age, gender and parental SES. The correlations between component loading factors and cognitive scores were corrected for multiple comparisons using FDR with significance threshold $p < 0.05$.

The correlation analyses indicated that three functional ERF components were significantly associated with the language, episodic memory and attention cognitive domains (see Table. 5). The results of two-tailed partial Pearson's correlation analyses are summarized in Table 5, Figs. 14 and 15.

Among the nine cognitive variables, the PICVOCAB (T), PSM (T), DCCS (T), and Connors 3 hyperactivity scores were significantly

correlated with ERF components. We present significant functional association patterns between the ERF components and the cognitive scores in Table 5.

The Occipital/M150 component in the VIS condition (Figs. 6a and 10a) was negatively correlated with the hyperactivity score in the HP group ($r_{\text{VIS-HYPERACTIVITY}}(86) = -0.233, p = 0.03$; Table 5; Fig. 15b). However, the correlation was not significant ($r_{\text{VIS-HYPERACTIVITY}}(167) = -0.194, p = 0.051$; Fig. 14a) in the full sample. The correlation was consistent with large group differences (HP > LP; $p < 0.0001$) in the occipital component loading factors identified by two-tailed t -tests (see Table 3).

The R.Temporal/M300a component (Figs. 6c-d and 11a-b) had a statistically significant positive correlation with the PSM (T) score in both groups (Table 3; Figs. 15c-d; HP: $r_{\text{AUD-PSM}}(86) = 0.342, p = 0.003$, $r_{\text{AV-PSM}}(86) = 0.364, p = 0.001$; LP: $r_{\text{AUD-PSM}}(78) = 0.291, p = 0.039$, $r_{\text{AV-PSM}}(78) = 0.297, p = 0.016$) and the full sample ($r_{\text{AUD-PSM}}(167) = 0.239, p = 0.004$; $r_{\text{AV-PSM}}(167) = 0.303, p = 0.001$; Figs. 14c-d).

Additionally, the correlation tests revealed significant associations between the R.Temporal/M300a component and DCCS (T) score (see Table 3 and Figs. 15e-f) in the HP group (AUD: $r_{\text{AUD-DCCS}}(86) = 0.267, p = 0.018$; AV: $r_{\text{AV-DCCS}}(86) = 0.277, p = 0.014$). The correlations between the R.Temporal/M300a component and PSM/DCCS scores were consistent with group differences (HP > LP; $p < 0.001$) in the component loading factors for the AUD and AV conditions (see Table 3).

The L.Central/M400 component in the VIS condition (Figs. 6e and 11c) was negatively correlated with the PICVOCAB (T) score in the HP group ($r_{\text{VIS-PICVOCAB}}(86) = -0.242, p = 0.013$; Table 5; Fig. 15a) and the full sample ($r_{\text{VIS-PICVOCAB}}(167) = -0.208, p = 0.017$; Fig. 14b).

No significant associations were found in the HP group, LP group or full sample for the other cognitive scores in the component loading factors ($p > 0.05$ all tests). There were no significant differences in partial correlations between the groups (HP vs. LP) or stimulus conditions (VIS vs. AUD vs. AV) within subject groups ($p > 0.05$ all tests).

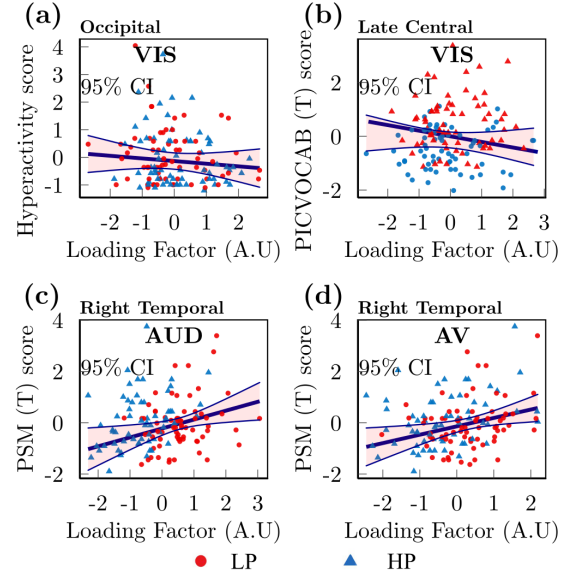


Fig. 14 Significant (FDR corrected, $p < 0.05$) two-tailed partial Pearson's correlations (correlation coefficient, r) between ERF components and neuropsychological (T) scores in the full sample. The linear fit and 95% confidence intervals (CIs) are shown. The blue dots denote HP group, and red dots denote LP group. (a) The occipital component was negatively correlated with the hyperactivity score in the VIS condition. (b) The late central component was negatively correlated with the PICVOCAB score in the VIS condition. (c)-(d) The right temporal component was positively correlated with the PSM (T) score in the AUD and AV conditions.

6. Discussion

This paper presents a tensor analysis-based model of MEG multi-subject data for identifying ERF components representative of typical brain developmental patterns in a healthy population of children and adolescents. The tensor analyses and tensor-based group-level statistical inferences outlined in this paper establish a foundational framework for extracting latent factors associated with children's brain development from MEG datasets.

We contribute to the developmental neuroscience literature on the relationship between MEG activity and cognition by correlating ERF components from a healthy pediatric population with neuropsychological age-adjusted cognitive (T) scores and attentional indices.

6.1. ERF Components Extraction

To the best of our knowledge, this is the first study that modeled event-related field MEG data

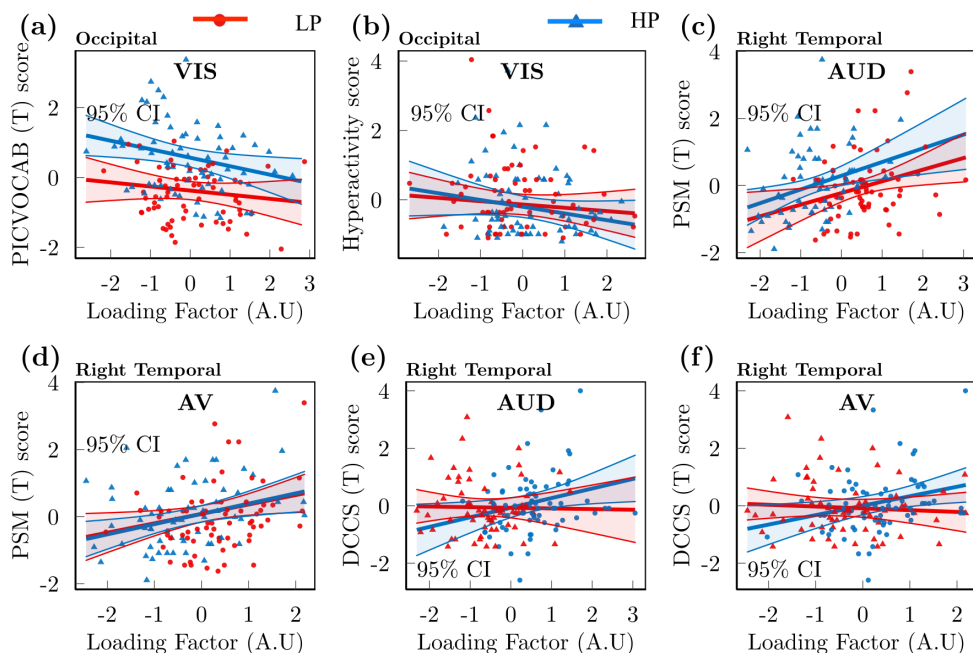


Fig. 15 Significant (FDR corrected, $p < 0.05$) two-tailed partial Pearson's correlations (correlation coefficient, r) between ERF components and neuropsychological (T) scores in the HP and LP groups (HP: blue dots, LP: red dots). The group linear fit and 95% CIs are shown (HP: blue line, LP: red line). (a) The late central (VIS) component was negatively correlated with the PICVOCAB score and significant in the HP group. (b) The occipital (VIS) component was negatively correlated with the hyperactivity score and significant in the HP group. (c)-(d) The right temporal component (AUD/AV) was positively correlated with the PSM score and significant in both groups. (e)-(f) The right temporal component (AUD/AV) was positively correlated with the DCCS score, significant in the HP group and showed a negative trend in the LP group. The correlation results are listed in Table 5.

as a low-rank third-order tensor. We demonstrated that CP factorization could produce latent factors that result in functionally relevant ERF components and reveal meaningful spatiotemporal brain patterns. The CP model was shown to be highly effective in capturing informative data representation. For example, Fig. 6 and Supplementary Table. S.2 illustrate that ERF components were well-matched with the average ERF waveforms and demonstrated significant correlations ($p < 0.001$) with the original datasets (see Supplementary Methods Section 3.2; Supplementary Table. S.3). The tensor analysis successfully identified latent brain developmental patterns across subjects for each stimulus condition. As expected, the Occipital/M150 component similar to the visual P100/M100 wave, was extracted from the VIS and AV conditions, as shown in Figs. 6a–b. We observed P300 component in the R.Temporal/M300a component (see Fig. 6c–d) extracted from the AUD and AV conditions. The absence of the visual Occipital/M150 component in the AUD condition and the absence of the

R.Temporal/M300a component in the VIS condition confirms that the CP model can extract meaningful patterns corresponding to expected ERF responses. The L.Central/M400 component (Figs. 6e–i) was extracted from the AUD, VIS, and AV conditions, representing activity in the left temporal-parietal sensors and likely capturing the motor response required in all three conditions.

In the present study, we found the M50/M100 subcomponents less dominant in terms of amplitude for the AUD and AV conditions. Our results agree with previous MEG studies (Cardy, Ferrari, Flagg, Roberts, & Roberts, 2004; Edgar et al., 2014; Kotecha et al., 2009), which reported that children do not show the M50/M100 adult-like waveforms until early adolescence. It was shown in (Bruneau, Roux, Guerin, Barthelemy, & Lelord, 1997; Kotecha et al., 2009; Ponton, Eggermont, Kwong, & Don, 2000) that the amplitude of the auditory component becomes more prominent with increasing and remains stable through adulthood. Our future work may include longitudinal studies of the same MEG dataset where we

can explore the latency and amplitude-age dependencies on the characteristics of the early auditory components.

6.2. Comparison of Group-Level Sensitivity Analyses

A significant problem in MEG research involves the detection of significant effects while controlling the FWER rate. In section 5.4, we compared the sensitivity and statistical power of the CP tensor decomposition model and conventional statistical nonparametric approach based on permutation testing. We have demonstrated that the CP model provides 1.5–2 times higher effect sizes and lower p -values (see Fig. 13). The lower sensitivity of the SnPM method is most likely caused by a higher significance threshold required to control the FWER. Evidently, better performance of the CP model is associated with a lower rate of false negatives. Thus, tensor-based group-level inference alleviates the multiple comparison problem due to dimensionality reduction and provides higher statistical power.

6.3. Functional Associations of Group-Discriminating Components

The CP decomposition of the multi-subject MEG data provides insight into how tensor analysis can be used to explore relationships between brain patterns and cognitive function in high- and low-performance subjects. The statistical tests confirmed the effect of the cognitive group on the relationship between the component loading factors and designated pediatric subgroups (see Section 5.3). We showed that children in the HP group were significantly different from those in the LP group (HP vs. LP; Table 3) in terms of the six components, with large effect sizes ($\eta_G^2 > 0.113$, see Section 5.3.1). The group differences (HP > LP) were consistent with the neuropsychological (T) score t -tests, with the HP group scoring significantly higher than the LP group on all cognitive and behavioral tests, except for the Connors 3 inattention and hyperactivity scores. The subjects in the HP and LP groups did not show differences in the spatial distributions of all components; however, they demonstrated significant differences

in the spatial activation strength and timecourse amplitude.

To identify ERF components as informative signatures of cognitive function, we correlated the ERF components with neuropsychological (T) scores in the full sample and each subject group. In the full sample, we found statistically significant correlations between neuropsychological (T) scores and specific ERF components (see details in Section 5.5), namely, the PSM (T), PICVOCAB (T) and Connors 3 hyperactivity scores.

The correlation analyses between the ERF components and neuropsychological scores revealed significant associations between ERF components and PSM (T) score in both the HP and LP pediatric groups, whereas the PICVOCAB (T), DCCS (T) and hyperactivity scores were significantly correlated with ERF components only in the HP group.

It was shown in the literature that the PSM test measures episodic memory (Dikmen et al., 2014), and the DCCS test measures cognitive flexibility and executive function (Weintraub et al., 2013). The PICVOCAB test measures verbal ability and language comprehension (Weintraub et al., 2013), and the Connors 3 hyperactivity score is an attentional index (Connors, 2008). These cognitive indices are foundational cognitive processes that change rapidly during development and vary across individuals.

6.3.1. Early Latency Components

The hyperactivity score was negatively correlated with the Occipital/M150 component in the VIS condition using the full sample data. In addition, the analyses revealed a significant negative correlation between the hyperactivity score and the Occipital/M150 component in the HP group. Finally, as shown in Supplementary Table S.1, the hyperactivity score was significantly lower in the HP group than in the LP group. These findings suggest that the Occipital/M150 component patterns were consistent between the correlation results and ERF component group-level analyses (HP > LP; Table 3), as well as with the group-level neuropsychological (T) score analyses (Supplementary Table S.1). It was shown in the literature (Ghani, Signal, Niazi, & Taylor, 2020; Sokhadze et al., 2017) that early and mid-latency ERP components (N100, N200, and

P200) are related to involuntary attention selection mechanisms. Our findings are consistent with the literature on healthy individuals (Allison & Polich, 2008; Ghani, Signal, Niazi, & Taylor, 2021; Kramer, Trejo, & Humphrey, 1995) and children with ADHD (Liotti, Pliszka, Higgins, Perez III, & Semrud-Clikeman, 2010; van Meel, Heslenfeld, Oosterlaan, & Sergeant, 2007), which reported lower early latency component amplitudes with increasing cognitive workload in patients relative to healthy controls.

6.3.2. Late Latency Components

The right temporal/M300a component in the AUD and AV conditions was positively associated with the PSM (T) score ($p < 0.01$) in the HP group, LP group, and the full sample. The partial correlation analyses for the PSM (T) score shown in Figs. 14c–d and 15c–d suggest that the increase in the value of the component loading factors was associated with a better PSM (T) score. Additionally, the right temporal/M300a component was significantly positively correlated with the DCCS (T) score in the HP group, which is consistent with the component loading factor group difference (HP > LP; Table 3). Similarly, the correlation analyses for the DCCS (T) score shown in Fig. 15e–f and Table 5 suggest that the increase in value of the component loading factors associated with a higher DCCS (T) score. The HP group scored significantly higher (Supplementary Table S.1; $p < 0.0001$) than the LP group in terms of both the PSM (T) and DCCS (T) scores. According to the literature, the P300a/M300a component amplitude represents an orienting response, reflecting involuntary orientations to attention-catching changes (Sur & Sinha, 2009). Additionally, the P300a/M300a component was categorized as an indicator of implicit memory and item familiarity (Friedman & Johnson Jr, 2000; Graf & Schacter, 1985; Rugg et al., 1998). The existing literature shows that the P300a/M300a component could indicate sustained attention and decreases in P300a amplitude with increasing cognitive workload (Berti & Schröger, 2003; Dyke et al., 2015; Horat et al., 2016).

The late central/M400 component showed a significant negative correlation with the PICVOCAB (T) score in the HP group and full sample. Similar to the N400 (Kutas & Federmeier, 2000)

ERP component, the spatial scalp distribution was maximally concentrated in the left temporal-parietal sensors.

The association of the late central/M400 component with the PICVOCAB (T) score (see Figs. 14b and 15a) demonstrated that the reduced M400 component amplitude results in a higher PICVOCAB (T) score. The literature indicates (Fitz & Chang, 2019) that N400 component ERPs may reflect prediction error signals needed for learning; thus, larger ERP amplitudes could be correlated with errors.

6.4. Comparison of Group-Level Imaging Methods

In group-level brain imaging studies, the goal is to determine spatiotemporal patterns of variability between/among groups or conditions. The sensitivity and statistical power of group-level inferences is dependent on the stability and unique presentation of these patterns to determine where and when a specific brain activity occurs.

In MEG research, the most common approach to identify a location of brain activity is to employ mass-univariate hypotheses testing methods (Groppe, Urbach, & Kutas, 2011). Mass-univariate hypothesis testing is based on executing multiple tests, which most often involves computing a parametric or nonparametric t -test for each timepoint/sensor. However, mass-univariate analyses in MEG have many shortcomings, such as (1) the high dimensionality of data requires a large number of tests corrected for multiple comparisons; (2) potentially overlapping sources of brain activity; (3) not taking into account interactions between timepoints/sensors; and (4) sensitivity of peak/mean amplitude measures to the analysis window (Luck & Gaspelin, 2017). Nonparametric approaches based on randomized permutation and cluster-based permutation tests (Groppe et al., 2011; Maris & Oostenveld, 2007) have been developed that inherently address multiple comparison problems (Westfall & Young, 1993) and locate the spatiotemporal effect of interest. However, the important drawback of nonparametric statistics is that with the increase in the number of tests, the power of the permutation test is diminished due to an overly conservative estimate of the significance threshold (Groppe et al.,

2011). Thus, with an increase in the dimensionality of multi-subject MEG data, the strong FWER control of the permutation method may impact the sensitivity of the analyses, resulting in Type II error.

A more promising approach to overcome the shortcomings of mass univariate approaches is to use an effective multivariate approach to summarize the data. Group-level tensor decomposition is a multivariate latent space group-analysis technique that has been shown to be capable of (1) localization of common unique patterns of brain activity for a group of subjects in a data-driven way (Cong et al., 2012; Tangwiriyaakul et al., 2019; D. Wang et al., 2018); (2) dimensionality reduction (Cichocki et al., 2016); (3) extraction of region-of-interest independent signatures for group-level inferences (Acar, Levin-Schwartz, Calhoun, & Adali, 2017; Cong et al., 2012; Tangwiriyaakul et al., 2019); (4) inherent alleviation of multiple-comparison problem; and (5) higher sensitivity by capturing complex spatiotemporal interactions (Acar et al., 2019; Kinney-Lang et al., 2019; Kinney-Lang, Spyrou, Ebied, Chin, & Escudero, 2017).

As we show next, we discuss the differences in statistical assessments of the CP model and conventional sensor/source level imaging methods. For the further discussion below, we assume that the time-frequency source reconstruction where subject datasets $\mathbf{S} \in \mathbb{R}^{C \times T \times F}$ joined in the subject mode forming a fourth-order tensor $\mathcal{X} \in \mathbb{R}^{K \times T \times C \times F}$ ($\mathbb{R}^{\text{subject} \times \text{time} \times \text{sensor} \times \text{frequency}}$).

6.4.1. Tensor Analysis in the Sensor/Source Level Space

After the MEG/EEG data are factorized by the CP model, the underlying component matrices can be readily analyzed by group-level statistical inference algorithms (Cong et al., 2015). Since the CP model performs simultaneous factorization and is fully multivariate, each factor of the latent CP component is identified at all levels of other factors.

Hence, the magnitude of the underlying CP component is quantified at each timepoint and sensor in the sensor-level space, which eliminates the necessity of selecting specific timepoints and sensor sites (timepoint, sensor and frequency at

the source level) for the group amplitude extraction in the group-level inferences. Thus, statistical inference can be directly applied to the selected component signatures. For example, as shown in Fig. 3c, to determine the discriminative groups in the subject mode, the r th subject loading factor $\mathbf{a}_r \in \mathbb{R}^K$ is used in multifactorial ANCOVAs to evaluate the experimental conditions. Similarly, the spatial $\mathbf{c}_r \in \mathbb{R}^C$, temporal signatures $\mathbf{b}_r \in \mathbb{R}^T$ or frequency signatures $\mathbf{f}_r \in \mathbb{R}^F$ can be statistically evaluated to determine the significance of the spatiotemporal extent or frequency bands. Notably, due to dimensionality reduction, the required number of tests to reach statistical significance is dramatically reduced since only the limited number of samples (order of 10^2) is used from each signature instead of all spatiotemporal/space-time-frequency features (order of 10^6 or higher).

Thus, the CP model reduces the problem of multiple comparisons in the group-level analyses since the extracted component signatures are used to determine discriminatory ERF components. As a result, the CP model could provide a higher sensitivity by reducing Type II error. In contrast, the univariate and nonparametric tests may fail to find a significant effect if they are applied to the full subject \times time \times sensor data cube.

In summary, the CP tensor decomposition offers the following advantages compared with univariate parametric and nonparametric statistical methods, such as data-driven source separation, a region-of-interest independent measure for group-level analyses, identification of common spatiotemporal patterns for a group of subjects, and alleviation of multiple comparison problems due to dimensionality reduction, which could result in higher statistical power and better sensitivity as shown in Sections 5.4 and 6.2.

6.4.2. Tensor Analysis for Source Localization

The localization of brain sources based on MEG/EEG recordings has been an ongoing topic of active research due to increased demand in clinical applications (Asadzadeh, Rezaii, Beheshti, Delpak, & Meshgini, 2020).

In the past decade several works have proposed tensor-based preprocessing (Becker, Albera, Comon, Gribonval, & Merlet, 2014; Becker,

Albera, Comon, Haardt, et al., 2014; De Vos et al., 2007; Mørup, Hansen, Herrmann, Parnas, & Arnfred, 2006) for source localization. Primarily, the proposed tensor-based source localization approaches are based on the transformation of the evoked field data in the sensor space using a space-time-frequency (STF) or space-time-wavevector (STWV) transform and subsequent application of the CP decomposition using the STF or the STWV transformed data. The details of transformations are described in (Becker, Albera, Comon, Haardt, et al., 2014). The tensor group-level analysis in the source space would be similar to the tensor analysis in the sensor space (see Sections 3.2, 6.4.1 and 5.3). As suggested in (Becker, Albera, Comon, Haardt, et al., 2014) in order to fit the dipole model, the STF data tensor of each subject should be constructed with one source per time and frequency under the hypothesis of oscillatory signals. We further refer our readers to the existing key papers (Asadzadeh et al., 2020; Becker, Albera, Comon, Gribonval, & Merlet, 2014) for the history, and various applications.

7. Limitations and Future Work

The proposed generative model using CP decomposition implies that all subjects have the same number of latent components R , and all subjects share the common matrices \mathbf{B} and \mathbf{C} . In other words, the CP model imposes strict assumptions such that the underlying brain patterns have identical timecourses and spatial maps across subjects. However, with real ERF MEG data, individual differences may exist in the timing and origin of the subject’s neural responses to the stimuli. For example, individual differences in the cognitive processing of the stimuli would result in differences in the timing and spatial distribution in MEG. To allow a variable number of components and spatiotemporal variability of brain patterns across subjects, a more flexible model can be used, such as constrained PARAFAC2 (Parallel Factor Analysis) (Afshar et al., 2018; Chatzichristos, Kofidis, Morante, & Theodoridis, 2019; Helwig & Snodgrass, 2019), or higher-order block term decomposition (BTD2) (Chatzichristos et al., 2019). It has been shown in (Harshman

et al., 1972; Helwig & Snodgrass, 2019) that PARAFAC2 can handle the heterogeneity of subject’s responses and allows a variable number of latent components per subject (Afshar et al., 2018) via sparsity constraints. To address this limitation, our future work may include using the PARAFAC2 or BTD2 models to account for subject’s individual differences.

Another limitation of our generative model is fitting each stimulus condition as a separate CP decomposition. Alternatively, the multi-task multi-subject MEG data could be modeled as a coupled tensor-tensor decomposition (CTTD) (Chatzichristos et al., 2022; Jonmohamadi et al., 2020), where each stimulus (VIS, AV, and AV) represented as a third-order tensor and coupled in the subject mode. The multi-task joint learning enables the use of the complementary information (Acar et al., 2013; Lahat et al., 2015) from multiple stimuli and thus could result in the latent components with a higher discriminative power.

The extracted ERF components could be used as bioimaging markers for classification or prediction. Specifically, the subject loading factors found from MEG data using the CP model can be interpreted as feature extraction. The combination of machine learning techniques and multi-task tensor decomposition of MEG data could identify more reliable bioimaging markers that may enable the exploration of neurological differences associated with symptom onset, enabling early intervention. Thus, the application of multi-subject MEG tensor decomposition in context of machine learning is a promising direction for future research in cognitive neuroscience.

8. Conclusion

We demonstrated that CP decomposition can be used for the effective identification and characterization of latent spatiotemporal components of multi-subject MEG data. We described the generative model for the multidimensional representation of multi-subject MEG data, latent component extraction and group-level statistical inference methodologies. We demonstrated that the group-level tensor decomposition recovers meaningful distinct brain patterns of varying spatiotemporal brain activity across subjects in healthy population of children/adolescents and in subgroups. The

advantages of the proposed method include successful identification of the underlying latent brain patterns in the form of factor matrices via tensor factorization to allow for statistical assessment of the identified sources. The presented tensor-based group-level inference using CP component matrices eliminates the need to select specific regions of interest, such as time windows or specific sensor sites.

Using the proposed approach, we show that the tensor group-level analyses and tensor-based feature extraction allow us to investigate differences in brain activity between different subject groups. Given the importance of group-level inferences in neuroimaging studies, the extracted latent ERF components could be used to study differences in brain patterns across groups and aid in understanding how spatiotemporal brain activity can explain cognitive function and developmental changes directly from electrophysiological measurements. The application of MEG tensor decomposition used in this study is a promising direction for future research on other populations with different age ranges or developmental disorders.

Information Sharing Statement

The code used in this manuscript can be found at <https://github.com/ibelyaeva/meg-tensor-decomposition>.

Electronic Supplementary Material. This article contains supplementary material.

Acknowledgments. This work was supported in part by NSF-NCS 1631838, NSF 2112455, NIH R01 MH118695, NIH R01MH123610, NIH R01AG073949, NIH R01 MH121101, and NIH P20 GM144641.

Data Availability. The data supporting this study's findings are available from the Dev-Cog MRN repository (Stephen et al., 2021)¹ upon user account request (contact email: vcalhoun@gsu.edu, jstephen@mrn.org).

Authors contributions. Irina Belyaeva – Investigation, Formal analysis, Software, Validation, Data curation, Visualization, Writing - original draft, Ben Gabrielson – Visualization, Validation, Writing - review & editing, Visualization, Yu-Ping Wang – Conceptualization, Validation, Writing – review & editing, Tony W. Wilson – Conceptualization, Resources, Data curation, Validation, Funding acquisition, Writing - review & editing, Vince D. Calhoun – Conceptualization, Methodology, Validation, Resources, Project Administration, Funding acquisition, Writing - review & editing, Julia M. Stephen - Conceptualization, Methodology, Data curation, Validation, Resources, Supervision, Project Administration, Funding acquisition, Writing - review & editing, Tülay Adalı – Conceptualization, Methodology, Resources, Validation, Supervision, Project Administration, Funding acquisition, Writing - review & editing.

Declarations

Competing Interests. All the authors declare that they have no competing interests.

¹<http://devcog.mrn.org>

References

- Ablin, P., Cardoso, J.-F., Gramfort, A. (2021). Spectral independent component analysis with noise modeling for M/EEG source separation. *Journal of Neuroscience Methods*, 356, 109144.
- Acar, E., Levin-Schwartz, Y., Calhoun, V.D., Adali, T. (2017). ACMTF for fusion of multi-modal neuroimaging data and identification of biomarkers. *2017 25th european signal processing conference (eusipco)* (pp. 643–647).
- Acar, E., Rasmussen, M.A., Savorani, F., Næs, T., Bro, R. (2013). Understanding data fusion within the framework of coupled matrix and tensor factorizations. *Chemometrics and Intelligent Laboratory Systems*, 129, 53–63.
- Acar, E., Schenker, C., Levin-Schwartz, Y., Calhoun, V.D., Adali, T. (2019). Unraveling diagnostic biomarkers of schizophrenia through structure-revealing fusion of multi-modal neuroimaging data. *Frontiers in neuroscience*, 13, 416.
- Adali, T., Anderson, M., Fu, G.-S. (2014). Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging. *IEEE Signal Processing Magazine*, 31(3), 18–33.
- Adali, T., Levin-Schwartz, Y., Calhoun, V.D. (2015). Multimodal data fusion using source separation: Application to medical imaging. *Proceedings of the IEEE*, 103(9), 1494–1506.
- Afshar, A., Perros, I., Papalexakis, E.E., Searles, E., Ho, J., Sun, J. (2018). COPA: constrained PARAFAC2 for sparse & large datasets. *Proceedings of the 27th acm international conference on information and knowledge management* (pp. 793–802).
- Akhavan, S., Baghestani, F., Kazemi, P., Karami, A., Soltanian-Zadeh, H. (2022). Dictionary learning for sparse representation of signals with hidden markov model dependency. *Digital Signal Processing*, 103420.
- Akhonda, M.A., Levin-Schwartz, Y., Bhinge, S., Calhoun, V.D., Adali, T. (2018). Consecutive independence and correlation transform for multimodal fusion: Application to EEG and fMRI data. *2018 ieee international conference on acoustics, speech and signal processing (icassp)* (pp. 2311–2315).
- Allison, B.Z., & Polich, J. (2008). Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biological psychology*, 77(3), 277–283.
- 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.
- Asadzadeh, S., Rezaii, T.Y., Beheshti, S., Delpak, A., Meshgini, S. (2020). A systematic review of EEG source localization techniques and their applications on diagnosis of brain abnormalities. *Journal of Neuroscience Methods*, 339, 108740.
- Bakdash, J.Z., & Marusich, L.R. (2017). Repeated measures correlation. *Frontiers in psychology*, 8, 456.
- Becker, H., Albera, L., Comon, P., Gribonval, R., Merlet, I. (2014). Fast, variation-based methods for the analysis of extended brain sources. *2014 22nd european signal processing conference (eusipco)* (pp. 41–45).
- Becker, H., Albera, L., Comon, P., Haardt, M., Birot, G., Wendling, F., ... Merlet, I. (2014). EEG extended source localization:

- tensor-based vs. conventional methods. *NeuroImage*, 96, 143–157.
- Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behavioural brain research*, 125(1-2), 279–284.
- Berti, S., & Schröger, E. (2003). Working memory controls involuntary attention switching: evidence from an auditory distraction paradigm. *European Journal of Neuroscience*, 17(5), 1119–1122.
- Boehler, C., Schoenfeld, M., Heinze, H.-J., Hopf, J.-M. (2008). Rapid recurrent processing gates awareness in primary visual cortex. *Proceedings of the National Academy of Sciences*, 105(25), 8742–8747.
- Boonyakitanont, P., Gabrielson, B., Belyaeva, I., Olikkal, P., Wang, Y.-P., Wilson, T., ... Tulay, A. (2022). An ICA-based framework for joint analysis of cognitive scores and MEG event-related fields. *44th annual international conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. *Accepted*.
- Bro, R., & Kiers, H.A. (2003). A new efficient method for determining the number of components in PARAFAC models. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 17(5), 274–286.
- Bro, R., & Smilde, A.K. (2003). Centering and scaling in component analysis. *Journal of Chemometrics*, 17(1), 16–33.
- Bruneau, N., Roux, S., Guerin, P., Barthelemy, C., Lelord, G. (1997). Temporal prominence of auditory evoked potentials (N1 wave) in 4-8-year-old children. *Psychophysiology*, 34(1), 32–38.
- Calhoun, V.D., & Adali, T. (2012). Multi-subject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery. *IEEE reviews in biomedical engineering*, 5, 60–73.
- Calhoun, V.D., Adali, T., Pearlson, G.D., Kiehl, K.A. (2006). Neuronal chronometry of target detection: fusion of hemodynamic and event-related potential data. *Neuroimage*, 30(2), 544–553.
- Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Human brain mapping*, 14(3), 140–151.
- Calhoun, V.D., Liu, J., Adali, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage*, 45(1), S163–S172.
- Cardy, J.E.O., Ferrari, P., Flagg, E.J., Roberts, W., Roberts, T.P. (2004). Prominence of M50 auditory evoked response over m100 in childhood and autism. *Neuroreport*, 15(12), 1867–1870.
- Carroll, J.D., & Chang, J.-J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of “eckart-young” decomposition. *Psychometrika*, 35(3), 283–319.
- Chatzichristos, C., Kofidis, E., Morante, M., Theodoridis, S. (2019). Blind fMRI source unmixing via higher-order tensor decompositions. *Journal of neuroscience methods*, 315, 17–47.
- Chatzichristos, C., Kofidis, E., Van Paesschen, W., De Lathauwer, L., Theodoridis, S., Van Huffel, S. (2022). Early soft and flexible fusion of

- electroencephalography and functional magnetic resonance imaging via double coupled matrix tensor factorization for multisubject group analysis. *Human brain mapping*, 43(4), 1231–1255.
- Cichocki, A., Lee, N., Oseledets, I., Phan, A.-H., Zhao, Q., Mandic, D.P. (2016). Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions. *Foundations and Trends® in Machine Learning*, 9(4-5), 249–429.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.
- Cong, F., Lin, Q.-H., Kuang, L.-D., Gong, X.-F., Astikainen, P., Ristaniemi, T. (2015). Tensor decomposition of EEG signals: a brief review. *Journal of neuroscience methods*, 248, 59–69.
- Cong, F., Phan, A.H., Zhao, Q., Huttunen-Scott, T., Kaartinen, J., Ristaniemi, T., ... Cichocki, A. (2012). Benefits of multi-domain feature of mismatch negativity extracted by non-negative tensor factorization from EEG collected by low-density array. *International journal of neural systems*, 22(06), 1250025.
- Conners, C.K. (2008). *Conners third edition (Conners 3)*. Los Angeles, CA: Western Psychological Services.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1), 9–21.
- De Vos, M., Vergult, A., De Lathauwer, L., De Clercq, W., Van Huffel, S., Dupont, P., ... Van Paesschen, W. (2007). Canonical decomposition of ictal scalp EEG reliably detects the seizure onset zone. *NeuroImage*, 37(3), 844–854.
- Dikmen, S.S., Bauer, P.J., Weintraub, S., Mungas, D., Slotkin, J., Beaumont, J.L., ... Heaton, R.K. (2014). Measuring episodic memory across the lifespan: NIH toolbox picture sequence memory test. *Journal of the International Neuropsychological Society*, 20(6), 611–619.
- Dyke, F.B., Leiker, A.M., Grand, K.F., Godwin, M.M., Thompson, A.G., Rietschel, J.C., ... Miller, M.W. (2015). The efficacy of auditory probes in indexing cognitive workload is dependent on stimulus complexity. *International Journal of Psychophysiology*, 95(1), 56–62.
- Edgar, J.C., Lanza, M.R., Daina, A.B., Monroe, J.F., Khan, S.Y., Blaskey, L., ... others (2014). Missing and delayed auditory responses in young and older children with autism spectrum disorders. *Frontiers in human neuroscience*, 8, 417.
- Efron, B., Halloran, E., Holmes, S. (1996). Bootstrap confidence levels for phylogenetic trees. *Proceedings of the National Academy of Sciences*, 93(14), 7085–7090.
- Fitz, H., & Chang, F. (2019). Language ERPs reflect learning through prediction error propagation. *Cognitive Psychology*, 111, 15–52.
- Friedman, D., & Johnson Jr, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy research and technique*, 51(1), 6–28.
- Gabrielson, B., Akhonda, M.A., Bhinge, S., Brooks, J., Long, Q., Adali, T. (2020). Joint-IVA for identification of discriminating

- features in EEG: Application to a driving study. *Biomedical Signal Processing and Control*, 61, 101948.
- Garcés, P., López-Sanz, D., Maestú, F., Pereda, E. (2017). Choice of magnetometers and gradiometers after signal space separation. *Sensors*, 17(12), 2926.
- Ghani, U., Signal, N., Niazi, I.K., Taylor, D. (2020). ERP based measures of cognitive workload: A review. *Neuroscience & Biobehavioral Reviews*, 118, 18–26.
- Ghani, U., Signal, N., Niazi, I.K., Taylor, D. (2021). Efficacy of a single-task ERP measure to evaluate cognitive workload during a novel exergame. *Frontiers in Human Neuroscience*, 519.
- Graf, P., & Schacter, D.L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, memory, and cognition*, 11(3), 501.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., ... Hämäläinen, M.S. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7(267), 1–13.
- 10.3389/fnins.2013.00267
- Groppe, D.M., Urbach, T.P., Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, 48(12), 1711–1725.
- Halgren, E., Dhond, R.P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J.D., Dale, A.M. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage*, 17(3), 1101–1116.
- Halkidi, M., Batistakis, Y., Vazirgiannis, M. (2002a). Clustering validity checking methods: Part ii. *ACM Sigmod Record*, 31(3), 19–27.
- Halkidi, M., Batistakis, Y., Vazirgiannis, M. (2002b). Cluster validity methods: part i. *ACM Sigmod Record*, 31(2), 40–45.
- Hämäläinen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O.V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of modern Physics*, 65(2), 413.
- Harshman, R.A., et al. (1970). Foundations of the PARAFAC procedure: Models and conditions for an” explanatory” multimodal factor analysis.
- Harshman, R.A., et al. (1972). PARAFAC2: Mathematical and technical notes. *UCLA working papers in phonetics*, 22(3044), 122215.
- He, B., & Liu, Z. (2008). Multimodal functional neuroimaging: integrating functional MRI and EEG/MEG. *IEEE reviews in biomedical engineering*, 1, 23–40.
- Helwig, N.E., & Snodgrass, M.A. (2019). Exploring individual and group differences in latent brain networks using cross-validated simultaneous component analysis. *NeuroImage*, 201, 116019.
- Hernández, D., Puupponen, A., Jantunen, T. (2022). The contribution of event-related potentials to the understanding of sign language processing and production in the

brain: Experimental evidence and future directions. *Frontiers in Communication*, 40.

- Horat, S.K., Herrmann, F.R., Favre, G., Terzis, J., Debatisse, D., Merlo, M.C., Missonnier, P. (2016). Assessment of mental workload: a new electrophysiological method based on intra-block averaging of ERP amplitudes. *Neuropsychologia*, 82, 11–17.
- Husson, F., Josse, J., Pagès, J. (2010). Principal component methods—hierarchical clustering, partitional clustering. *Why would we need to choose for visualizing data*.
- Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5), 411–430.
- Ikeda, S., & Toyama, K. (2000). Independent component analysis for noisy data—MEG data analysis. *Neural Networks*, 13(10), 1063–1074.
- Jaccard, P. (1912). The distribution of the flora in the alpine zone. 1. *New phytologist*, 11(2), 37–50.
- Jin, R., Dontaraju, K.K., Kim, S.-J., Akhonda, M.A.B.S., Adali, T. (2020). Dictionary learning-based fMRI data analysis for capturing common and individual neural activation maps. *IEEE Journal of Selected Topics in Signal Processing*, 14(6), 1265–1279.
- Jonmohamadi, Y., Muthukumaraswamy, S., Chen, J., Roberts, J., Crawford, R., Pandey, A. (2020). Extraction of common task features in EEG-fMRI data using coupled tensor-tensor decomposition. *Brain Topography*, 33(5), 636–650.
- Jung, T.-P., Makeig, S., McKeown, M.J., Bell, A.J., Lee, T.-W., Sejnowski, T.J. (2001). Imaging brain dynamics using independent component analysis. *Proceedings of the IEEE*, 89(7), 1107–1122.
- Kinney-Lang, E., Ebied, A., Auyeung, B., Chin, R.F., Escudero, J. (2019). Introducing the joint EEG-development inference (JEDI) model: A multi-way, data fusion approach for estimating paediatric developmental scores via eeg. *IEEE transactions on neural systems and rehabilitation engineering*, 27(3), 348–357.
- Kinney-Lang, E., Spyrou, L., Ebied, A., Chin, R., Escudero, J. (2017). Elucidating age-specific patterns from background electroencephalogram pediatric datasets via PARAFAC. *2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 3797–3800).
- Kinney-Lang, E., Spyrou, L., Ebied, A., Chin, R.F., Escudero, J. (2018). Tensor-driven extraction of developmental features from varying paediatric EEG datasets. *Journal of Neural Engineering*, 15(4), 046024.
- Kolda, T.G., & Bader, B.W. (2009). Tensor decompositions and applications. *SIAM review*, 51(3), 455–500.
- Kossaifi, J., Panagakis, Y., Anandkumar, A., Pantic, M. (2019). Tensorly: Tensor learning in python. *Journal of Machine Learning Research*, 20(26), 1-6. Retrieved from <http://jmlr.org/papers/v20/18-277.html>
- Kotecha, R., Pardos, M., Wang, Y., Wu, T., Horn, P., Brown, D., ... Xiang, J. (2009). Modeling the developmental patterns of auditory evoked magnetic fields in children. *PLoS One*, 4(3), e4811.

- Kramer, A.F., Trejo, L.J., Humphrey, D. (1995). Assessment of mental workload with task-irrelevant auditory probes. *Biological psychology*, 40(1-2), 83–100.
- Kroonenberg, P.M., & De Leeuw, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika*, 45(1), 69–97.
- Kruskal, J.B. (1977). Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear algebra and its applications*, 18(2), 95–138.
- Kuang, L.-D., Lin, Q.-H., Gong, X.-F., Cong, F., Wang, Y.-P., Calhoun, V.D. (2019). Shift-invariant canonical polyadic decomposition of complex-valued multi-subject fMRI data with a phase sparsity constraint. *IEEE transactions on medical imaging*, 39(4), 844–853.
- Kutas, M., & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, 4(12), 463–470.
- Labounek, R., Bridwell, D.A., Mareček, R., Lamoš, M., Mikl, M., Slavíček, T., ... others (2018). Stable scalp EEG spatio-spectral patterns across paradigms estimated by group ICA. *Brain topography*, 31(1), 76–89.
- Lahat, D., Adali, T., Jutten, C. (2015). Multi-modal data fusion: an overview of methods, challenges, and prospects. *Proceedings of the IEEE*, 103(9), 1449–1477.
- Liotti, M., Pliszka, S.R., Higgins, K., Perez III, R., Semrud-Clikeman, M. (2010). Evidence for specificity of ERP abnormalities during response inhibition in ADHD children: A comparison with reading disorder children without ADHD. *Brain and cognition*, 72(2), 228–237.
- Liu, W., Wang, X., Xu, J., Chang, Y., Hämäläinen, T., Cong, F. (2021). Identifying oscillatory hyperconnectivity and hypoconnectivity networks in major depression using coupled tensor decomposition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29, 1895–1904.
- Luck, S.J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, 54(1), 146–157.
- Marinković, K. (2004). Spatiotemporal dynamics of word processing in the human cortex. *The Neuroscientist*, 10(2), 142–152.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, 164(1), 177–190.
- Mørup, M., & Hansen, L.K. (2009). Automatic relevance determination for multi-way models. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 23(7-8), 352–363.
- Mørup, M., Hansen, L.K., Herrmann, C.S., Parnas, J., Arnfred, S.M. (2006). Parallel factor analysis as an exploratory tool for wavelet transformed event-related EEG. *NeuroImage*, 29(3), 938–947.
- Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? *Journal of classification*, 31(3), 274–295.

- Nichols, T.E., & Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, 15(1), 1–25.
- Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological methods*, 8(4), 434.
- Pantazis, D., Nichols, T.E., Baillet, S., Leahy, R.M. (2003). Spatiotemporal localization of significant activation in MEG using permutation tests. *Biennial international conference on information processing in medical imaging* (pp. 512–523).
- Pfefferbaum, A., Ford, J.M., Weller, B.J., Kopell, B.S. (1985). ERPs to response production and inhibition. *Electroencephalography and clinical neurophysiology*, 60(5), 423–434.
- Pinner, J.F., Coffman, B.A., Stephen, J.M. (2020). Covariation between brain function (MEG) and structure (DTI) differentiates adolescents with fetal alcohol spectrum disorder from typically developing controls. *Neuroscience*, 449, 74–87.
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology*, 118(10), 2128–2148.
- Ponton, C.W., Eggermont, J.J., Kwong, B., Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical neurophysiology*, 111(2), 220–236.
- Rugg, M.D., Mark, R.E., Walla, P., Schloerscheidt, A.M., Birch, C.S., Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, 392(6676), 595–598.
- Salman, M.S., Du, Y., Lin, D., Fu, Z., Fedorov, A., Damaraju, E., ... others (2019). Group ICA for identifying biomarkers in schizophrenia: ‘adaptive’ networks via spatially constrained ICA show more sensitivity to group differences than spatio-temporal regression. *NeuroImage: Clinical*, 22, 101747.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461–464.
- Sidiropoulos, N.D., & Bro, R. (2000). On the uniqueness of multilinear decomposition of N-way arrays. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 14(3), 229–239.
- Sokhadze, E.M., Casanova, M.F., Casanova, E.L., Lamina, E., Kelly, D.P., Khachidze, I. (2017). Event-related potentials (ERP) in cognitive neuroscience research and applications. *NeuroRegulation*, 4(1), 14–14.
- Stephen, J.M., Coffman, B.A., Jung, R.E., Bustillo, J.R., Aine, C., Calhoun, V.D. (2013). Using joint ICA to link function and structure using MEG and DTI in schizophrenia. *Neuroimage*, 83, 418–430.
- Stephen, J.M., Solis, I., Janowich, J., Stern, M., Frenzel, M.R., Eastman, J.A., ... others (2021). The developmental chronnectogenomics (Dev-CoG) study: A multimodal study on the developing brain. *NeuroImage*, 225, 117438.
- Stoica, P., & Selen, Y. (2004). Model-order selection: a review of information criterion rules. *IEEE Signal Processing Magazine*, 21(4), 36–47.
- Sullivan, G.M., & Feinn, R. (2012). Using effect size—or why the p value is not enough. *Journal of graduate medical education*, 4(3), 279–282.

- Sur, S., & Sinha, V.K. (2009). Event-related potential: An overview. *Industrial psychiatry journal*, 18(1), 70.
- Székel, 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.
- Tangwiriyaakul, C., Premoli, I., Spyrou, L., Chin, R.F., Escudero, J., Richardson, M.P. (2019). Tensor decomposition of TMS-induced EEG oscillations reveals data-driven profiles of antiepileptic drug effects. *Scientific reports*, 9(1), 1–11.
- Taulu, S., & Kajola, M. (2005). Presentation of electromagnetic multichannel data: the signal space separation method. *Journal of Applied Physics*, 97(12), 124905.
- Taulu, S., & Simola, J. (2006). Spatiotemporal signal space separation method for rejecting nearby interference in meg measurements. *Physics in Medicine & Biology*, 51(7), 1759.
- Team, R.C., et al. (2013). R: A language and environment for statistical computing.
- Tomasi, G., & Bro, R. (2005). PARAFAC and missing values. *Chemometrics and Intelligent Laboratory Systems*, 75(2), 163–180.
- Vanderperren, K., Mijović, B., Novitskiy, N., Vanrumste, B., Stiers, P., Van den Bergh, B.R., ... others (2013). Single trial ERP reading based on parallel factor analysis. *Psychophysiology*, 50(1), 97–110.
- van Meel, C.S., Heslenfeld, D.J., Oosterlaan, J., Sergeant, J.A. (2007). Adaptive control deficits in attention-deficit/hyperactivity disorder (ADHD): the role of error processing. *Psychiatry research*, 151(3), 211–220.
- Vogel, E.K., Luck, S.J., Shapiro, K.L. (1998). Electrophysiological evidence for a post-perceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, 24(6), 1656.
- Wang, D., Zhu, Y., Ristaniemi, T., Cong, F. (2018). Extracting multi-mode ERP features using fifth-order nonnegative tensor decomposition. *Journal of neuroscience methods*, 308, 240–247.
- Wang, X., Liu, W., Toivainen, P., Ristaniemi, T., Cong, F. (2020). Group analysis of ongoing EEG data based on fast double-coupled nonnegative tensor decomposition. *Journal of neuroscience methods*, 330, 108502.
- Wechsler, D. (2011). *WASI-II: Wechsler abbreviated scale of intelligence*. PsychCorp.
- Weintraub, S., Bauer, P.J., Zelazo, P.D., Wallner-Allen, K., Dikmen, S.S., Heaton, R.K., ... others (2013). I. NIH Toolbox Cognition Battery (CB): introduction and pediatric data. *Monographs of the Society for Research in Child Development*, 78(4), 1–15.
- Westfall, P.H., & Young, S.S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment* (Vol. 279). John Wiley & Sons.
- Wu, W., Wu, C., Gao, S., Liu, B., Li, Y., Gao, X. (2014). Bayesian estimation of ERP components from multicondition and multichannel EEG. *NeuroImage*, 88, 319–339.

- Zhang, W., & Luck, S.J. (2009). Feature-based attention modulates feedforward visual processing. *Nature neuroscience*, *12*(1), 24–25.
- Zhu, Y., Liu, J., Ye, C., Mathiak, K., Astikainen, P., Ristaniemi, T., Cong, F. (2020). Discovering dynamic task-modulated functional networks with specific spectral modes using MEG. *NeuroImage*, *218*, 116924.