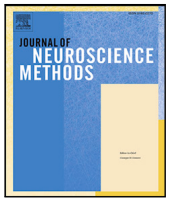
Journal of Neuroscience Methods 385 (2023) 109768

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth



Single-trial-based temporal principal component analysis on extracting event-related potentials of interest for an individual subject

Guanghui Zhang [a,b,*], Xueyan Li [c], Yingzhi Lu [d], Timo Tiihonen [b], Zheng Chang [b], Fengyu Cong [a,b,e,f,*]

[a] School of Biomedical Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, 116024, China
[b] Faculty of Information Technology, University of Jyväskylä, Jyväskylä, 40014, Finland
[c] School of Foreign Languages, Dalian University of Technology, Dalian, 116024, China
[d] School of Psychology, Shanghai University of Sport, Shanghai, 200438, China
[e] School of Artificial Intelligence, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, 116024, China
[f] Key Laboratory of Integrated Circuit and Biomedical Electronic System, Dalian University of Technology, Dalian, 116024, China

ARTICLE INFO

Keywords:
Event-related potentials
Principal component analysis
Single-trial analysis
Back-projection
Individual subject

ABSTRACT

Background: Temporal principal component analysis (tPCA) has been widely used to extract event-related potentials (ERPs) at group level of multiple subjects ERP data and it assumes that the underlying factor loading is fixed across participants. However, such assumption may fail to work if latency and phase for one ERP vary considerably across participants. Furthermore, effect of number of trials on tPCA decomposition has not been systematically examined as well, especially for within-subject PCA.
New method: We reanalyzed a real ERP data of an emotional experiment using tPCA to extract N2 and P2 from single-trial EEG of an individual. We also explored influence of the number of trials (consecutively increased from 10 to 42 trials) on PCA decomposition by comparing temporal correlation, the statistical result, Cronbach's alpha, spatial correlation of both N2 and P2 for the proposed method with the conventional time-domain analysis, trial-averaged group PCA, and single-trial-based group PCA.
Results: The results of the proposed method can enhance spatial and temporal consistency. We could obtain stable N2 with few trials (about 20) for the proposed method, but, for P2, approximately 30 trials were needed for all methods.
Comparison with Existing Method(s): About 30 trials for N2 were required and the reconstructed P2 and N2 were poor correlated across participants for the other three methods.
Conclusion: The proposed approach may efficiently capture variability of one ERP from an individual that cannot be extracted by group PCA analysis.

1. Introduction

Event-related potentials (ERPs) are considerably small compared with other signals (e.g., noise, spontaneous EEG, etc.) and they are often temporally and spatially mixed with other signals to some extent (Beauducel et al., 2000; Beauducel and Debener, 2003). Imaging that several sources are active at a specific time point, the voltage at that time point will be the sum of those sources (i.e., temporal overlap issue). Similarly, the voltage of collected signals at a specific electrode site is also contribution of multiple sources (i.e., spatial overlap issue) (Scharf et al., 2022). Although signal-to-noise (SNR) is improved considerably after some preprocessing steps (e.g., filtering, artifact correction via independent component analysis), the preprocessed ERP signals still contain considerable noise and the overlapping problem for the ERP components still exists. To address this, temporal principal component analysis (tPCA) has been widely used to estimate those temporally and/or spatially mixed ERP components (Fogarty et al., 2020; Male and Gouldthorp, 2020; Bonmassar et al., 2020; Dien, 2012, 1998; Dien et al., 2007; Kayser and Tenke, 2003, 2006a,b).

* Corresponding authors at: School of Biomedical Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, 116024, China.
E-mail addresses: zhang.guanghui@foxmail.com (G. Zhang), cong@dlut.edu.cn (F. Cong).

https://doi.org/10.1016/j.jneumeth.2022.109768
Received 26 March 2022; Received in revised form 26 November 2022; Accepted 12 December 2022
Available online 15 December 2022

1.1. Model and use of temporal PCA, and potential problems of its current application

As shown in Fig. 1a, matrix $\mathbf{Z} \in \mathfrak{R}^{T\times M}$ formed by preprocessed ERP data from multiple participants with T variables (i.e., time samples) and M observations (i.e., combinations of electrodes, conditions, and participants) can be represented using blind source separation (BSS) model (see Fig. 1) (Cong et al., 2011a,b; Makeig et al., 1997, 1999):

$$\mathbf{Z} = \mathbf{a}_1 \circ \mathbf{s}_1 + \cdots + \mathbf{a}_r \circ \mathbf{s}_r + \cdots + \mathbf{a}_R \circ \mathbf{s}_R = \mathbf{AS}, \quad (1)$$

$\mathbf{A} \in \mathfrak{R}^{T\times R}$ is coefficient matrix and the magnitude of each element for its rth column $\mathbf{a}_r$ is the coefficient for rth source at a specific time point (see Fig. 1b) which determines signal amplitude at that time point. $\mathbf{S} \in \mathfrak{R}^{R\times M}$ is source matrix and its each row $\mathbf{s}_r$ is spatial distribution which is related to stimulus-locked activity, spontaneous EEG, or noise (see Fig. 1c). Therefore, the amplitude $\mathbf{z}(t_0) \in \mathfrak{R}^{1\times M}$ for all channels at time point t_0 is the sum of multiplication between the coefficient $a_r(t_0) \in \mathfrak{R}^{1\times 1}$ and the corresponding spatial source $\mathbf{s}_r \in \mathfrak{R}^{1\times M}$ or $\mathbf{z}(t_0) = \sum_{r=1}^{R} a_r(t_0) \cdot \mathbf{s}_r$ (see Fig. 1). R is the number of sources. " $\circ$ " represents outer-product between two vectors. "$\cdot$" represents dot-product.

In practice, both $\mathbf{s}_r$ and $\mathbf{a}_r$ are unknown. Fortunately, some algorithms, for example, singular value decomposition (SVD) and Promax rotation, have been developed to estimate them. Thus, the matrix $\mathbf{Z}$ can be approximately estimated by the sums of outer-product of several factors related to stimulus onset[1]:

$$\mathbf{Z} \approx \mathbf{u}_1 \circ \mathbf{y}_1 + \cdots + \mathbf{u}_k \circ \mathbf{y}_k + \cdots + \mathbf{u}_K \circ \mathbf{y}_K = \mathbf{UY}, \quad (2)$$

$\mathbf{U} \in \mathfrak{R}^{T\times K} = [\mathbf{u}_1, \ldots, \mathbf{u}_k, \ldots, \mathbf{u}_K]$ and it is the estimate of $\mathbf{A}$ which is known as factor loading or factor coefficient. Each row of $\mathbf{U}$ contains T coefficients corresponding to T time points and the columns are factor variances that are in descending order. $\mathbf{Y} \in \mathfrak{R}^{K\times M} = [\mathbf{y}_1, \ldots, \mathbf{y}_k, \ldots, \mathbf{y}_K]^T$ represents $\mathbf{Z}$ in the factor (principal component) space, and its rows correspond to factors and columns are observations. That is, $\mathbf{y}_k$ estimates $\mathbf{s}_r$ and it is topographical distribution of kth factor which is also known as factor score. K, the number of factors, is estimated by some methods, for example, cumulative explained variance, and it is usually smaller than the number of sources R because we are only interested in the sources related to stimulus-onset activities. The basic mathematical procedure for tPCA from the view of BSS can be found in Appendix A and more details about PCA algorithm based on SVD are available in Abdi and Williams (2010), and more information about Promax rotation and other rotation methods can be found in Richman (1986), Hendrickson and White (1964).

ERP components are usually extracted from ERP data[2] of all participants under all conditions by means of tPCA (Fogarty et al., 2020; Male and Gouldthorp, 2020; Bonmassar et al., 2020; Dien, 2012, 1998; Dien et al., 2007; Kayser and Tenke, 2003, 2006a,b). In detail, ERP datasets are first concatenated across electrodes, conditions, and participants to form a matrix. The size of the matrix is time samples multiplied by the combinations of electrodes, conditions, and participants. PCA is then conducted to decompose this matrix into several factors. The corresponding PCA procedure refers to trial-averaged group PCA (TGPCA) in this study and it assumes (i.e., measurement invariance) that all subjects and conditions share the same time-course. That is, the underlying factor loading is fixed across participants across conditions:

$$\mathbf{u}_k = \mathbf{u}(1)_k = \cdots = \mathbf{u}(p)_k = \cdots = \mathbf{u}(P)_k, \quad (3)$$

k is the order of a particular factor, p is the order of a particular participant, and P is the number of participants.

However, even if the experimental environment is well controlled, the responses for the stimuli are not exactly same across participants because, for example, brain structure substantially varies among people at least. Another problem is that, because single-trial EEG are first averaged for each condition before PCA performance, the potential variability of ERP of interest in single-trial EEG among participants may be averaged out during such an averaging procedure.

To study the potential variability of ERP components that may be averaged out during averaging procedure, some researchers have performed PCA on the single-trial EEG data for all conditions for all participants to isolate ERP components (Rushby and Barry, 2009; MacDonald and Barry, 2017; MacDonald et al., 2015; Rushby et al., 2005). We name this procedure as single-trial-based group PCA (SGPCA). Similarly, Barry and his colleagues have tried to use a separate PCA to explore the latency difference of ERP components of interest between conditions among all participants (Barry et al., 2016). Note that the core ideas of the above works are similar to TGPCA.

1.2. Effects of the number of trial on the study of ERP

It is widely accepted that SNR of ERPs linearly increases with the increase of square root of the number of trials N, i.e., SNR $\sim \sqrt{N}$ (Luck, 2014). That is, the more number of trials, the higher SNR. However, more trials may lead to fatigue, influencing the task performance of participants and resulting in the enhancement of unexpected signals (e.g., alpha band). Therefore, it is critical and necessary to keep a balance between data quality and experimental time by optimizing the trial numbers involved in experiments.

Several early studies have provided data-driven evidence and guidelines on how many trials should be contained in specific ERP experimental paradigms. For example, N2, P3, N1, P1, P2, error-related negativity (ERN), Late positive potential (LPP), N170, and error positivity (Pe) have been systematically investigated (Clayson, 2020; Larson et al., 2010; Huffmeijer et al., 2014; Pontifex et al., 2010; Olvet and Hajcak, 2009; Fischer et al., 2017; Thigpen et al., 2017; Cohen and Polich, 1997; Rietdijk et al., 2014; Kleene et al., 2021; Clayson and Larson, 2013). Most of these studies have quantified how to use minimum number of trials to obtain the stable and reliable measures of ERPs of interest for conventional time-domain analysis. Although Steele et al. (2016) has studied how to use minimal number of trials to obtain a reliable ERN for PCA decomposition, the impact of trial number on the PCA decomposition of single-subject (within-subject PCA or WPCA) was poorly investigated and the comparison between group PCA and WPCA was not fully explored either.

1.3. Organization of the present study

To address the above mentioned problems, we divided the current study into two main parts. In the first part, we proposed a method (WPCA) to study the variability across individuals that extracted two ERP components from single-trial EEG of an individual by means of tPCA in an emotional ERP experiment, and the following steps were involved (see Fig. 2a). Firstly, single-trial EEG data for each subject with all the remaining trials were stacked along electrodes to form a matrix separately. Secondly, tPCA and Promax rotation were employed to decompose the matrices separately. Next, the factors associated with an ERP of interest for each subject were selected and projected onto the electrode fields (i.e., in microvolt) that reconstructed the single-trial EEG waveform. Back-projection procedure can refer to Fig. 2b. Afterwards, the reconstructed waveforms of single-trial EEG were averaged separately under each experimental condition for each subject. Finally, the mean amplitudes of the ERPs were quantified within a predefined time-window at some electrodes to make comparisons between conditions.

[1] Term "factor" is used here to represent principal component extracted by PCA and Promax rotation, which is to avoid confusion with ERP component. K is also used to avoid confusion with R.

[2] We called the averages of single-trial EEG as "ERP data" in this study because noise signals are assumed to be totally removed during averaging procedure and only information evoked by stimulus is contained.

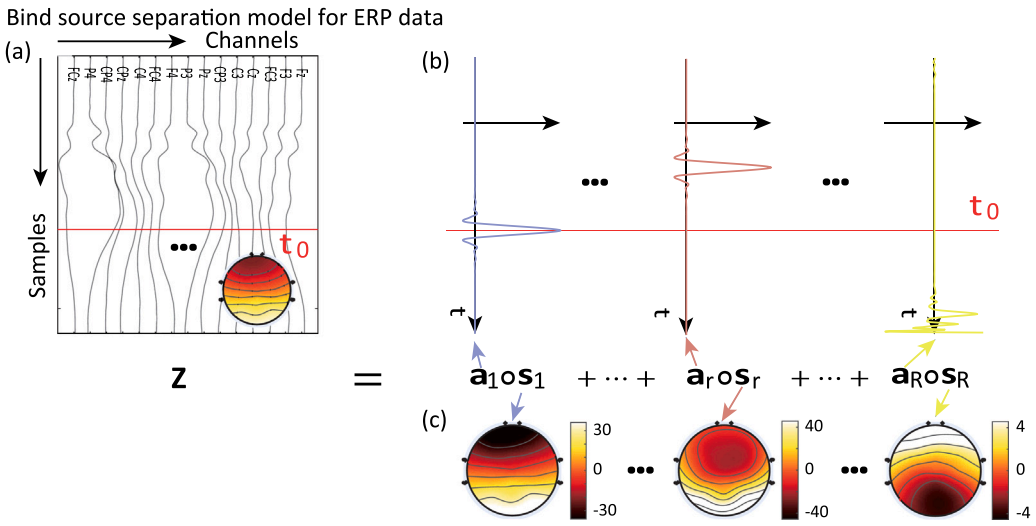


Fig. 1. Blind source separation (BSS) model for ERP/EEG data (decomposition of temporal principal component analysis). For the preprocessed single-trial EEG or ERP data $\mathbf{Z} \in \Re^{T\times M}$, it can be considered to be the sum of outer-product between the coefficients $\mathbf{a}_r \in \Re^{T\times 1}$ and spatial source $\mathbf{s}_r \in \Re^{1\times M}$, that is, $\mathbf{Z} = \mathbf{a}_1 \circ \mathbf{s}_1 + \cdots + \mathbf{a}_r \circ \mathbf{s}_r + \cdots + \mathbf{a}_R \circ \mathbf{s}_R$. R is the number of sources. For the scalp data at point t_0 (red line), i.e., the topographic distribution $\mathbf{z}(t_0) \in \Re^{1\times M}$ is regarded to be multiplication between coefficient for the first source at time point t_0 and the first spatial source $\mathbf{s}_1$: $\mathbf{z}(t_0) = a_1(t_0) \cdot \mathbf{s}_1$ if the coefficients for the other sources at this time point are zero (i.e., $a_1(t_0) = \cdots = a_R(t_0) = 0$).

The other three methods, namely, conventional time-domain analysis, TGPCA, and SGPCA, were also used to quantify ERP components of interest. We examined the differences among four used methods by comparing the corresponding ERP waveforms, topographical distributions, spatial correlation coefficient (CCs), Cronbach's alpha, and statistical analysis results. Note that the comparison is based on the data for all remaining trials in this part.

In the second part, we investigated whether the number of trials (consecutively increased from 10 to 42) has a significant effect on three PCA-based methods and conventional time-domain analysis compared with the yields of all the remaining trials. We first measured temporal CCs of P2/N2 waveforms for all participants between the averages with small number of trials (e.g., 10 trials) and the grand average with all remaining trials, and afterwards, we computed the Cronbach's alpha for waveforms of the increased trials. Computing consistency of topographies among participants for all cases as well, and we finally examined the statistical results for all cases.

2. Data description and method

In this section, we briefly described a published ERP experiment (Lu et al., 2016) and introduced PCA procedures in step-by-step. The data and Matlab code scripts used in this study are available at: http://zhangg.net/publications/.

2.1. Participants

22 healthy undergraduate students from Shanghai University of Sport participated in the study as paid volunteers. This experiment had been approved by the local ethics committee and all participants had signed their written informed consent before the experiment. The age of participants was 22.05 ± 1.53 years old (from 20 to 24), including 10 males and 12 females.

2.2. Stimuli and task

2.2.1. Stimuli

We used 60 pictures in this experiment including 20 disgusting pictures, 20 fearful pictures, and 20 neutral pictures. The disgusting items are, for example, dead animals, feces, maggots, and moldy pictures. The fear items are those pictures that may make people feel fearful, such as guns, snakes, accidents, dogs, etc. Neutral items are those pictures showing peaceful scenes, for example, household objects and tranquil animals. More details about how to select those pictures can be found in Lu et al. (2016).

2.2.2. Task

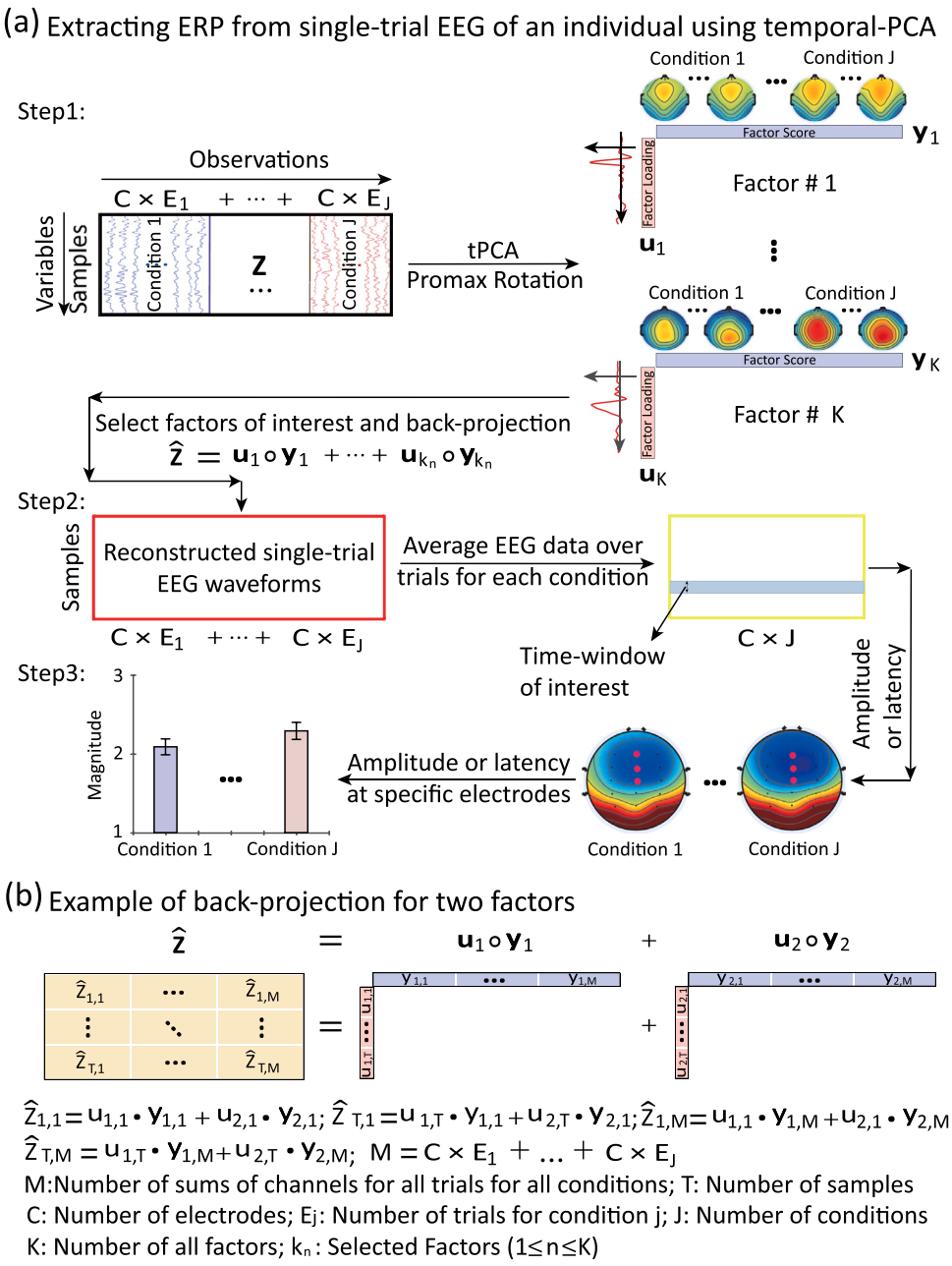
Before starting the ERP experiment, participants were required to fill out the Beck Anxiety Inventory and Self-rating Anxiety Scale to measure the state of anxiety that may reflect typical reactions to disgusting and fearful items. Participant were allowed to do the following ERP task if both measurements were within the normal range.

ERP part is a modified oddball distinction task in which participants were required to respond to different emotional pictures. At the beginning, participants had a practice with ten trials to better understand the procedure. At the beginning of each trial, a black fixation (fixed time is 300 ms) was displayed in the center of the computer screen, and then a blank white screen was displayed with a random time-window between 500 ms and 1500 ms. Afterward, the stimuli were displayed to participants, and then disappeared after 1000 ms or terminated by key press ('J' for the deviant pictures and 'F' for standard ones). Another blank screen with 1000 ms followed the response before the next trial.

We used a blocked-design method for fear and disgust induction to avoid contamination across emotional emotions for random designs, that is, we featured a single category of emotional (deviant) stimuli in each block. Six blocks featuring disgusting, neutral, and standard stimuli were included in the experiment. Another six blocks with fearful, neutral, and standard stimuli were also included. In each block, there were 100 trials that comprised 70 standard stimuli and 30 deviant stimuli. A natural scene of a chair was used for standard stimuli. Deviant stimuli included 10 moderately/extremely disgusting pictures, 10 moderately/extremely fearful pictures, and 10 neutral pictures, separately. Participants received accuracy (ACC) rates for both standard and deviant stimuli at the end of each block. More details about the task can be found in Lu et al. (2016).

2.3. EEG recording and preprocessing

Nineteen electrodes (F3, FC3, C3, CP3, P3, Fz, FCz, Cz, CPz, Pz, F4, FC4, C4, CP4, P4, TP9, TP10, VEOG, and HEOG) were used to record EEG based on the international 10–20 system with 1000 Hz sampling rate. EEG recordings were referenced at FCz (Brain Products GmbH, Germany). The impedance was less than 5 kΩ for each sensor of each subject.



$$\hat{Z}_{1,1}=u_{1,1}\cdot y_{1,1}+u_{2,1}\cdot y_{2,1};\ \hat{Z}_{T,1}=u_{1,T}\cdot y_{1,1}+u_{2,T}\cdot y_{2,1};\ \hat{Z}_{1,M}=u_{1,1}\cdot y_{1,M}+u_{2,1}\cdot y_{2,M}$$

$$\hat{Z}_{T,M}=u_{1,T}\cdot y_{1,M}+u_{2,T}\cdot y_{2,M};\quad M=C\times E_1+\ldots+C\times E_J$$

M:Number of sums of channels for all trials for all conditions; T: Number of samples
C: Number of electrodes; Ej: Number of trials for condition j; J: Number of conditions
K: Number of all factors; k_n: Selected Factors ($1\le n\le K$)

Fig. 2. Pipeline for extracting ERP of interest from single-trial EEG of an individual using temporal principal component analysis (tPCA) (a), and example of back-projection for two factors (b). (a) Step 1: Single-trial EEG for multiple conditions for an individual are arranged into a matrix $\mathbf{Z}\in\Re^{T\times M}$ along electrodes (numbers of trials for condition 1, ···, and J are E_1, ···, and E_J, respectively; C is number of electrodes), and the matrix is decomposed into R factors using tPCA; Step 2: Selecting and projecting factors of interest back onto electrode fields for reconstructing the waveforms of single-trial EEG $\hat{\mathbf{Z}}\in\Re^{T\times M}$; Note that only information related to the identified factors is contained in $\hat{\mathbf{Z}}$; Step 3: Averaging single-trial EEG for each condition, calculating the amplitudes or latencies for ERP for different conditions within a defined time-window at some specific electrodes, and computing the difference between conditions based on these obtained amplitudes or latencies. (b) Example of back projection for two factors that computing the sums of outer-products between temporal parts and spatial parts for both Factor 1 and Factor 2. "·" represents dot-product. " ∘ " represents outer-product between two vectors.

The collected EEG data were preprocessed offline. Firstly, the left and right mastoids were set to offline references, and the sampling rate was set to 500 Hz. Secondly, EEG data were filtered by an infinite impulse response (IIR) band-pass filter: the lower cut-off was 0.1 Hz, and the higher cut-off was 30 Hz, with 48 dB/oct slope. Next, the filtered EEG data were segmented from 200 ms before stimulus onset to 900 ms after stimulus onset. Those epochs whose magnitudes exceeded ±80 μV were discarded and the remaining epochs were baseline corrected by subtracting the mean amplitudes of baseline period (from −200 ms to 0 ms) from the amplitude of the whole epoch. Those epochs can be classified into six categories, namely, extreme disgusting (ED), moderate disgusting (MD), neutral disgusting (ND), extreme fearful (EF), moderate fearful (MF), and neutral fearful (NF).

We demonstrated that a linear filter (i.e., wavelet filter) plays an important role in denoising and reducing the number of factors related to noisy signals for BSS application and it can yield better overall results than BSS algorithm without filter (Cong et al., 2011c). Therefore, wavelet filter was used to improve SNR of single-trial EEG before any further analysis. The parameters were set for wavelet filter as below: the number of decomposition level was 9; the select mother wavelet was 'rbio6.8'; the detail coefficients at levels 4, 5, 6, 7, and 8 were used for signal reconstruction.

The preprocessed EEG data of 20 subjects were involved in N2 analysis and 17 subjects were for P2. Note that neural disgusting and neural fearful conditions were merged into one condition in the previous study (Lu et al., 2016). More details for data collection and experiment can be found in Lu et al. (2016).

2.4. Procedures for PCA application

To extract ERPs of interest from preprocessed ERP/EEG data, PCA was employed and the following steps are included: arrange preprocessed ERP/EEG datasets into a matrix, estimate the number of sources, select the rotation method, identify factors of interest, and project the identified factors onto electrode fields (i.e., in microvolt). Here, trial-averaged data are taken as an example to explain the procedures about how to use PCA to extract ERPs.

The first step is to arrange ERP datasets into a matrix. The recorded EEG datasets are born with temporal and spatial dimensions for multi-condition and multi-subject trial-averaged datasets, and therefore, two types of matrices can be formed (Dien, 2012; Dien et al., 2007, 2005). For the first type of matrix, time-samples are variables and the combinations of electrodes, conditions, and subjects are observations. The related PCA procedure is named as "tPCA". For the second type of matrix, electrodes are variables and the other variances (i.e., time samples and condition-subjects) are merged into be observations. Likewise, we call the performance of PCA on this type of matrix as "spatial PCA" (sPCA). In this study, the former type is formed based on the following reasons. For one thing, regarding the performance of tPCA and sPCA, tPCA can yield overall better results than sPCA (Dien, 1998, 2012). For another thing, the desired ERP is easily mixed with others in the spatial domain to some degree due to the volume conduction (Dien, 2012).

The second step is to estimate the number of factors related to stimulus onset. This step aims to use few factors to represent the whole data. Several approaches have been developed to realize this goal, such as Parallel Test (Horn, 1965; Dien, 2010a), gap measure (He et al., 2010; Cong et al., 2013), and cumulative explained variance (Zhang et al., 2020; Huster and Raud, 2018; Arbel et al., 2013). Parallel test and cumulative explained variance are frequently used to determine how many factors should be retained. Parallel Test suggests few factors to be extracted in the case that the factors are highly correlated, for example, ERP data are recorded from few sensors (Lim and Jahng, 2019; Beauducel, 2001). The previous studies have demonstrated that much stronger biases can be obtained from the few extracted factors than from many more factors (Wood et al., 1996; De Winter and Dodou, 2012).

Therefore, we here used cumulative explained variance to estimate the number of sources and it is actually to calculate the percentage L between the sums of first K lambda values (one lambda corresponds to one factor) over the sums of all lambda values:

$$L = \frac{\sum_{k=1}^{K} \lambda_k}{\sum_{t=1}^{T} \lambda_t} \times 100\%, \tag{4}$$

where K is the number of estimated sources associated with stimulus onset; T is the number of variables/time samples here ($K \ll T$); The lambda values are in descending order: $\lambda_1 \geq \cdots \geq \lambda_t = \cdots = \lambda_T = \sigma^2$ and they are the non-zero eigenvalues of the matrix $\mathbf{Z}^T\mathbf{Z}$ or $\mathbf{Z}\mathbf{Z}^T$. λ_k is the variance for kth factor. Once L is given with a specific value, for example, 95%, 99%, and so on, the corresponding number of sources can be then obtained.

The third step is to select the rotation method. The goal of rotation is to rearrange the factors into simple and interpretable structures so that one factor corresponds to one ERP (Dien, 2012). Some rotation methods, such as Promax, Varimax, can be utilized. Those rotations methods can roughly be oblique and orthogonal methods. The key difference between these two methods is that the former one allows the factors to be correlated with each other but the later requires the factors to be uncorrelated. Due to the volume conduction, ERP components overlap in temporal and spatial domain to some extent. Thus, the factors are required to be correlated substantially (Dien, 1998; Scharf et al., 2022). Besides, the results of actual and stimulated ERP datasets indicated that Promax rotation showed more accurate results than Varimax rotation (Dien et al., 2005; Dien, 1998; Dien et al., 2003). And, simulation ERP studies also revealed that Promax generated overall better results for tPCA, and Infomax yielded better separation for sPCA (Dien et al., 2007; Dien, 2010b). Therefore, Promax rotation was used in all cases in this study.

The fourth step is to identify factors of interest. Although some preprocessing steps are utilized to remove some artifacts or noise, like DC offset, slow drift of sensors, eye movement, line noise, and muscle contraction (Jung et al., 2000; Delorme et al., 2007; de Cheveigné and Nelken, 2019; de Cheveigné, 2020; Sai et al., 2017), the preprocessed data still contain spontaneous EEG brain activities, ERP components of non-interest, ERP components of interest, and noise activities. Therefore, we need to identify those factors associated with ERPs of interest for further analysis. Generally, the identification of desired factors is based on two aspects (Dien, 2012; Barry et al., 2020; Zhang et al., 2020): (1) the polarity and latency of temporal factor; (2) the polarity and topographical distribution of spatial factor.

The fifth step is to project the identified factors back onto electrode fields (i.e., rescale them to microvolt) for correcting the polarity and the variance indeterminacies between different factors. When performing PCA on an ERP dataset of multiple subjects, an ERP of interest may be decomposed into several factors because the differences are found in latencies or phases of this ERP across different subjects to some degree. For example, for N2, the latencies of some subjects' ERP are from 200 to 250 ms, and the others are located in 250–300 ms. Two factors whose latency is 136 ms and 168 ms respectively will be extracted by PCA (as shown in Fig. 3 for the PCA decomposition of all remaining trials). Therefore, all of the decomposed factors related to this ERP of interest should be back-projected onto electrode fields (Dien, 2012, 1998; Zhang et al., 2020).

3. Quantification of P2 and N2

In this study, ERPs of interest, P2 and N2, were reanalyzed because they reflected the emotional cognitive processing, which were reported in Lu et al. (2016). For the proposed method and other three alternative methods, both ERPs were separately quantified from all remaining trials and consecutively increased trials (i.e., 10, 11, ⋯, 41, and 42 trials). Notably, 10 is applied to ensure the number of observations is larger than the number of the variables when using PCA. 42 is used as the upper number because it is the maximal common number of trials across all conditions across all subjects. For the case with few trials, the number of trials is the same across conditions across subjects. For the case of all remaining trials, the number of trials may vary across conditions across subjects.

To better identify the factor(s) extracted by means of three PCA-based methods, a time-window was set for P2 (130–190 ms) and N2 (190–310 ms) separately. That is, factors are considered for further analysis only when peak latencies of corresponding temporal factors are within the defined time-window. According to the previous study (Lu et al., 2016), we measured the mean amplitudes for both P2 and N2 at six electrodes (FC3, FCz, FC4, C3, Cz, and C4 electrode sites) based on the inspection of topographical distribution of all remaining trials.

To distinguish the procedures of three PCA-based methods, we here define that single-trial EEG for pth participant $\mathbf{Z}^p \in \Re^{T\times(C\times E_p)}$ has T time points, C electrodes, and E_p trials ($E_p = E_{p,1} + \cdots + E_{p,j} + \cdots + E_{p,J}$ is the trial number of all conditions J). The trial-averaged ERP data of pth participant are then defined by using $\bar{\mathbf{Z}}^p \in \Re^{T\times(C\times J)} = \frac{1}{E_p}\sum_{e=1}^{E_p}\mathbf{Z}_e^p$. Therefore, the trial-averaged data of multi-participants is $\bar{\mathbf{Z}} \in \Re^{T\times(C\times J\times P)} = [\bar{\mathbf{Z}}^1, \ldots, \bar{\mathbf{Z}}^p, \ldots, \bar{\mathbf{Z}}^P]$. Single-trial EEG data of all participants is $\mathbf{\Phi} \in \Re^{T\times(C\times P\times(E_1+\cdots+E_p+\cdots+E_P))} = [\mathbf{Z}^1, \ldots, \mathbf{Z}^p, \ldots, \mathbf{Z}^P]$. The specific back-projection model of three PCA-based methods is in Appendix B.

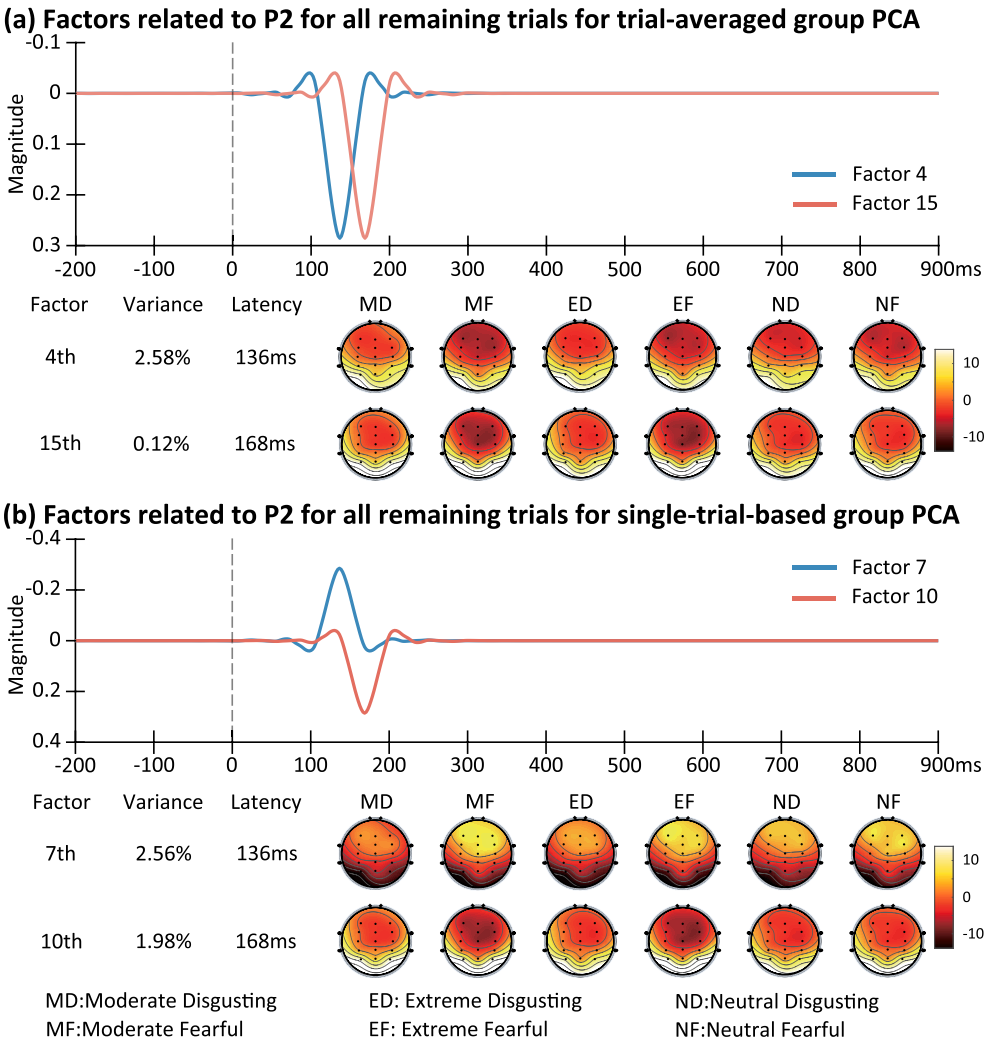


Fig. 3. Identified factors associated with P2 that were extracted from all remaining trials for trial-averaged group PCA (a) and single-trial-based group PCA (b). Topographical distributions were averaged across 17 participants for each condition. Higher percentage of explained variance for one factor implicates that this factor contributes more to P2 and latencies of P2 for more participants are near the peak point of that factor.

3.1. Conventional time-domain analysis

To quantify P2 and N2 in the conventional time-domain analysis, the following steps are included. Note that the mentioned single-trial EEG from this section were preprocessed based on the pipeline in Section 2.3.

For the analysis of all remaining trials, single-trial EEG data for each condition for each participant were first averaged to obtain ERP waveform. Similarly, for analysis of the increased trials (from 10, 11, ⋯, 41, and 42), single-trial EEG were also averaged separately. Mean amplitude was then quantified within the defined time-window from these averaged ERP waveforms for P2/N2 at the defined electrode sites. Consequently, for each case, we obtained 102 amplitude values for P2 (i.e., 17 participants and 6 conditions) and 120 amplitude values for N2 (i.e., for 20 participants and 6 conditions).

3.2. Trial-averaged group PCA

To isolate P2/N2 from ERP data for all participants by means of tPCA, six steps are included.

Step 1: Averaging single-trial EEG for both few trials and all remaining trials for each condition for each participant separately. For example, single-trial EEG data with 10 trials were averaged, resulting in a third-order tensor (15 channels × 550 samples × 6 conditions) for each participant.

Step 2: Arranging the averaged ERP data for all subjects and all conditions into a big-size matrix $\bar{\mathbf{Z}}$ along channels. The size of $\bar{\mathbf{Z}}$ for N2 is 550 samples × (15 channels × 6 conditions × 20 subjects), and it is 550 samples × (15 channels × 6 conditions × 17 subjects) for P2.

Step 3: Using PCA to decompose the matrix $\bar{\mathbf{Z}}$ into 550 factors that are identical to the number of samples, and 40 factors were remained and rotated using Promax rotation for all cases in Matlab environment (Version 2018b, the Mathworks, Inc., Natick, MA; functions: *pca.m* and *rotatefactors.m*). These 40 factors were considered to come from sources of stimulus-related signals, and we used them to represent the whole original data. The other 510 factors were discarded because they were regarded to be associated with noise or other signals. After inspection for all cases, these 40 factors accounted for more than 99% of overall variance for each case.

Step 4: Identifying the factors that were extracted from P2/N2 based on corresponding temporal factor $\mathbf{u}_k^{TGPCA}$ and spatial factor $\mathbf{y}_k^{TGPCA}$. For example, P2 is a positive-going ERP (with negative voltage) that is defined from 130 ms and 190 ms and is distributed around centrofrontal sites in this study. Factors 4 and 15 were finally selected to reconstruct the waveforms of P2 for all remaining trials because of the following four aspects (see Fig. 3a). For the first aspect, their latencies are 136 ms and 168 ms separately, and corresponding polarities are

negative for both factors. For the second one, we also observed the maximal amplitude is at centrofrontal sites, and polarities are negative for both factors. For the third one, the correlation of spatial parts between the two factors is about 0.8 which indicates that they are highly correlated. For the last one, the identification of the factor was also based on correlation of spatial factor between participants (i.e., the correlation for P2 should be more than 0.1 and N2 is more than 0.3). Note that we only described the correlations between identified factors for all remaining trials in the manuscript.

Step 5: Using the identified factors to reconstruct the waveform for ERP for all conditions for all participants based on back-projection theory (see Fig. 2b). For example, using Factor 4 and Factor 15 to reconstruct the P2 waveform for all remaining trials, that is, $\bar{\mathbf{Z}}_{P2}^{TGPCA} = \mathbf{u}_4^{TGPCA} \circ \mathbf{y}_4^{TGPCA} + \mathbf{u}_{15}^{TGPCA} \circ \mathbf{y}_{15}^{TGPCA}$ (see Eq. (B.1) in Appendix B). Note that the information of P2 is assumed to be contained in $\bar{\mathbf{Z}}_{P2}^{TGPCA}$ and other information is totally removed.

Step 6: Measuring the mean amplitude for P2/N2 within the defined time-window at specific electrode sites for all cases.

3.3. Single-trial-based group PCA

We also quantified P2 and N2 separately from all remaining trials and from few trials using single-trial-based group PCA. The same steps are as for the procedure of the trial-averaged group PCA (see Section 3.2) except for the following two aspects.

For the first aspect, in the first step, we organized single-trial EEG for all conditions for all subjects into a big-size matrix instead of averaging single-trial EEG. For example, a matrix $\mathbf{\Phi}$ with 550 samples × (15 channels × 6 conditions × 17 subjects × 10 trials) was formed to extract P2 from 10 trials for different conditions for all participants. For the second aspect, for the Step 5 in Section 3.2, we averaged single-trial EEG reconstructed by the identified factors for each condition for each participant. Note that 40 factors were also remained for all cases which explained more than 99% of total variance.

3.4. Within-subject PCA

The following steps were conducted to extract P2/N2 from the preprocessed single-trial EEG data for each subject with few trials and all remaining trials (see Fig. 2a).

Step 1: Single-trial EEG for all conditions for pth participant with a specific number of trials (e.g., 10 trials) were arranged into a matrix $\mathbf{Z}^p$ with size of $T \times (C \times (E_{p,1} + \cdots + E_{p,j} + \cdots + E_{p,J}))$. T is the number of samples, C is the number of channels, and $E_{p,j}$ is the number of trials for jth condition ($j = 1, \ldots, J-1$, and J). Note that, for the analysis of all remaining trials, $E_{p,1}$, $\cdots$, $E_{p,J-1}$, and $E_{p,J}$ may vary across conditions across participants because of artifact rejection during preprocessing procedures. Whereas, for the analysis of adding trial numbers, the number of trials is the same for all conditions and all participants (i.e., trial number is 10, 11, $\cdots$, 41, and 42, separately).

Step 2: The matrix $(\mathbf{Z}^p)^T$ was decomposed into T (i.e., 550) factors, and $K^{(p)}$ factors were then remained and rotated by means of Promax rotation ($K^p \leq T$) (see Appendix B). Here, aiming to extract ERPs of interest successfully, K^p was set to 40 for both few trials (consecutively increased from 10 to 42 trials) and all remaining trials. These 40 factors accounted for more than 99% of total variance for each case.

Step 3: The factors associated with P2/N2 were chosen for the next procedure according to following aspects. The first aspect was to judge if the latency of kth temporal factor $\mathbf{u}_k^{WPCA}(p)$ fell in the predefined time window. In this study, 130–190 ms was for P2 and 190–310 is for N2. The second one was to judge if the topographical distribution of kth spatial factor $\mathbf{y}_k^{WPCA}(p)$ was in accordance with spatial property of P2/N2. More details about how to identify factors of interest were described in fifth step of Section 3.2. Thirdly, the correlations between factor scores for the selected factors were computed and examined. Lastly, correlation of spatial factor between participants (i.e., the correlation for P2 should be more than 0.1 and N2 is more than 0.3)was computed and examined.

Step 4: The identified factors were projected onto electrode fields for correcting their variance and polarity indeterminacies, and these factors were then reconstruct a new matrix $\mathbf{Z}^p = \mathbf{u}_{k_1}^{WPCA}(p) \circ \mathbf{y}_{k_1}^{WPCA}(p) + \cdots + \mathbf{u}_{k_n}^{WPCA}(p) \circ \mathbf{y}_{k_n}^{WPCA}(p)$ (see Eq. (B.3)). Besides, although the identified factors and the corresponding order/number may vary across participants, we can use back-projection theory to address this issue and reconstruct the waveforms of single-trial EEG at microvolt level. The reconstructed single-trial EEG data were then averaged for each experimental condition separately, resulting in a matrix with 550 samples × (15 channels × 6 conditions) for each participant for each case.

Step 5: Mean amplitudes of P2/N2 of different conditions for that participant were finally measured within a defined time-window at some electrodes.

Step 6: Repeating Steps 1–5 until both P2 and N2 were extracted from the single-trial EEG of each subject and quantified from the reconstructed ERP data.

Here is main difference among four used methods. Conventional time-domain analysis assumes that only the information associated with the analyzed ERP (i.e., one brain activity) is contained within the defined time-window although the analyzed ERP may overlap with other ERPs (Beauducel and Debener, 2003; Donchin and Heffley, 1978). PCA is able to disentangle those spatially and temporally overlapped ERP components efficiently (Dien, 2012; Kayser and Tenke, 2003, 2006a,b). For the two group PCA, they assume that the time-course (i.e., factor loading) is the same across all participants across all conditions, whereas within-subject PCA allows the time course to vary across participants.

3.5. Internal consistency analysis and statistical analysis

For the analysis of all the remaining trials, to evaluate the performance for four used methods, we respectively computed the corresponding Cronbach's alpha. During computation of Cronbach's alpha, participants were then serving as 'items' and samples were observed to assess the extent of internal consistency of waveforms for P2/N2 across participants. Moreover, we also computed spatial Pearson correlation coefficients (CCs) of topographies among participants to assess how P2 and N2 are spatially similar across participants. The topography of one subject is characterized by 90 values (15 electrode sites and 6 conditions).

For the few trials, besides the Cronbach's alpha and spatial CCs, we explored the relationships of waves between few trials and all remaining trials by computing CCs between the averages of small trial numbers and the grand averaged N2/P3 (i.e., all remaining trials were included). For example, computing the CC for waveforms between the average with 10 trials and grand averages with all remaining trials for conventional time-domain analysis. The aim is to examine the consistency between averages of small trials and the grand averages.

We used two-ways repeated measurement analysis of variance (rm-ANOVA) with valence (extreme, moderate, and neutral) × negative-category (disgusting and fearful) as within-subject factors to compute the statistical results of the measured mean amplitudes for all trials and increased trials separately. The uses of ANOVA obey the following rationales: (1) individual measurements are mutually independent; (2) the data follow an additive statistical model including fixed effects and random errors; (3) the random errors are normal distribution; and (4) the random errors are as homogeneous as possible. Greenhouse–Geisser method was used for correcting the number of degrees of freedom.

Table 1
Orders, latencies (i.e., peak latencies of the selected factor loadings), percentages (%) of explained variance of the identified factors extracted from all remaining trials for P2 for each participant for within-subject PCA.

Subject #	# 1	# 2	# 3	# 4	# 5	# 6	# 7	# 8	# 9
Factor Order	9,11	7,15	11,12	5,7	32	4	16	4,10	9
Latency (ms)	168,136	168,136	136,168	136,168	168	136	168	136,168	168
Variance (%)	1.32,1.06	2.64, 1.33	1.35,1.19	4.20,2.89	0.18	4.37	1.41	4.91,1.87	1.92
Subject #	# 10	# 11	# 12	# 13	# 14	# 15	# 16	# 17	–
Factor Order	11,14	6	13,25	10,30	23	16	10,13	4,20	–
Latency (ms)	136,168	136	136,168	168,136	136	136	136,168	168,136	–
Variance (%)	1.21,1.08	3.47	1.17,0.41	1.85,0.41	0.41	1.44	1.82,1.45	3.98,0.90	–

Table 2
Orders, latencies (i.e., peak latency of the selected factor loadings), percentages (%) of explained variance of identified factors extracted from all remaining trials for N2 for each participant for within-subject PCA.

Subject #	# 1	# 2	# 3	# 4	# 5
Factor Order	6,8,15,32	8,11,12,13	16	1,10,14,30	3,9,11,32
Latency (ms)	232,264,296,200	232,200,264,296	296	264,232,296,200	200, 296,264,232
Variance (%)	2.13,1.62,0.76,0.25	2.39,1.59,1.58,1.46	1.44	55.08,1.44,0.93,0.26	7.24,2.42,1.90,0.32
Subject #	# 6	# 7	# 8	# 9	# 10
Factor Order	3,7,10,33	10,12,13,14	2,4,9,30	10,11,12	3,32
Latency (ms)	200,264,296,232	200,232,264,296	264,200,296,232	296,232,264	296,264
Variance (%)	8.47,2.47,1.48,0.17	1.80,1.60,1.47,1.29	8.02,4.46,2.07,0.33	1.66,1.59,1.53	7.63,0.28
Subject #	# 11	# 12	# 13	#1 4	# 15
Factor Order	11,30	3,12,32	8,9,13	6,11,12,29	5,16,17,33
Latency (ms)	264,232	232,264,200	232,264,296	232,264,296,200	264,296,200,232
Variance (%)	1.53,0.27	5.57,1.17,0.17	2.86,2.46,1.60	2.94,1.25,1.22,0.21	3.46,1.01,0.96,0.35
Subject #	# 16	# 17	# 18	# 19	# 20
Factor Order	2,6,10	15	2,4,7	2,8,11,32	1,8,31,25
Latency (ms)	264,232,296	232	296,200,232	232,296,264,200	264,200,232,296
Variance (%)	10.76,2.53,1.33	1.85	10.65,6.28,3.37	9.21,1.99,1.65,0.29	47.54,2.00,0.73,0.56

4. Results

4.1. Results for all remaining trials

4.1.1. Factors extracted by three PCA-based methods for P2/N2

In this section, we descried the factors used to recover waveforms of P2/N2 for three PCA-based methods for all remaining trials.

P2: For trial-averaged group PCA, as shown in Fig. 3a, factors 4 (latency is 136 ms) and 15 (latency is 168 ms) were selected, and they explained 2.58% and 0.12% of total variance separately. The correlation between their spatial parts was 0.8. As for single-trial-based group PCA (Fig. 3b), we observed that the properties of factors 7 and 10 were consistent with P2 and they explained 2.56% and 1.98% of total variance separately. Latency for both factors is 136 ms and 168 ms separately. The correlation between their spatial parts was −0.80. For within-subject PCA, we identified about 1.59 ± 0.51 factors (explained 2.95 ± 1.97% of total variance) that were related to P2 from single-trial EEG for each participant for within-subject PCA. The identified factors and corresponding information for each participant were described in Table 1.

N2: Factors 2 (232 ms), 5 (200 ms), 9 (296 ms), 26 (264 ms) were identified for trial-averaged group PCA (Fig. 4a) and they explained 8.53%, 1.87%, 0.41%, and 0.04% of total variance separately (Fig. 4b). Their spatial correlations were highly correlated (Fig. 4c) and the negative correlation meant that the polarities between two factors were opposite. Similarly, for single-trial-based group PCA (Figs. 4d–f), four factors were also selected, namely, Factor 3 (232 ms), Factor 10 (200), Factor 12 (264 ms), and Factor 16 (296 ms) which accounted for 7.54%, 2%,1.64%, and 1.25% of total variance, respectively. They were also highly related based on the correlations between the corresponding spatial parts (Fig. 4f). For within-subject PCA, 3.18 ± 1.07 factors that explained 12.85 (± 15.04)% of total variance for each subject were used to reconstruct N2 and we listed the identified factors and the related information in Table 2.

4.1.2. Reconstructed grand averaged waveform and topography

Figs. 5 and 6 display reconstructed grand averaged waveforms and topographies for P2/N2 for those four used techniques when all the remaining trials were involved based on the identified factors as described in Section 4.1.1.

Comparing with topographies and waveforms of P2 (Fig. 5) and N2 (Fig. 6) by conventional time-domain analysis, they were reconstructed well for P2/N2 using three PCA-based methods. Note that the waveforms and topographies of P2/N2 for trial-averaged group PCA and single-trial-based group PCA were no different (Figs. 5 and 6b, c). But it does not mean that they are the same (although PCA is a linear decomposition algorithm) because the proportion of explained variance of selected factors that used to reconstruct those waveforms of P2/N2 for trial-averaged group PCA were different from those for single-trial-based group PCA. The proposed method reconstructed the waveforms and topographies of P2/N2 successfully as the other two group PCA methods did (Fig. 5d and Fig. 6d).

We observed that the P2 was a positive-going ERP component and had negative voltage in the original waveform, whereas, as shown in Figs. 5b–d, the reconstructed waveforms of P2 of three PCA-based methods were negative-going because PCA just performed on the voltage of P2 on the original data rather than the corresponding polarity. Likewise, the reconstructed N2 waveforms were negative-going that consistent with original waveforms for three PCA-based methods as shown in Fig. 6.

4.1.3. Internal consistency and statistical results for P2/N2

P2: Both Cronbach's alpha (excellent level > 0.9) and spatial sim-

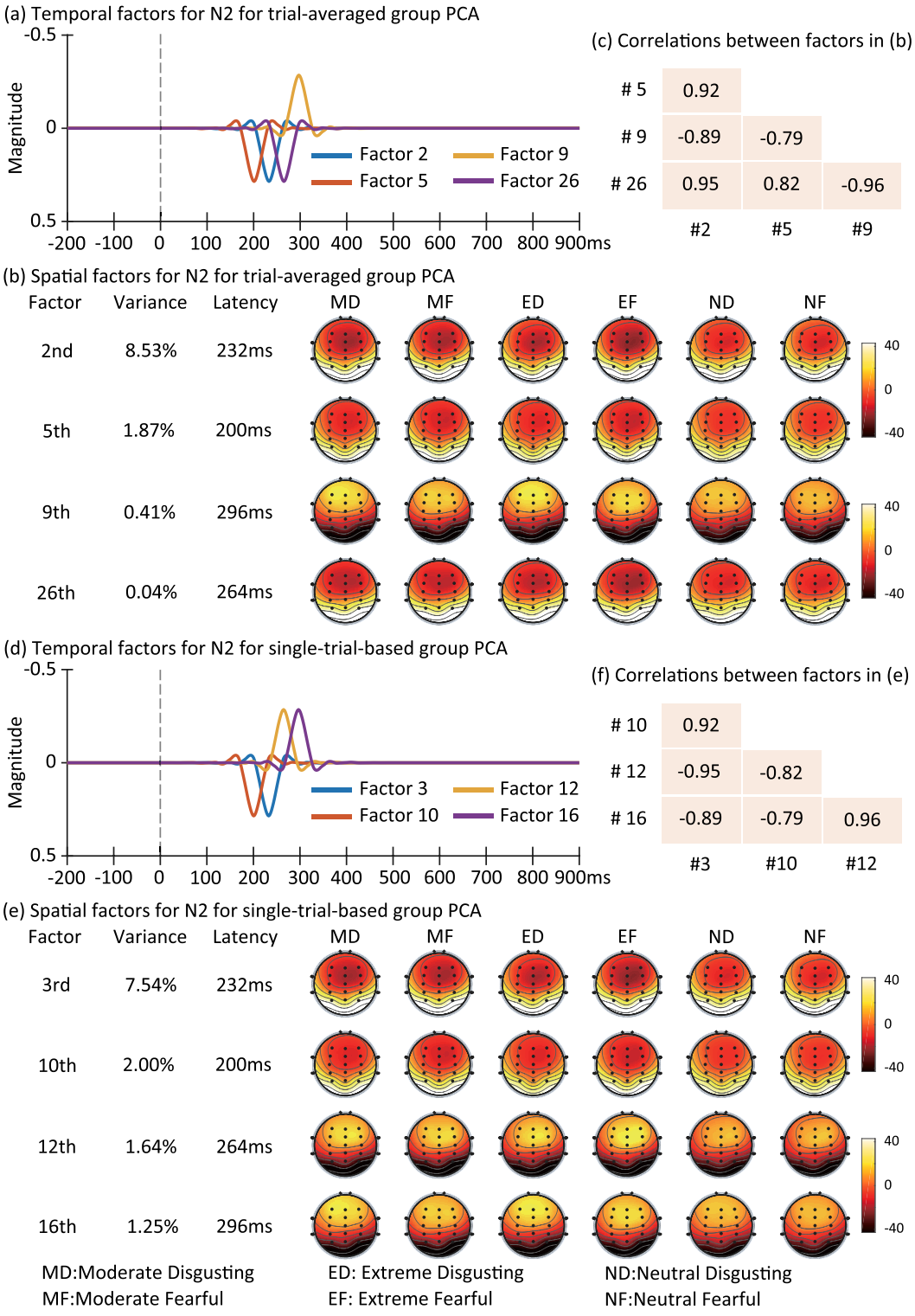


Fig. 4. Identified factors related to N2 that were extracted from all remaining trials and their correlations for trial-averaged group PCA (a–c) and single-trial-based group PCA (d–f). Topographical distributions were obtain by averaging across 20 participants for each condition (b and e) and correlations was based on those topographies of all conditions and participants for two factors (c and f). Higher proportion of explained variance for one factor implicates that this factor contributes more to N2 and latencies of N2 for more participants are nearly distributed to the peak point of that factor. Negative correlation as shown in c and f between two factors reflects their polarities are opposite.

ilarity for two group PCA methods are lower for the conventional time domain and within-subject PCA. The spatial topographies for the within-subject PCA were much more consistent across participants than for the conventional time domain (see Fig. 7a and b). For the statistical analysis, as shown in Table 3, there was no significant difference for all used techniques between disgusting and fearful stimuli. We found significant main effect for negative-category and the interaction effect between valence and negative-category.

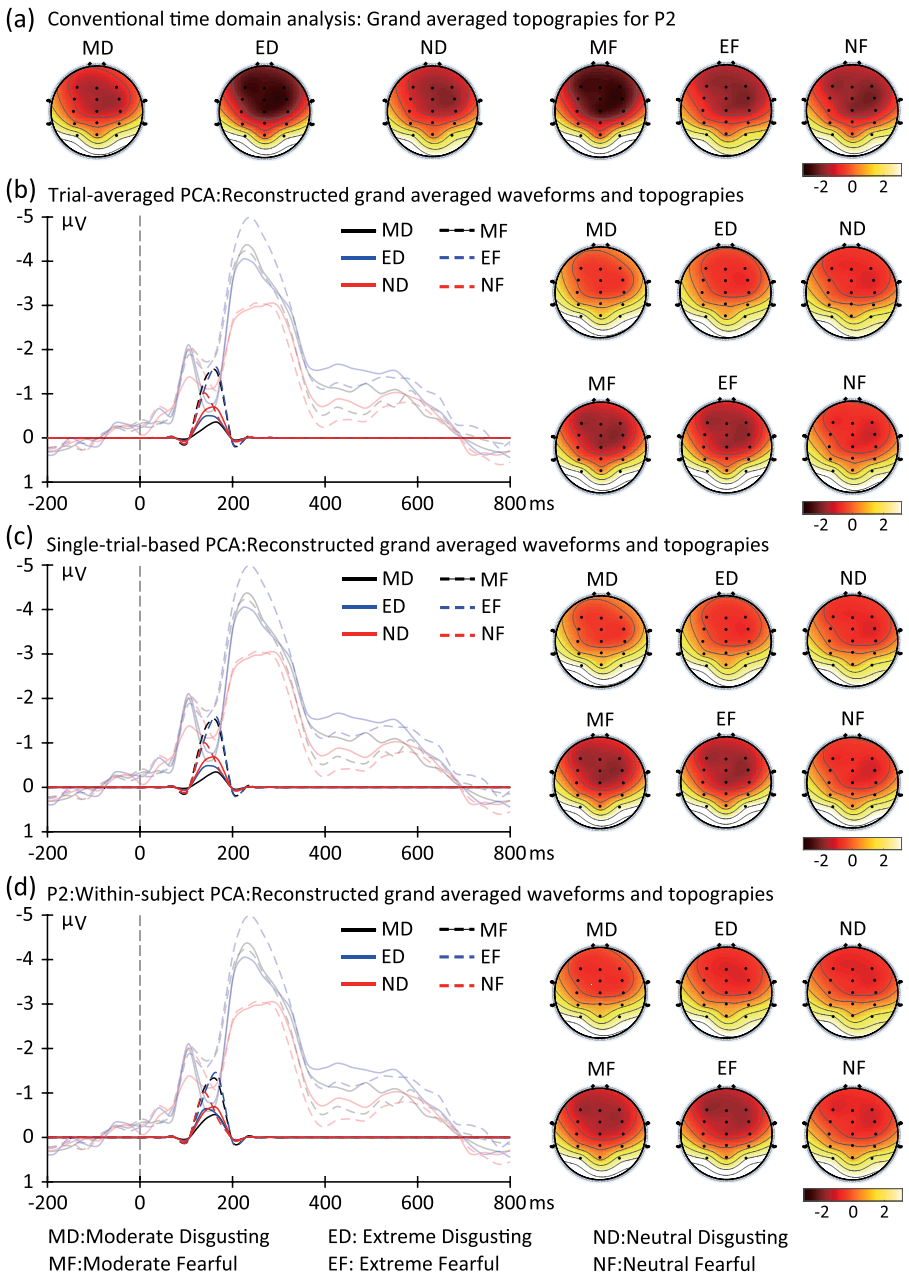


Fig. 5. Reconstructed grand averaged waveforms (at FC3, FC4, FCz, C3, C4, and Cz electrode sites) and topographies of P2 under different conditions for all remaining trials for time-domain analysis (a), trial-averaged group PCA (b), single-trial-based group PCA (c), and within-subject PCA (d). Note that the transparent lines in panels b–d are grand averaged waveforms for time-domain analysis. Topographies were measured from 130 to 190 ms for P2 across 17 participants. Time window from −200 ms to 800 ms was used here to show P2 clearly.

N2: There is no difference in Cronbach's alpha among the four used methods and the three PCA-based methods yielded slightly high Cronbach's alpha (see Fig. 7c). Spatial similarity is improved for the proposed method compared with the other three methods (see Fig. 7d). For the statistical analysis, we observed a significant main effect for valence for all used techniques but not for negative-category. The interaction effect between valence and negative-category was also significant (see Table 4).

4.2. Effect of the number of trials on quantification of P2/N2

Here we displayed the internal consistency of P2/N2 based on temporal CCs and Cronbach's alpha, spatial similarity, and statistical analysis to show the influence of the number of trials on quantification of P2/N2. The selected number of factors related to P2/N2 and the explained variance of these factors for three PCA-based methods are shown in Appendix C (see Figs. C.1 and C.2).

4.2.1. Internal consistency of P2/N2

We used CCs to examine the relationships between averages of small trials and grand average of all remaining trials for P2/N2 (see Fig. 8 a and b).

Regarding P2 (Fig. 8a), the excellent CCs (>0.9) were observed after about 25 trials for within-subject PCA, whereas few trials (about 10 trials) were needed to obtain excellent CCs for the other three approaches. The CCs for two group PCA methods was slightly higher than for conventional time-domain analysis because both methods extracted common features of P2 across participants, resulting in the similar reconstructed P2. Besides the common feature of P2, within-subject PCA also captures variability of P2 for one participant. As

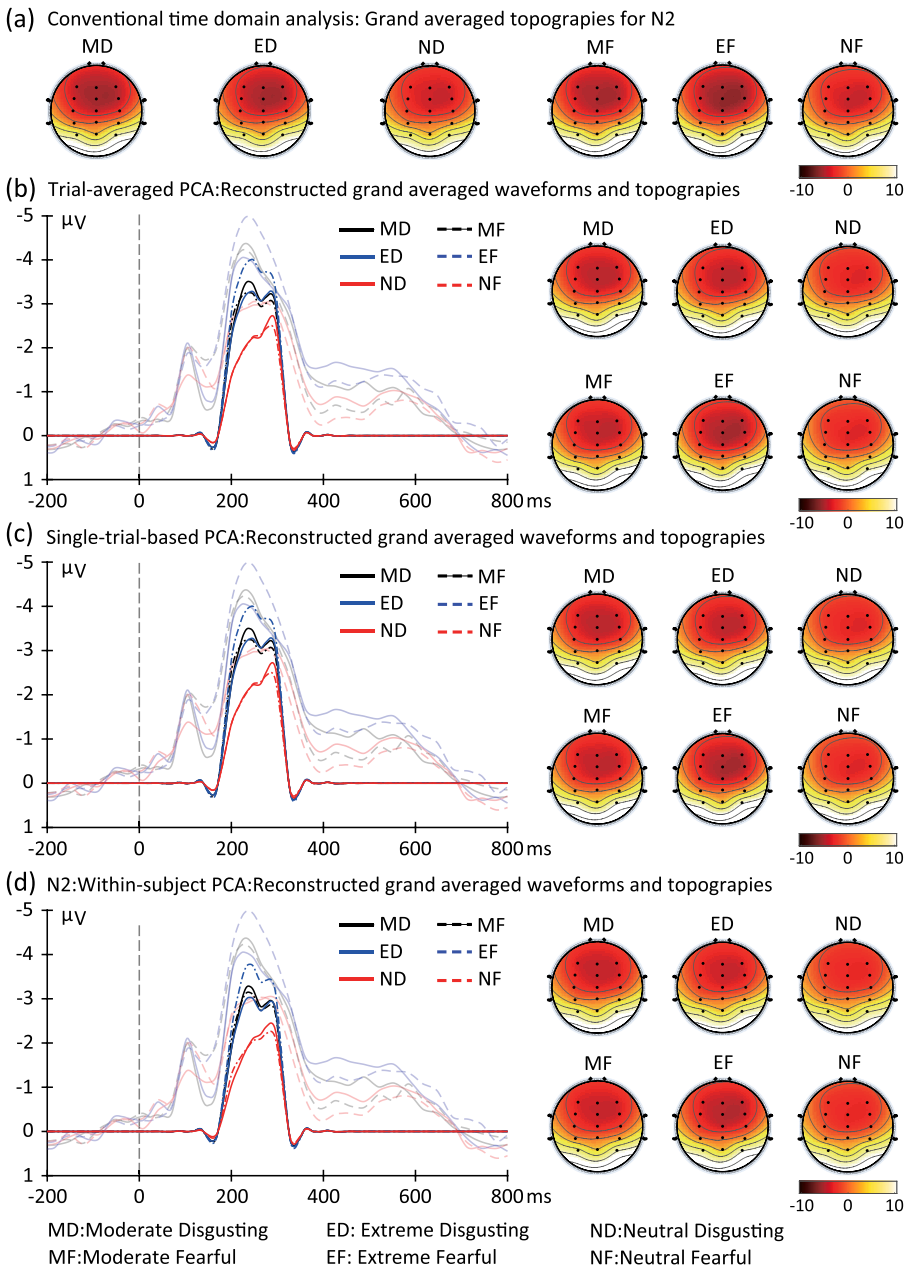


Fig. 6. Reconstructed grand averaged waveforms (at FC3, FC4, FCz, C3, C4, and Cz electrode sites) and topographies of N2 for all remaining trials for conventional time-domain analysis (a), trial-averaged group PCA (b), single-trial-based group PCA (c), and within-subject PCA (d). Note that the transparent lines in panels b–d are grand averaged waveforms for time-domain analysis. Topographies were measured from 190 to 310 ms for N2 and were averages across 20 participants for all conditions. Time window from −200 ms to 800 ms was used here to show N2 clearly.

for N2 (Fig. 8b), the three PCA-based approaches had higher CCs than for conventional time-domain analysis as shown in the previous study (Steele et al., 2016). There was no difference for CCs of P2/N2 between the two group PCA methods, which is consistent with the claim that PCA is a linear decomposition algorithm (Wold et al., 1987; Anowar et al., 2021)

We also computed the Cronbach's alpha of P2 and N2 for those four approaches separately when trials were added to the averages (see Fig. 8 c and d). For both P2 and N2, all the methods showed a high intralsubject consistency. That is, Cronbach's alpha is greater than 0.95. Specifically speaking, for P2, Cronbach's alpha value was lowest for both group PCA methods, slightly higher for the other two methods (Fig. 8c) which is contrast with the findings in Steele et al. (2016). As for N2, three PCA-based methods helped to improve intralsubject consistency slightly compared with conventional time-domain analysis which is consistent with Steele et al. (2016). That is, Cornbach's alpha was lowest for conventional time-domain analysis. Note that those differences between the used four methods were negligible (Fig. 8d).

The following issues need to notice. Firstly, CCs (or Cronbach's alpha) for P2 was different from that for N2 because different numbers of participants were involved and N2 was a dominant part for the whole time period. Secondly, the difference in the number of participants between P2 and N2 was mainly reflected in CCs rather than in Cronbach's alpha. Thirdly, as we described in Section 4.1.1, the lines of CCs/Cronbach's alpha for both group PCA methods were coincidence, but it did not mean that they were the same.

4.2.2. *Change of correlation coefficients of topographies among participants*

To examine the performance of those four used techniques, we assessed the intersubject consistency of P2/N2 by measuring the CCs

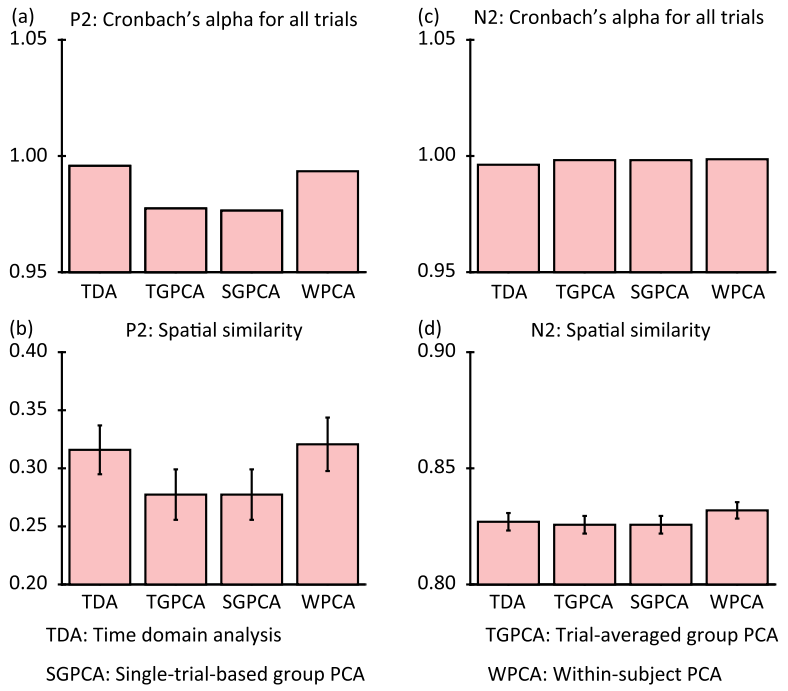


Fig. 7. Cronbach's alpha spatial similarity of both P2 and N2 for all remaining trials for the four used methods.

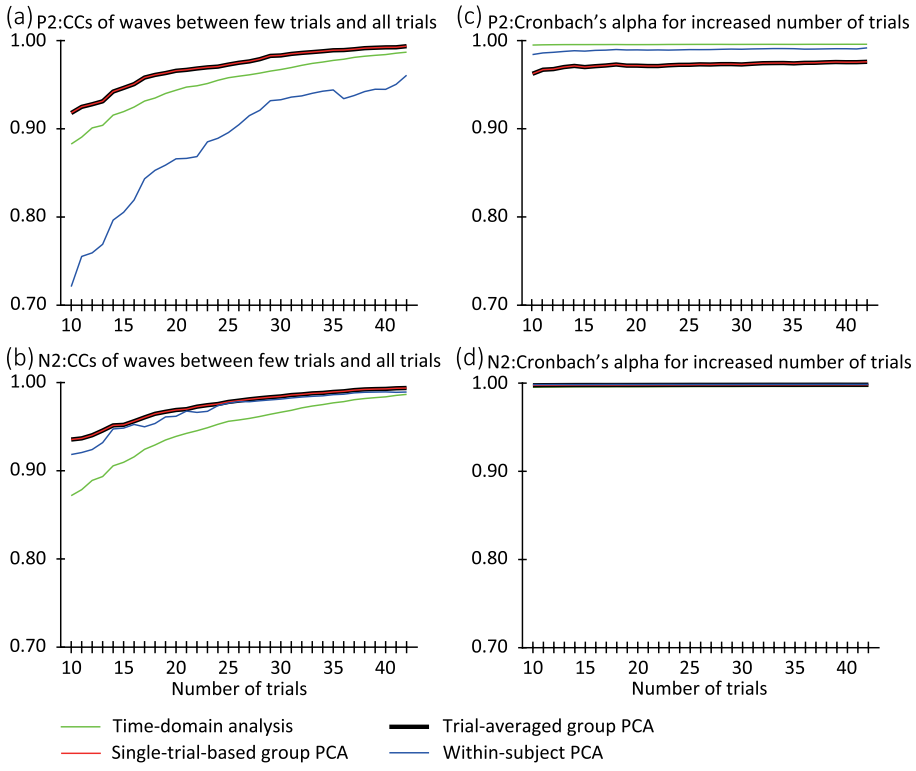


Fig. 8. Internal consistency analysis for P2/N2. (a) and (b): Correlation coefficients (CCs) between P2/N2 averages of consecutive increased trials (i.e., 10, 11, …, 41, and 42) and all remaining trials. (c) and (b) are Cronbach's alpha for adding trials. Electrodes are used for both P2 and N2 at FC3, FCz, FC4, C3, Cz, and C4 electrodes.

of topographies ('spatial similarity') between participants for increased trials and all remaining trials (see Fig. 9).

For P2 (see Fig. 9 upper panel), spatial similarity for both conventional time-domain analysis and within-subject PCA was higher than for the other two group PCA methods in most of cases. Spatial similarity for within-subject PCA went up with the number of trials increases and it changed slightly for conventional time-domain analysis across trials. Additionally, the difference of spatial similarity between the two methods decreased gradually with the number of trials increases. Topographies for the two group PCA methods were the same but were less similar between participants.

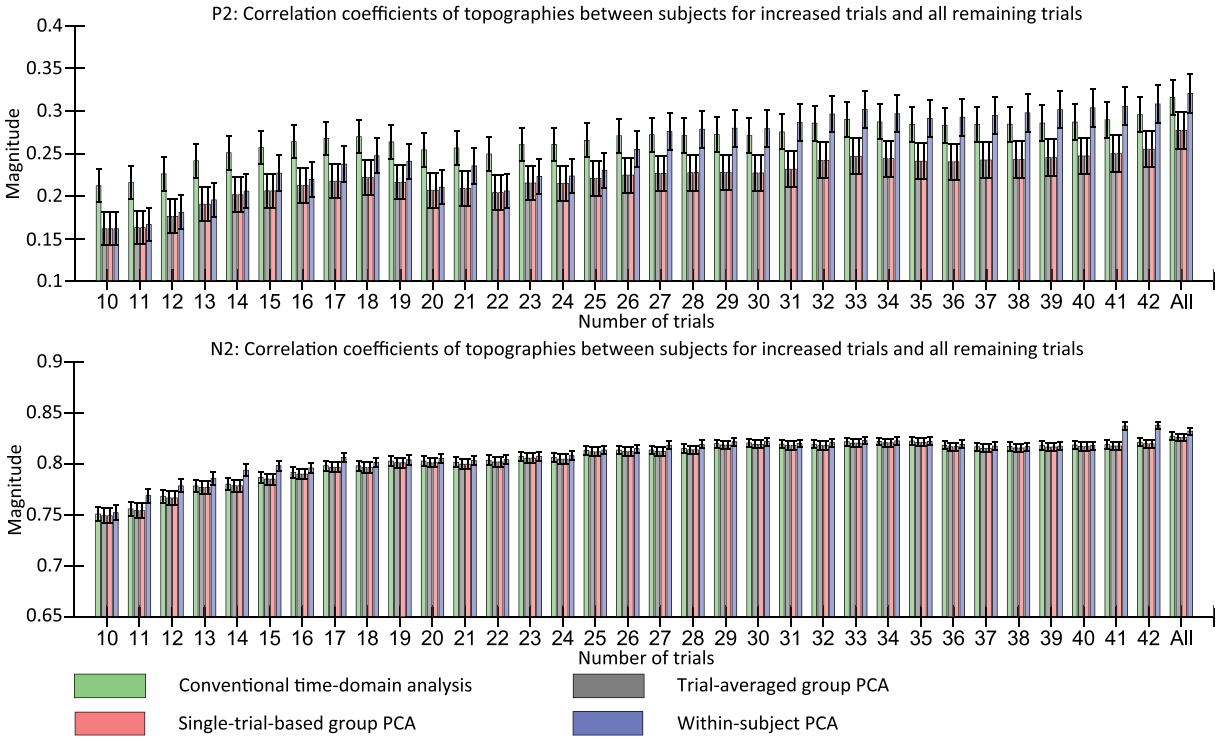


Fig. 9. Correlation coefficients (CCs) of topographies between any two different subjects for adding trial numbers and all remaining trials for the four used techniques for P2 (upper panel) and N2 (lower panel). Note that the topographies are for six conditions within one participant. Topographies were measured from 130 to 190 ms for P2 and 190–310 ms for N2 for all participants for each case. Error bars were the standard error of the mean of the possible pairs. The number of pairs is J*P*(P-1)/2 (P is number of participants and J is number of conditions).

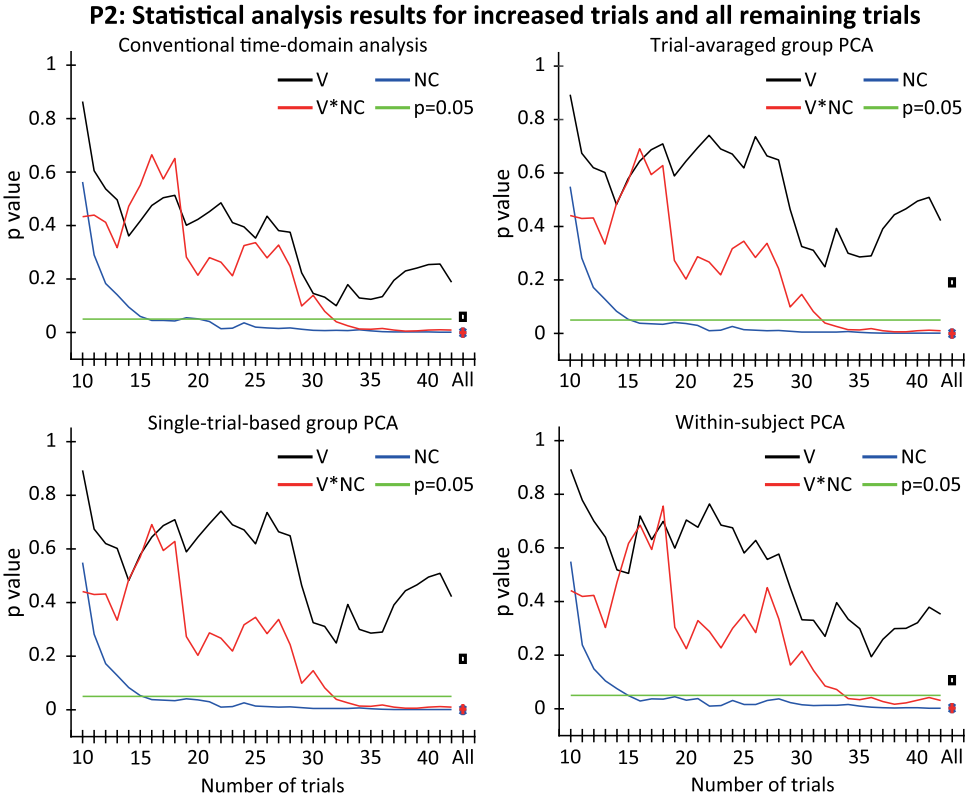


Fig. 10. Statistical analysis results for P2 averages of increased trials and grand average of all remaining trials using four different methods. Valence (V) × Negative-category (NC) are the within-subject factors. V*NC is the interactive effect between valence and negative-category factors.

Table 3
Statistical results of P2 (130–190 ms) for four used methods for all remaining trials. Valence (V) and Negative-category (NC) are the within-subject factors. TDA: Time-domain analysis. TGPCA: trial-averaged group PCA. SGPCA: Single-trial-based group PCA. WPCA: Within-subject PCA.

Methods	V			NC			$V * NC$		
	F	η_p^2	p	F	η_p^2	p	F	η_p^2	p
TDA	3.079	0.161	0.060	19.392	0.548	**< 0.001**	10.355	0.393	**< 0.001**
TGPCA	1.752	0.099	0.190	21.303	0.571	**< 0.001**	10.484	0.396	**0.001**
SGPCA	1.752	0.099	0.190	21.303	0.571	**< 0.001**	10.484	0.396	**0.001**
WPCA	2.387	0.130	0.108	17.868	0.528	**0.001**	7.245	0.312	**0.004**

Table 4
Statistical results of N2 (190–310 ms) for the four used methods when all trials are averaged. Valence (V) and Negative-category (NC) are the within-subject factors. TDA: Time-domain analysis. TGPCA: trial-averaged group PCA. SGPCA: Single-trial-based group PCA. WPCA: Within-subject PCA.

Methods	V			NC			$V * NC$		
	F	η_p^2	p	F	η_p^2	p	F	η_p^2	p
TDA	27.431	0.591	**< 0.001**	0.415	0.021	0.527	5.081	0.211	**0.012**
TGPCA	27.455	0.591	**< 0.001**	0.452	0.023	0.510	5.198	0.215	**0.011**
SGPCA	27.455	0.591	**< 0.001**	0.452	0.023	0.510	5.198	0.215	**0.011**
WPCA	24.397	0.562	**< 0.001**	0.746	0.038	0.399	5.553	0.226	**0.010**

Regarding N2 (see Fig. 9, lower panel), spatial similarities for within-subject PCA were slightly higher than for the other three approaches. There were no much more differences of spatial CCs for all used methods across trials (Fig. 9, lower panel).

Furthermore, the spatial similarities in all cases were approximately three times as great for P2 (about 0.25) as for N2 (roughly 0.75) for all approaches. Compared with N2 (about 15 trials), we need more trials for P2 (about 25 trials) to yield similar results as for all remaining trials.

4.2.3. Number of trials affects the statistical results of P2/N2

To evaluate the influences of the number of trials on the statistical results, we applied two-ways rm-ANOVA to calculate the statistical results based on the mean amplitudes of P2/N2 that were obtained from few trials and all remaining trials, separately (see Figs. 10 and 11).

Fig. 10 displays the statistical results of increased trials and all remaining trials for P2 for the four used approaches. We found that the p values of main effects of Valence/Negative-category, and their interaction showed decreased trends with number of trials increases. These trends suggest that adding trials had very a slight effect on the statistical results after a certain number of trials, for example, 30 trials. For the proposed method, 34 trials or more trials could yield the same statistical results as all remaining trials do, but only 32 trials were needed for the other three methods.

Similarly, regarding statistical results for N2, as shown in Fig. 11, we also observed the downward trends for p values of both main effects of Valence/Negative-category and their interaction effect with the increase of number of trials. Few trials (about 19) for within-subject PCA could produce the stable and equivalent results as all trials do, but approximately 33 trials were for the other three techniques.

5. Discussion

Factor analysis methods, for example, temporal principal component analysis (tPCA), have been widely used to extract event-related potentials (ERPs) of interest from ERP data of all participants (called trial-averaged group PCA) or single-trial EEG of all participants (named single-trial-based group PCA). Both methods assume that the underlying factor loading is identical across experimental conditions and across participants which is also famous as measurement invariance (Meredith, 1993). However, this assumption may fail to work if the latency and phase for one ERP component vary across participants considerably. For example, for N2 in this study, latencies for some participants are around 200–250 ms and the others maybe 250–300 ms. Consequently, the factor loadings may be different across participants. As were reported in previous psychometric studies (Borkenau and Ostendorf, 1998; Molenaar and Campbell, 2009; Nesselroade and Molenaar, 2016; Marsh et al., 2018), in the first part, we extracted ERP components from single-trial EEG of an individual (refer Fig. 2a) by comparing with the factors of P2/N2 obtained by the other two group PCA methods in a modified oddball emotional experiment (for all remaining trials). In the second part, we investigated how the number of trials affects quantification of P2 and N2. That is, how many minimum number of trials can yield the stable results as all remaining trials do which is explored by computing the Pearson correlation coefficients (CCs) of waves between few trials and all remaining trials, Cronbach's alpha, spatial similarities, and statistical analysis.

5.1. Differential results between PCA-based methods/ERP components for all remaining trials

As illustrated by Eqs. (1) and (2), ERP data are assumed to be the mixtures of brain activities and noise signals (i.e., PCA is based on the linear transformation (Anowar et al., 2021; Wold et al., 1987)), it is therefore no surprising that the yields for trial-averaged group PCA (TGPCA) and single-trial-based group PCA (SGPCA) are almost the same if all else being equal. Mathematically (see Appendix B), the PCA decomposition from the trial-averaged data $\bar{\mathbf{Z}}$ of all participants is approximately equal to the mean of reconstructed single-trial EEG $\mathbf{\Phi}$ (the number of trials is E), i.e., $PCA[\bar{\mathbf{Z}}] \approx \frac{1}{E}\sum_{e=1}^{E}(PCA[\mathbf{\Phi}])_e$ ($\mathbf{\Phi} = [\mathbf{Z}_1, \ldots, \mathbf{Z}_e, \ldots, \mathbf{Z}_E]$ and $\mathbf{Z}_e$ is eth single-trial EEG of all conditions and all participants). PCA[·] represents the whole PCA decomposition procedure, and $\bar{\mathbf{Z}} = \frac{1}{E}\sum_{e=1}^{E}\mathbf{Z}_e$. We will get highly similar reconstructed waveforms and topographies if the numbers of the estimated factors are the same (see Appendix B, K_1 is TGPCA for and K_2 is for SGPCA) and all factors associated with an ERP are identified for back-projection for both group PCA methods (see factors are for P2 and N2 in Figs. 3 and 4). Simply, if we assume only one factor is related to ERP component of interest, the back-projection $\mathbf{u}_{k_1}^{TGPCA} \circ \mathbf{y}_{k_1}^{TGPCA}$ of that ERP for TGPCA is approximately equal to the average of the back-projection of single-trial EEG $\frac{1}{E}\sum_{e=1}^{E}\mathbf{u}_{k_2}^{SGPCA} \circ \mathbf{y}_{k_2}^{SGPCA}(e)$ for SGPCA. k_1 may be different with k_2 (see Figs. 3 and 4).

The above mentioned discussion is demonstrated by the P2 and N2 results for the two group PCA methods. For example, the reconstructed P2 wave from the PCA decomposition of the trial-averaged ERP data ($\bar{\mathbf{Z}}_{P2} = \mathbf{u}_4^{TGPCA} \circ \mathbf{y}_4^{TGPCA} + \mathbf{u}_{15}^{TGPCA} \circ \mathbf{y}_{15}^{TGPCA}$) is approximately equal to the average of the reconstructed single-trial P2 wave of all participants $\frac{1}{E}\sum_{e=1}^{E}(\mathbf{\Phi})_e^{P2} = \frac{1}{E}\sum_{e=1}^{E}(\mathbf{u}_7^{SGPCA} \circ \mathbf{y}_7^{SGPCA}(e) + \mathbf{u}_{10}^{SGPCA} \circ \mathbf{y}_{10}^{SGPCA}(e))$. Note that the latencies for P2 factors extracted by the two group PCA methods are the same (see Fig. 3a and b). Similarly, for N2, the almost equivalent reconstructed waves were also obtained for both

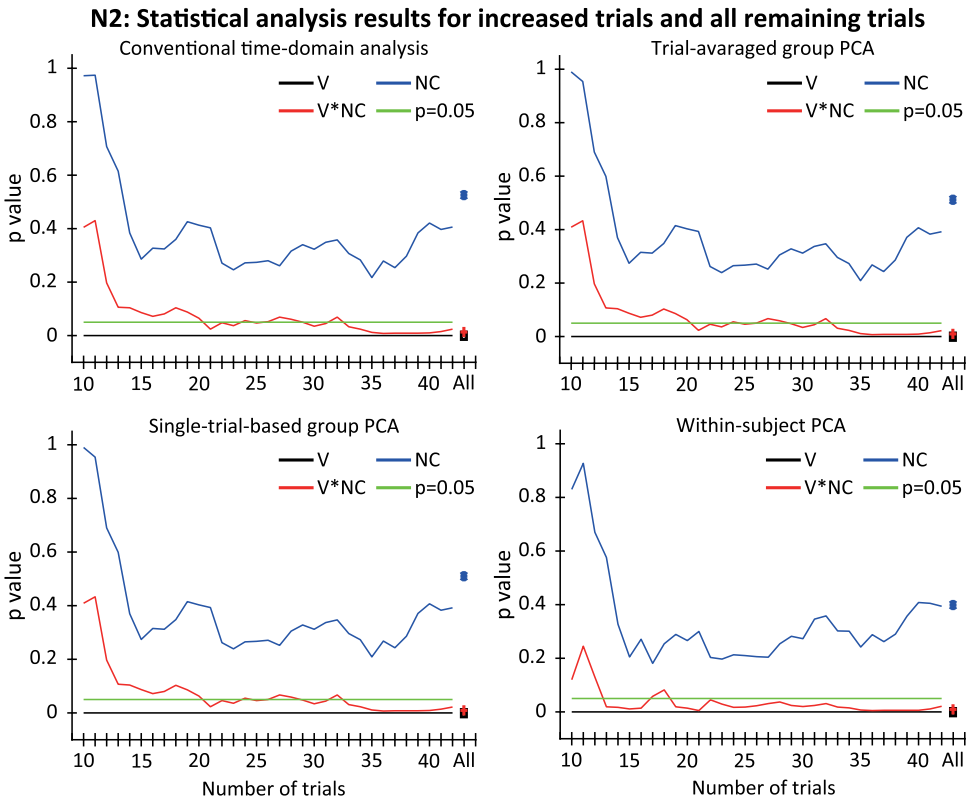


Fig. 11. Statistical analysis results for N2 averages of increased trials and grand average of all remaining trials using four different methods. Valence (V) × Negative-category (NC) are the within-subject factors. V∗NC is the interactive effect between valence and negative-category factors.

trial-averaged group PCA ($\bar{\mathbf{Z}}_{N2} = \sum \mathbf{u}_{k_1}^{TGPCA} \circ \mathbf{y}_{k_1}^{TGPCA}$, k_1 is 2, 5, 9, and 26 separately) and single-trial-based group PCA ($\frac{1}{E}\sum_{e=1}^{E}(\mathbf{\Phi})_e^{N2} = \frac{1}{E}\sum_{e=1}^{E}(\sum \mathbf{u}_{k_2}^{SGPCA} \circ \mathbf{y}_{k_2}^{SGPCA}(e))$, k_2 is 3, 10, 12, and 16 separately) when all four factor associated with N2 were chosen (see Fig. 4). Therefore, as shown in Section 4.1, we therefore observed the highly similar reconstructed waves, topographies, CCs, Cronbach's alpha, spatial similarity, and statistical analysis results for the two group PCA methods.

Similar to one of concerns in the analysis of human behavioral data (Nesselroade and Molenaar, 2016, 2010), the proposed method (within-subject PCA) tries to explore the interindividual variability (i.e., the factors related to ERP components vary across participants) which is contrast with the group PCA methods that the factors are invariant. That is, all the participants completely share the same factor loadings (see Figs. 3 and 4) for the two group PCA methods, whereas the factors related to P2/N2 for the proposed method may vary as shown in Tables 1 and 2. For example, regarding the extraction of P2, Factors 7 and 10 (latencies are 136 ms and 168 ms, respectively, see Fig. 3b) are identified in single-trial-based group PCA analysis, but only one factor (latency is 136 ms, see Table 1) for Subject 6 is selected for within-subject PCA analysis. Therefore, only the information near 136 ms for the reconstructed P2 for Subject 6 is preserved for within-subject PCA (WPCA): $\mathbf{Z}_{WPCA}^{6} = \mathbf{u}_4^{WPCA}(6) \circ \mathbf{y}_4^{WPCA}(6)$. Whereas the information of the reconstructed P2 around 136 ms and 168 ms via single-trial-based group PCA (we called it SGPCA) are retained: $\mathbf{\Phi}_{SGPCA}^{6} = \mathbf{u}_7^{SGPCA} \circ \mathbf{y}_7^{SGPCA}(6) + \mathbf{u}_{10}^{SGPCA} \circ \mathbf{y}_{10}^{SGPCA}(6)$. As a result, we can observe the differences of the reconstructed P2/N2 wave and topographies vary among the three PCA-based methods (see Figs. 5 and 6) (because $\mathbf{\Phi}_{WPCA}^{6}$ is not identical/similar to $\mathbf{Z}_{SGPCA}^{6}$). Similarly, the different factors used to reconstruct P2 and N2 (see Figs. 3 and 4, Tables 1 and 2) lead to differential results between P2 and N2 when using the same PCA method.

5.2. Effect of the number of trials on P2/N2 results

We found that the effect of the number of trials on PCA decomposition depends on the analyzed ERP components like the conclusion of early report for conventional time domain analysis (Boudewyn et al., 2018).

Pearson CCs of waveforms for P2/N2 between the averages with few trials and the grand average of all remaining trials were affected by the number of trials, but such affect depends on the analyzed ERP components and the used analysis method. Consistent with the previous study (Steele et al., 2016; Beauducel et al., 2000), as shown in Fig. 8a and b, CCs of two group PCA methods were higher than for the conventional time-domain analysis, that is, few trials were required for group PCA methods to obtain the reliable P2 and N2. It is no surprising that CCs of both P2 and N2 for the two group PCA methods was higher than the proposed PCA method because both group PCA methods try to extract the common features of ERP components from all participants. Similarly, few trials were needed to obtain a stable N2 for the proposed method than for the conventional time-domain analysis. However, the proposed method needed more trials than for the conventional time-domain analysis for P2.

Also, we calculated the Cronbach's alpha for increased trials to show the extent of homogeneous for the waveforms of all participants for a specific number of trials, which also showed variant across ERP components and across the measure methods (see Fig. 8c and d). Similar to the findings by Thigpen et al. (2017), we also observed that the consistency of N2 for the used methods is slightly better than for P2. In contrast with the previous report (Steele et al., 2016; Beauducel et al., 2000), Cronbach's alpha of P2 for conventional time-domain analysis is higher than the one for three PCA-based methods. All Cronbach's alpha values for four methods reached the excellent level (more than 0.9).

Moreover, we computed the spatial similarity of topographies between participants across the number of trials to examine the spatial internal consistency across participants. As far as we know, this is the

first time to use spatial similarity to evaluate the homogeneity of an ERP in spatial domain. More trials (about 30) were needed for P2 with low spatial similarity to obtain spatial stable P2 and about 15 trials are required for N2 with high spatial similarity (0.75). This suggests that ERP component with high spatial similarity needs few trials to get stable measures than the other ERP components with low spatial similarity.

Finally, the statistical results of P2/N2 for increased trials reflected that the effect of the number of trials also depends on the ERP components being analyzed (Figs. 10 and 11). Specifically, we can use about 30 trials to obtain the equivalent statistical results of P2 as all remaining trials do for all used methods (Fig. 10). For N2, as shown in Fig. 11, about 20 trials for the proposed method can achieve the same results as for all trials, but more trials will be needed for other methods (e.g., 30 trials). This indicates that the proposed method may capture the variability of an individual that group PCA methods cannot do.

5.3. Limitations and future directions

There are some potential limitations in the present study. Firstly, we only extracted P2/N2 from neurotypical participants, and we did not compared with other populations (older people, infants, etc.) and influence of the number of trials on PCA decomposition of other ERP components has not been further studied as well. Secondly, the comparison of yields between PCA and ICA was not included in this study. Many applications of ICA indicated that ERPs of interest can also be efficiently extracted from both single-trial EEG and averaged ERP datasets (Lee et al., 2016; Wessel, 2018; Cong et al., 2013; Rissling et al., 2014) so that it is worth investigating the information of the desired ERP using ICA from either average traces or single-trial traces. Thirdly, the identification of factors associated with the desired ERP components seems to be a subjective method in this study, which only takes temporal and spatial properties into consideration. As described in the application of ICA on the single-trial EEG from single subject (Rissling et al., 2014), we can also identify the factors related to ERPs of interest by using characteristics of the factor topography, factor waveform, factor spectra, and factor dipoles. Furthermore, we merely investigated the effects of trial numbers on the internal consistency of ERPs and statistical results, but the interactive effects between subjects and trials are not further studied as in past report (Boudewyn et al., 2018) because few subjects (i.e., 22) are involved in current study. Finally, the performance of PCA was evaluated by computing the correlation of peak amplitudes/topographies/latency between the original simulated data and PCA-extracted data in previous studies (Beauducel et al., 2000; Beauducel and Debener, 2003; Dien et al., 2007; Dien, 2010b), but we did not do such work in this study because latency scores and peak amplitude scores could not be validly obtained for many participants as in previous studies (Kappenman et al., 2021; Luck, 2014). Although we concluded that the proposed method did overall better results than the other three methods, this conclusion is needed to further verify via the simulation data. Following this line, we also need to further examine if the measurement invariant across subjects was seriously violated for group PCA methods compared with within-subject PCA.

6. Conclusion

Overall, it is first time to obtain P2 and N2 from single-trial EEG data of individuals for extracting ERP variability using PCA. We also extended the investigation of the minimum of trials that is used to obtain a stable P2 and N2 to the proposed method and the two group PCA methods. Moreover, the reconstructed P2 and N2 by the proposed method are more spatially consistent than the other three methods. N2 has internally consistent after about 30 trials are averaged for conventional time-domain analysis, trial-averaged group PCA, and single-trial-based group PCA, whereas within-subject PCA can yield a stable N2 with about 20 trials based on correlation coefficients of waves, Cronbach's alpha, spatial similarity, and statistical analysis.

Disclosure

The draft version of the manuscript has been submitted as a preprint: https://www.biorxiv.org/content/10.1101/2021.03.10.434892v3.

CRediT authorship contribution statement

Guanghui Zhang: Conceptualization, Data curation, Methodology, Software, Visualization, Writing – original draft. **Xueyan Li:** Visualization, Writing – original draft. **Yingzhi Lu:** Data curation; Data preprocessing, Writing – original draft. **Timo Tiihonen:** Supervision, Writing – original draft. **Zheng Chang:** Supervision, Writing – original draft. **Fengyu Cong:** Conceptualization, Project administration, Supervision, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have share the link to the used data and codes in manuscript.

Acknowledgments

We would like to thank the reviewers who contributed their time to review manuscripts for the three rounds of revision, and thanks for their remarkable comments which significantly improved the manuscript quality.

This work was supported by National Natural Science Foundation of China (Grant No. 91748105), National Foundation in China (No. JCKY2019110B009 & 2020-JCJQ-JJ-252) and the Fundamental Research Funds for the Central Universities in Dalian University of Technology in China [DUT20LAB303 & DUT20LAB308], and the scholarships from China Scholarship Council (No. 201806060165). This study is to memorize Prof. Tapani Ristaniemi for his great help to Guanghui Zhang, Fengyu Cong and the other authors.

Open access funding provided by University of Jyväskylä (JYU).

Appendix A. Model for temporal PCA from view of blind source separation

Simply, we here taking the procedure of trial-averaged group PCA as an example to explain how to extract ERP components at the group level with the view of blind source separation (BSS). As we described in Section 1.1, for the matrix $\mathbf{Z} \in \Re^{T\times M}$ (rows of the matrix correspond to the samples, and columns are multiplications of electrodes, conditions, and participants) formed by the ERP data of multi-participants with multi-conditions, it can be represented by Eq. (1).

Because the signals in Eq. (1) are mixtures of stimulus related signals and other signals (e.g., noise), we usually use few sources (number is K) associated with stimulus onset to represent the whole original data. We therefore need to estimate the number of stimulus related sources which is realized by some methods, for example, the accumulative explained variance. This procedure is known as dimensionality reduction. Thus, $\mathbf{Z}$ can be converted to a new matrix $\mathbf{X} \in \Re^{K\times M}$ (Cong et al., 2011a,b; Zhang et al., 2020).

$$\mathbf{X} = \mathbf{V}^T\mathbf{Z} = \mathbf{V}^T\mathbf{A}\mathbf{S} = \mathbf{H}\mathbf{S}, \tag{A.1}$$

where $\mathbf{V}^T \in \Re^{K\times T}$ is the dimensionality reduction matrix obtained by applying some PCA algorithms to matrix $\mathbf{Z}^T$ and the kth column of $\mathbf{V}$ is the kth eigenvector of the covariance matrix $\mathbf{Z}\mathbf{Z}^T$ (Cao et al., 2003). $\mathbf{X}$ is the principal component (PC) matrix and each row represents one PC (Cao et al., 2003). $\mathbf{H} = \mathbf{V}^T\mathbf{A} \in \Re^{K\times K}$ is a mixing matrix.

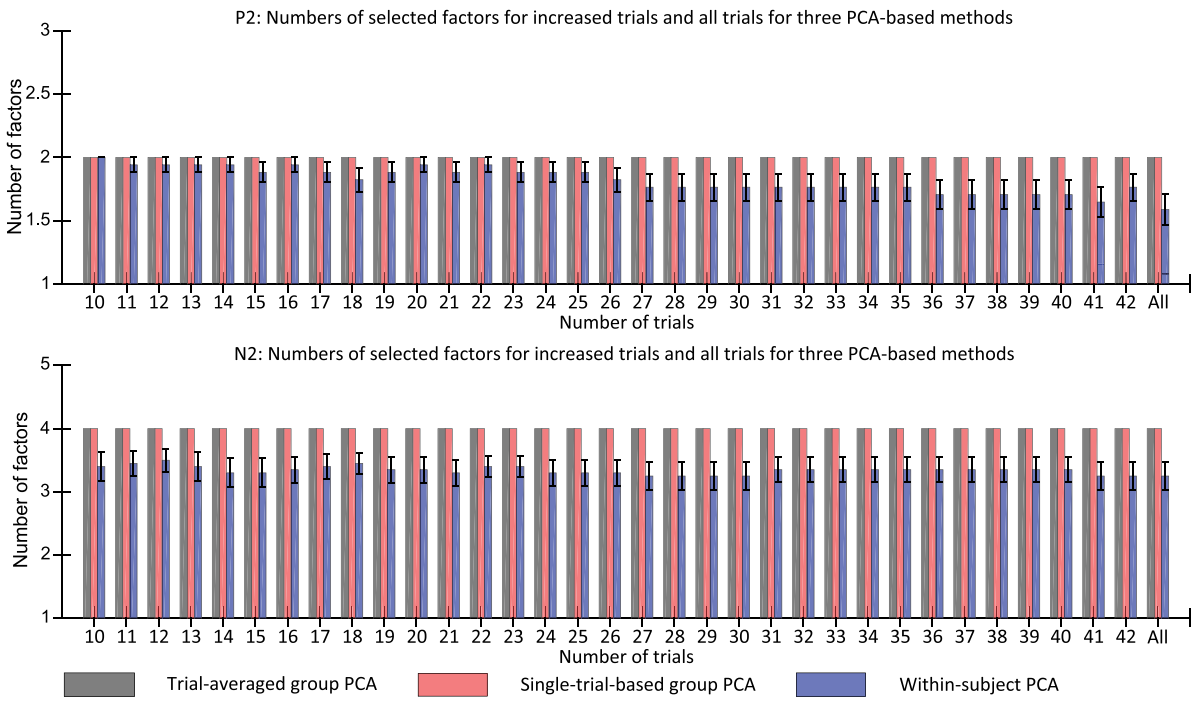


Fig. C.1. Numbers of selected factors associated with P2/N2 for the increasing trials and all trials for three PCA-based strategies. Each bar represents the number of identified factors for a specific number of trials for each PCA-based method. Error bars show the standard error of the mean of single participant values for within-subject PCA.

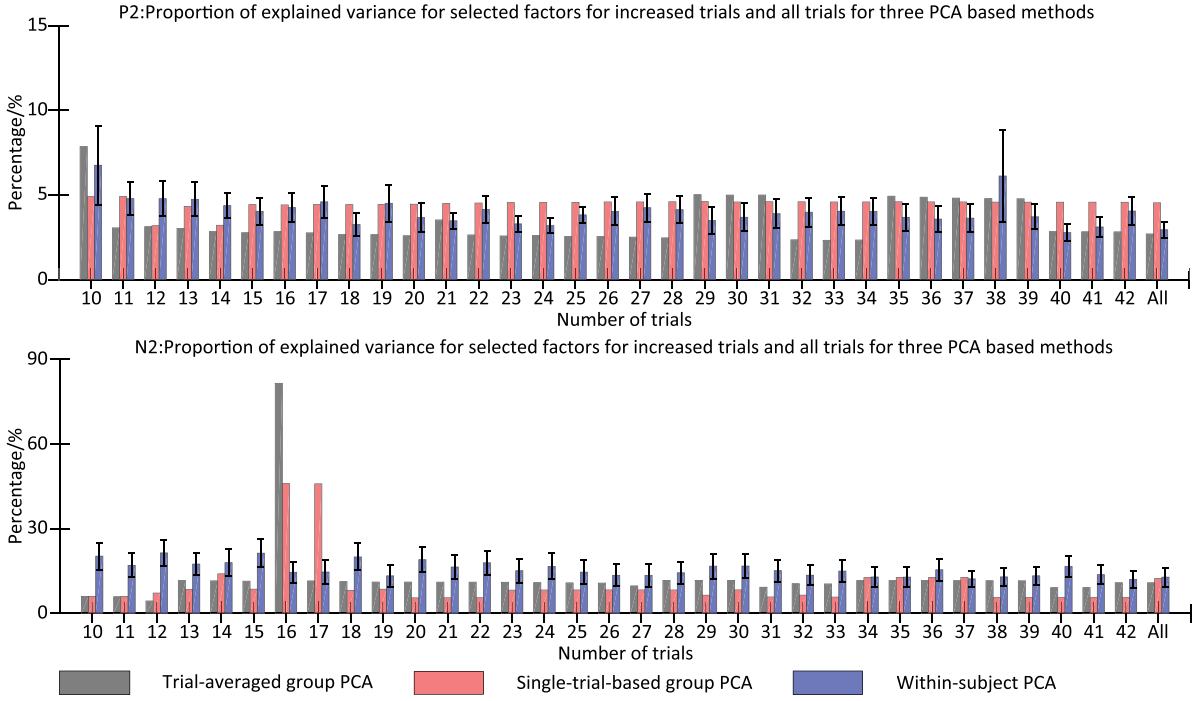


Fig. C.2. Proportion of explained variance for selected factors for both P2 (upper panel) and N2 (lower panel) for three PCA-based methods. Each bar represents total proportion of the factors for a specific number of trials for each PCA-based method. Error bars show the standard error of the mean of single participant values for within-subject PCA.

In this study, we seek an un-mixing matrix $\mathbf{W} \in \Re^{K\times K}$ (or transformation matrix) by using Promax rotation (Richman, 1986; Hendrickson and White, 1964) and dimensionality reduction matrix $\mathbf{V}$ to decompose $\mathbf{X}$ into a sum of several factors. $\mathbf{W}$ is typically normalized and $diag(\mathbf{W}^T\mathbf{W}) = \mathbf{I} \in \Re^{K\times K}$. Once the un-mixing matrix is generated, its inverse matrix $\mathbf{B} = \mathbf{W}^{-1}$ is used to estimate $\mathbf{H}$. And we can also use the unmixing matrix $\mathbf{W}$ to convert $\mathbf{X}$ into an estimated component matrix:

$$\mathbf{Y} = \mathbf{WX} = \mathbf{WHS} = \mathbf{CS}, \tag{A.2}$$

here, each row of the estimated source matrix $\mathbf{Y} \in \Re^{K\times M}$ represents topography of each factor (i.e., rotated factor scores) and it is scaled version of $\mathbf{S}$; $\mathbf{C} = \mathbf{WH}$ is a global matrix.

Generally, we need to choose several factors derived from $\mathbf{X}$ for further analysis (Comon and Jutten, 2010), and thus, the theory of

back-projection is used to analyze these factors simultaneously as applied in the previous studies (Dien, 1998; Cong et al., 2011a,b; Makeig et al., 1997, 1999). In the matrix–vector form, the back-projection is the outer product of kth column of $\mathbf{B}$ with kth row of estimated factor matrix $\mathbf{Y}$:

$$\mathbf{Q}_k = \mathbf{b}_k \circ \mathbf{y}_k, \tag{A.3}$$

where $\mathbf{Q}_k \in \Re^{K\times M}$ represents the back-projected signals at all the electrodes for kth selected factor and it estimates the partial information from $\mathbf{X}$; ' $\circ$ ' denotes the outer product of two vectors. k is used to avoid confusion with r in Eq. (1).

Under global optimization, only one nonzero element exists in each row and each column of matrix $\mathbf{C}$. That is, the extracted kth factor can uniquely represent rth unknown scaled source (Cong et al., 2011a,b):

$$\mathbf{Q}_k = \mathbf{b}_k \circ \mathbf{y}_k = \mathbf{h}_r \circ \mathbf{s}_r. \tag{A.4}$$

Note that k may be not equal to r because we do not know what both $\mathbf{h}_r$ and $\mathbf{s}_r$ are.

Turning back to Eq. (1), kth factor generated from $\mathbf{Z}$ is selected to project back onto electrode fields and this procedure can be described as below:

$$\hat{\mathbf{Z}}_k = \mathbf{u}_k \circ \mathbf{y}_k, \tag{A.5}$$

$$\mathbf{U} = \mathbf{VB}, \tag{A.6}$$

where $\mathbf{U} \in \Re^{T\times K}$ is the estimation of $\mathbf{A}$ in Eq. (1)(i.e., the kth rotated factor loading) and its columns correspond to the factor variance and rows are time courses.

In the application of PCA, a desired ERP is often decomposed into several factors because the latency and phase of the ERP vary among different subjects. Therefore, those factors need to project back onto the electrode fields simultaneously based on the following rule (Cong et al., 2011a,b; Zhang et al., 2020):

$$\begin{aligned}\hat{\mathbf{Z}} &= \left[\mathbf{u}_{k_1}, \ldots, \mathbf{u}_{k_i}\right]\left[\mathbf{y}_{k_1}, \ldots, \mathbf{y}_{k_i}\right]^T \\ &= \mathbf{u}_{k_1} \circ \mathbf{y}_{k_1} + \cdots + \mathbf{u}_{k_i} \circ \mathbf{y}_{k_i},\end{aligned} \tag{A.7}$$

where $k_1, \ldots,$ and k_i $(1 \le k_i < K)$ represent the orders of the identified factors. Each column of $\hat{\mathbf{Z}} \in \Re^{T\times M}$ is the reconstructed waveform of one or multiple ERP components for one condition for one participant at one electrode and each row represents the topography at a specific time point.

Appendix B. Mathematical models for three PCA-based methods

To simply explain the difference and common of the models among three PCA-based methods, we here define eth single-trial EEG data for pth participant $\mathbf{Z}_e^p \in \Re^{T\times(C\times J)}$ have T time points, C electrodes, and J conditions (E is number of trials and we assume that all conditions of each participant have the same number of trials E). Likewise, the eth single-trial EEG for all conditions and all participants are defined by using $\mathbf{Z}_e \in \Re^{T\times(C\times J\times P)}$ (P is the number of all participants). The trial-averaged ERP data of pth participant are then defined by using $\bar{\mathbf{Z}}^p \in \Re^{T\times(C\times J)} = \frac{1}{E}\sum_{e=1}^{E}\mathbf{Z}_e^p$.

According to Appendix A, for the trial-averaged group PCA (TGPCA), the trial-averaged ERP data $\bar{\mathbf{Z}} \in \Re^{T\times(C\times J\times P)} = [\bar{\mathbf{Z}}^1, \ldots, \bar{\mathbf{Z}}^p, \ldots, \bar{\mathbf{Z}}^P]$ or $\bar{\mathbf{Z}} = \frac{1}{E}\sum_{e=1}^{E}\mathbf{Z}_e$ for all participants can be decomposed into the sum of outer-product between K_1 temporal factors and K_1 spatial factors:

$$\bar{\mathbf{Z}} = \sum_{k_1=1}^{K_1} \mathbf{u}_{k_1}^{TGPCA} \circ \mathbf{y}_{k_1}^{TGPCA}, \tag{B.1}$$

where $\mathbf{y}_{k_1}^{TGPCA} \in \Re^{1\times(C\times J\times P)} = [\mathbf{y}_{k_1}^{TGPCA}(1), \ldots, \mathbf{y}_{k_1}^{TGPCA}(p), \ldots, \mathbf{y}_{k_1}^{TGPCA}(P)]$ and $\mathbf{y}_{k_1}^{TGPCA}(p) \in \Re^{1\times(C\times J)}$ is the topographies of J conditions for pth participant (see Fig. 3a). $\mathbf{u}_{k_1}^{TGPCA} \in \Re^{T\times 1}$ is the time course for $k_1{}^{th}$ factor and it is the common temporal feature across all conditions across all participants (see Fig. 3a and Fig. 4b).

Similarly, for single-trial-based group PCA (SGPCA), the single-trial EEG data of all participants is defined to be $\mathbf{\Phi} \in \Re^{T\times(C\times J\times P\times E)} = [\mathbf{Z}^1, \ldots, \mathbf{Z}^p, \ldots, \mathbf{Z}^P]$ or $\mathbf{\Phi} = [\mathbf{Z}_1, \ldots, \mathbf{Z}_e, \ldots, \mathbf{Z}_E]$, the matrix $\mathbf{\Phi}$ can be transformed into:

$$\mathbf{\Phi} = \sum_{k_2=1}^{K_2} \mathbf{u}_{k_2}^{SGPCA} \circ \mathbf{y}_{k_2}^{SGPCA}, \tag{B.2}$$

$\mathbf{u}_{k_2}^{SGPCA} \in \Re^{T\times 1}$ is also the common temporal wave across all trials across all conditions across all participants (see Fig. 4d). $\mathbf{y}_{k_2}^{SGPCA} \in \Re^{1\times(C\times J\times P\times E)} = [\mathbf{y}_{k_2}^{SGPCA}(e,1), \ldots, \mathbf{y}_{k_2}^{SGPCA}(e,p), \ldots, \mathbf{y}_{k_2}^{SGPCA}(e,P)]$ and $\mathbf{y}_{k_2}^{SGPCA}(e,p) \in \Re^{1\times(C\times J)}$ is the topography of J conditions for eth trial for pth participant (see Fig. 4e).

For within-subject PCA (WPCA), the single-trial EEG data $\mathbf{Z}^p \in \Re^{T\times(C\times J\times E)}$ for pth participant can be represented:

$$\mathbf{Z}^p = \sum_{k_3=1}^{K_3} \mathbf{u}_{k_3}^{WPCA}(p) \circ \mathbf{y}_{k_3}^{WPCA}(p), \tag{B.3}$$

$\mathbf{u}_{k_3}^{WPCA} \in \Re^{T\times 1}$ is the common temporal wave across all trials and all conditions for pth participant. $\mathbf{y}_{k_3}^{WPCA}(p) \in \Re^{1\times(C\times J\times E)} = [\mathbf{y}_{k_3}^{WPCA}(p,1), \ldots, \mathbf{y}_{k_3}^{WPCA}(p,e), \ldots, \mathbf{y}_{k_3}^{WPCA}(p,E)]$ and $\mathbf{y}_{k_3}^{WPCA}(p,e) \in \Re^{1\times(C\times J)}$ is the topography of J conditions for eth trial for pth participant.

Appendix C. Number of factors and explained variance for increased trials for three PCA-based methods

See Fig. C.1 and Fig. C.2.

References

Abdi, H., Williams, L.J., 2010. Principal component analysis. Wiley Interdiscip. Rev. Comput. Stat. 2 (4), 433–459.

Anowar, F., Sadaoui, S., Selim, B., 2021. Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE). Comp. Sci. Rev. 40, 100378. http://dx.doi.org/10.1016/j.cosrev.2021.100378.

Arbel, Y., Goforth, K., Donchin, E., 2013. The good, the bad, or the useful? The examination of the relationship between the feedback-related negativity (FRN) and long-term learning outcomes. J. Cogn. Neurosci. 25 (8), 1249–1260. http://dx.doi.org/10.1162/jocn_a_00385.

Barry, R.J., De Blasio, F.M., Fogarty, J.S., Karamacoska, D., 2016. ERP Go/NoGo condition effects are better detected with separate PCAs. Int. J. Psychophysiol. 106, 50–64. http://dx.doi.org/10.1016/j.ijpsycho.2016.06.003.

Barry, R.J., Steiner, G.Z., De Blasio, F.M., Fogarty, J.S., Karamacoska, D., MacDonald, B., 2020. Components in the P300: Don't forget the novelty P3! Psychophysiology 57 (7), e13371. http://dx.doi.org/10.1111/psyp.13371.

Beauducel, A., 2001. Problems with parallel analysis in data sets with oblique simple structure. Methods Psychol. Res. Online 6 (2), 141–157.

Beauducel, A., Debener, S., 2003. Misallocation of variance in event-related potentials: Simulation studies on the effects of test power, topography, and baseline-to-peak versus principal component quantifications. J. Neurosci. Methods 124 (1), 103–112. http://dx.doi.org/10.1016/S0165-0270(02)00381-3.

Beauducel, A., Debener, S., Brocke, B., Kayser, J., 2000. On the reliability of augmenting/reducing: Peak amplitudes and principal component analysis of auditory evoked potentials. J. Psychophysiol. 14 (4), 226. http://dx.doi.org/10.1027/0269-8803.14.4.226.

Bonmassar, C., Widmann, A., Wetzel, N., 2020. The impact of novelty and emotion on attention-related neuronal and pupil responses in children. Dev. Cogn. Neurosci. 42, 100766. http://dx.doi.org/10.1016/j.dcn.2020.100766.

Borkenau, P., Ostendorf, F., 1998. The big five as states: How useful is the five-factor model to describe intraindividual variations over time? J. Res. Personal. 32 (2), 202–221. http://dx.doi.org/10.1006/jrpe.1997.2206.

Boudewyn, M.A., Luck, S.J., Farrens, J.L., Kappenman, E.S., 2018. How many trials does it take to get a significant ERP effect? It depends. Psychophysiology 55 (6), e13049. http://dx.doi.org/10.1111/psyp.13049.

Cao, L., Chua, K.S., Chong, W., Lee, H., Gu, Q., 2003. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. Neurocomputing 55 (1–2), 321–336. http://dx.doi.org/10.1016/S0925-2312(03)00433-8.

Clayson, P.E., 2020. Moderators of the internal consistency of error-related negativity scores: A meta-analysis of internal consistency estimates. Psychophysiology 57 (8), e13583. http://dx.doi.org/10.1111/psyp.13583.

Clayson, P.E., Larson, M.J., 2013. Psychometric properties of conflict monitoring and conflict adaptation indices: Response time and conflict N2 event-related potentials. Psychophysiology 50 (12), 1209–1219. http://dx.doi.org/10.1111/psyp.12138.

Cohen, J., Polich, J., 1997. On the number of trials needed for P300. Int. J. Psychophysiol. 25 (3), 249–255. http://dx.doi.org/10.1016/S0167-8760(96)00743-X.

Comon, P., Jutten, C., 2010. Handbook of Blind Source Separation: Independent Component Analysis and Applications. Academic Press.

Cong, F., He, Z., Hämäläinen, J., Leppänen, P.H., Lyytinen, H., Cichocki, A., Ristaniemi, T., 2013. Validating rationale of group-level component analysis based on estimating number of sources in EEG through model order selection. J. Neurosci. Methods 212 (1), 165–172. http://dx.doi.org/10.1016/j.jneumeth.2012.09.029.

Cong, F., Kalyakin, I., Chang, Z., Ristaniemi, T., 2011a. Analysis on subtracting projection of extracted independent components from EEG recordings. Biomed. Eng./Biomed. Tech. 56 (4), 223–234. http://dx.doi.org/10.1515/bmt.2011.102.

Cong, F., Kalyakin, I., Ristaniemi, T., 2011b. Can back-projection fully resolve polarity indeterminacy of independent component analysis in study of event-related potential? Biomed. Signal Process. Control 6 (4), 422–426. http://dx.doi.org/10.1016/j.bspc.2010.05.006.

Cong, F., Leppänen, P.H., Astikainen, P., Hämäläinen, J., Hietanen, J.K., Ristaniemi, T., 2011c. Dimension reduction: additional benefit of an optimal filter for independent component analysis to extract event-related potentials. J. Neurosci. Methods 201 (1), 269–280. http://dx.doi.org/10.1016/j.jneumeth.2011.07.015.

de Cheveigné, A., 2020. ZapLine: A simple and effective method to remove power line artifacts. NeuroImage 207, 116356. http://dx.doi.org/10.1016/j.neuroimage.2019.116356.

de Cheveigné, A., Nelken, I., 2019. Filters: When, why, and how (not) to use them. Neuron 102 (2), 280–293. http://dx.doi.org/10.1016/j.neuron.2019.02.039.

De Winter, J.C., Dodou, D., 2012. Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. J. Appl. Stat. 39 (4), 695–710. http://dx.doi.org/10.1080/02664763.2011.610445.

Delorme, A., Sejnowski, T., Makeig, S., 2007. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. NeuroImage 34 (4), 1443–1449. http://dx.doi.org/10.1016/j.neuroimage.2006.11.004.

Dien, J., 1998. Addressing misallocation of variance in principal components analysis of event-related potentials. Brain Topogr. 11 (1), 43–55. http://dx.doi.org/10.1023/A:1022218503558.

Dien, J., 2010a. The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. J. Neurosci. Methods 187 (1), 138–145. http://dx.doi.org/10.1016/j.jneumeth.2009.12.009.

Dien, J., 2010b. Evaluating two-step PCA of ERP data with geomin, infomax, oblimin, promax, and varimax rotations. Psychophysiology 47 (1), 170–183. http://dx.doi.org/10.1111/j.1469-8986.2009.00885.x.

Dien, J., 2012. Applying principal components analysis to event-related potentials: A tutorial. Dev. Neuropsychol. 37 (6), 497–517. http://dx.doi.org/10.1080/87565641.2012.697503.

Dien, J., Beal, D.J., Berg, P., 2005. Optimizing principal components analysis of event-related potentials: Matrix type, factor loading weighting, extraction, and rotations. Clin. Neurophysiol. 116 (8), 1808–1825. http://dx.doi.org/10.1016/j.clinph.2004.11.025.

Dien, J., Frishkoff, G.A., Cerbone, A., Tucker, D.M., 2003. Parametric analysis of event-related potentials in semantic comprehension: Evidence for parallel brain mechanisms. Cogn. Brain Res. 15 (2), 137–153. http://dx.doi.org/10.1016/S0926-6410(02)00147-7.

Dien, J., Khoe, W., Mangun, G.R., 2007. Evaluation of PCA and ICA of simulated ERPs: Promax vs. Infomax rotations. Hum. Brain Mapp. 28 (8), 742–763. http://dx.doi.org/10.1002/hbm.20304.

Donchin, E., Heffley, E., 1978. Multivariate analysis of event-related potential data: A tutorial review. In: Otto, D. (Ed.), Multidisciplinary Perspectives in Event-Related Brain Potential Research. pp. 555–572.

Fischer, A.G., Klein, T.A., Ullsperger, M., 2017. Comparing the error-related negativity across groups: The impact of error-and trial-number differences. Psychophysiology 54 (7), 998–1009. http://dx.doi.org/10.1111/psyp.12863.

Fogarty, J.S., Barry, R.J., Steiner, G.Z., 2020. Auditory stimulus-and response-locked ERP components and behavior. Psychophysiology 57 (5), e13538. http://dx.doi.org/10.1111/psyp.13538.

He, Z., Cichocki, A., Xie, S., Choi, K., 2010. Detecting the number of clusters in n-way probabilistic clustering. IEEE Trans. Pattern Anal. Mach. Intell. 32 (11), 2006–2021. http://dx.doi.org/10.1109/TPAMI.2010.15.

Hendrickson, A.E., White, P.O., 1964. Promax: A quick method for rotation to oblique simple structure. Br. J. Stat. Psychol. 17 (1), 65–70. http://dx.doi.org/10.1111/j.2044-8317.1964.tb00244.x.

Horn, J.L., 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30 (2), 179–185. http://dx.doi.org/10.1007/BF02289447.

Huffmeijer, R., Bakermans-Kranenburg, M.J., Alink, L.R., Van IJzendoorn, M.H., 2014. Reliability of event-related potentials: The influence of number of trials and electrodes. Physiol. Behav. 130, 13–22. http://dx.doi.org/10.1016/j.physbeh.2014.03.008.

Huster, R.J., Raud, L., 2018. A tutorial review on multi-subject decomposition of EEG. Brain Topogr. 31 (1), 3–16. http://dx.doi.org/10.1007/s10548-017-0603-x.

Jung, T.-P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin. Neurophysiol. 111 (10), 1745–1758. http://dx.doi.org/10.1016/S1388-2457(00)00386-2.

Kappenman, E.S., Farrens, J.L., Zhang, W., Stewart, A.X., Luck, S.J., 2021. ERP CORE: An open resource for human event-related potential research. NeuroImage 225, 117465. http://dx.doi.org/10.1016/j.neuroimage.2020.117465.

Kayser, J., Tenke, C.E., 2003. Optimizing PCA methodology for ERP component identification and measurement: Theoretical rationale and empirical evaluation. Clin. Neurophysiol. 114 (12), 2307–2325. http://dx.doi.org/10.1016/S1388-2457(03)00241-4.

Kayser, J., Tenke, C.E., 2006a. Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. Clin. Neurophysiol. 117 (2), 348–368. http://dx.doi.org/10.1016/j.clinph.2005.08.034.

Kayser, J., Tenke, C.E., 2006b. Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates. Clin. Neurophysiol. 117 (2), 369–380. http://dx.doi.org/10.1016/j.clinph.2005.08.033.

Kleene, V., Lang-Keller, N., Steffen, M., Dreismann, V., Leue, A., 2021. Reliability of the N2-component in a modified 3-stimulus concealed information test: On the importance of excellent measurement accuracy. Biol. Psychol. 159, 108026. http://dx.doi.org/10.1016/j.biopsycho.2021.108026.

Larson, M.J., Baldwin, S.A., Good, D.A., Fair, J.E., 2010. Temporal stability of the error-related negativity (ERN) and post-error positivity (Pe): The role of number of trials. Psychophysiology 47 (6), 1167–1171. http://dx.doi.org/10.1111/j.1469-8986.2010.01022.x.

Lee, W.L., Tan, T., Falkmer, T., Leung, Y.H., 2016. Single-trial event-related potential extraction through one-unit ICA-with-reference. J. Neural Eng. 13 (6), 066010. http://dx.doi.org/10.1088/1741-2560/13/6/066010.

Lim, S., Jahng, S., 2019. Determining the number of factors using parallel analysis and its recent variants. Psychol. Methods 24 (4), 452. http://dx.doi.org/10.1037/met0000230.

Lu, Y., Luo, Y., Lei, Y., Jaquess, K.J., Zhou, C., Li, H., 2016. Decomposing valence intensity effects in disgusting and fearful stimuli: An event-related potential study. Soc. Neurosci. 11 (6), 618–626. http://dx.doi.org/10.1080/17470919.2015.1120238.

Luck, S.J., 2014. An Introduction to the Event-Related Potential Technique. MIT Press.

MacDonald, B., Barry, R.J., 2017. Significance and novelty effects in single-trial ERP components and autonomic responses. Int. J. Psychophysiol. 117, 48–64. http://dx.doi.org/10.1016/j.ijpsycho.2017.03.007.

MacDonald, B., Barry, R.J., Bonfield, R.C., 2015. Trials and intensity effects in single-trial ERP components and autonomic responses in a dishabituation paradigm with very long ISIs. Int. J. Psychophysiol. 98 (3), 394–412. http://dx.doi.org/10.1016/j.ijpsycho.2015.08.002.

Makeig, S., Jung, T.-P., Bell, A.J., Ghahremani, D., Sejnowski, T.J., 1997. Blind separation of auditory event-related brain responses into independent components. Proc. Natl. Acad. Sci. 94 (20), 10979–10984. http://dx.doi.org/10.1073/pnas.94.20.10979.

Makeig, S., Westerfield, M., Jung, T.-P., Covington, J., Townsend, J., Sejnowski, T.J., Courchesne, E., 1999. Functionally independent components of the late positive event-related potential during visual spatial attention. J. Neurosci. 19 (7), 2665–2680. http://dx.doi.org/10.1523/JNEUROSCI.19-07-02665.1999.

Male, A.G., Gouldthorp, B., 2020. Hemispheric differences in perceptual integration during language comprehension: An ERP study. Neuropsychologia 139, 107353. http://dx.doi.org/10.1016/j.neuropsychologia.2020.107353.

Marsh, H.W., Guo, J., Parker, P.D., Nagengast, B., Asparouhov, T., Muthén, B., Dicke, T., 2018. What to do when scalar invariance fails: The extended alignment method for multi-group factor analysis comparison of latent means across many groups. Psychol. Methods 23 (3), 524.

Meredith, W., 1993. Measurement invariance, factor analysis and factorial invariance. Psychometrika 58 (4), 525–543. http://dx.doi.org/10.1007/BF02294825.

Molenaar, P.C., Campbell, C.G., 2009. The new person-specific paradigm in psychology. Curr. Direct. Psychol. Sci. 18 (2), 112–117. http://dx.doi.org/10.1111/j.1467-8721.2009.01619.x.

Nesselroade, J.R., Molenaar, P., 2010. Emphasizing intraindividual variability in the study of development over the life span: Concepts and issues. http://dx.doi.org/10.1002/9780470880166.hlsd001002.

Nesselroade, J.R., Molenaar, P.C., 2016. Some behaviorial science measurement concerns and proposals. Multivar. Behav. Res. 51 (2–3), 396–412. http://dx.doi.org/10.1080/00273171.2015.1050481.

Olvet, D.M., Hajcak, G., 2009. The stability of error-related brain activity with increasing trials. Psychophysiology 46 (5), 957–961. http://dx.doi.org/10.1111/j.1469-8986.2009.00848.x.

Pontifex, M.B., Scudder, M.R., Brown, M.L., O'Leary, K.C., Wu, C.-T., Themanson, J.R., Hillman, C.H., 2010. On the number of trials necessary for stabilization of error-related brain activity across the life span. Psychophysiology 47 (4), 767–773. http://dx.doi.org/10.1111/j.1469-8986.2010.00974.x.

Richman, M.B., 1986. Rotation of principal components. J. Climatol. 6 (3), 293–335. http://dx.doi.org/10.1002/joc.3370060305.

Rietdijk, W.J., Franken, I.H., Thurik, A.R., 2014. Internal consistency of event-related potentials associated with cognitive control: N2/P3 and ERN/Pe. PLoS One 9 (7), e102672. http://dx.doi.org/10.1371/journal.pone.0102672.

Rissling, A.J., Miyakoshi, M., Sugar, C.A., Braff, D.L., Makeig, S., Light, G.A., 2014. Cortical substrates and functional correlates of auditory deviance processing deficits in schizophrenia. NeuroImage: Clin. 6, 424–437. http://dx.doi.org/10.1016/j.nicl.2014.09.006.

Rushby, J.A., Barry, R.J., 2009. Single-trial event-related potentials to significant stimuli. Int. J. Psychophysiol. 74 (2), 120–131. http://dx.doi.org/10.1016/j.ijpsycho.2009.08.003.

Rushby, J.A., Barry, R.J., Doherty, R.J., 2005. Separation of the components of the late positive complex in an ERP dishabituation paradigm. Clin. Neurophysiol. 116 (10), 2363–2380. http://dx.doi.org/10.1016/j.clinph.2005.06.008.

Sai, C.Y., Mokhtar, N., Arof, H., Cumming, P., Iwahashi, M., 2017. Automated classification and removal of EEG artifacts with SVM and wavelet-ICA. IEEE J. Biomed. Health Inf. 22 (3), 664–670. http://dx.doi.org/10.1109/JBHI.2017.2723420.

Scharf, F., Widmann, A., Bonmassar, C., Wetzel, N., 2022. A tutorial on the use of temporal principal component analysis in developmental ERP research–opportunities and challenges. Dev. Cogn. Neurosci. 101072. http://dx.doi.org/10.1016/j.dcn.2022.101072.

Steele, V.R., Anderson, N.E., Claus, E.D., Bernat, E.M., Rao, V., Assaf, M., Pearlson, G.D., Calhoun, V.D., Kiehl, K.A., 2016. Neuroimaging measures of error-processing: Extracting reliable signals from event-related potentials and functional magnetic resonance imaging. NeuroImage 132, 247–260. http://dx.doi.org/10.1016/j.neuroimage.2016.02.046.

Thigpen, N.N., Kappenman, E.S., Keil, A., 2017. Assessing the internal consistency of the event-related potential: An example analysis. Psychophysiology 54 (1), 123–138. http://dx.doi.org/10.1111/psyp.12629.

Wessel, J.R., 2018. A neural mechanism for surprise-related interruptions of visuospatial working memory. Cerebral Cortex 28 (1), 199–212. http://dx.doi.org/10.1093/cercor/bhw367.

Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemometr. Intell. Lab. Syst. 2 (1–3), 37–52. http://dx.doi.org/10.1016/0169-7439(87)80084-9.

Wood, J.M., Tataryn, D.J., Gorsuch, R.L., 1996. Effects of under-and overextraction on principal axis factor analysis with varimax rotation. Psychol. Methods 1 (4), 354. http://dx.doi.org/10.1037/1082-989X.1.4.354.

Zhang, G., Li, X., Cong, F., 2020. Objective extraction of evoked event-related oscillation from time-frequency representation of event-related potentials. Neural Plast. 2020, http://dx.doi.org/10.1155/2020/8841354.