Title

Assessing the effectiveness of spatial PCA on SVM-based decoding of EEG data

Authors

Zhang, Guanghui

Carrasco, Carlos D

Winsler, Kurt

et al.

Publication Date

2024-06-01

DOI

10.1016/j.neuroimage.2024.120625

Peer reviewed

Published in final edited form as:

Neuroimage. 2024 June ; 293: 120625. doi:10.1016/j.neuroimage.2024.120625.

Assessing the effectiveness of spatial PCA on SVM-based decoding of EEG data

Guanghui Zhang[a,b,*], **Carlos D. Carrasco**[b], **Kurt Winsler**[b], **Brett Bahle**[b], **Fengyu Cong**[c,d,e], **Steven J. Luck**[b]

[a]Research Center of Brain and Cognitive Neuroscience, Liaoning Normal University, Dalian, Liaoning, 116029, China

[b]Center for Mind and Brain, University of California-Davis, Davis, California, 95618, USA

[c]School of Biomedical Engineering, Faculty of Medicine, Dalian University of Technology, Dalian, Liaoning, 116024, China

[d]Faculty of Information Technology, University of Jyvaskyla, Jyvaskyla, 40014, Finland

[e]Key Laboratory of Social Computing and Cognitive Intelligence, Ministry of Education, Dalian University of Technology, Dalian, Liaoning, 116024, China

Abstract

Principal component analysis (PCA) has been widely employed for dimensionality reduction prior to multivariate pattern classification (decoding) in EEG research. The goal of the present study was to provide an evaluation of the effectiveness of PCA on decoding accuracy (using support vector machines) across a broad range of experimental paradigms. We evaluated several different PCA variations, including group-based and subject-based component decomposition and the application of Varimax rotation or no rotation. We also varied the numbers of PCs that were retained for the decoding analysis. We evaluated the resulting decoding accuracy for seven common event-related potential components (N170, mismatch negativity, N2pc, P3b, N400, lateralized readiness potential, and error-related negativity). We also examined more challenging decoding tasks, including decoding of face identity, facial expression, stimulus location, and stimulus orientation. The datasets also varied in the number and density of electrode sites. Our findings indicated that none of the PCA approaches consistently improved decoding performance related to no PCA, and the application of PCA frequently reduced decoding performance. Researchers should therefore be cautious about using PCA prior to decoding EEG data from similar experimental paradigms, populations, and recording setups.

Keywords

EEG; MVPA; group-based PCA; subject-based PCA; dimensionality reduction

[*]Corresponding author zhang.guanghui@foxmail.com; Address:Research Center of Brain and Cognitive Neuroscience, Liaoning Normal University, No. 850, Huanghe Road, Shahekou District, Dalian 116029, China. (Guanghui Zhang).

Declaration of interest

The authors declare no competing financial interests.

1. Introduction

Over the past few decades, there has been an increasing interest in applying machine learning (ML) classification to the electroencephalogram (EEG) and to event-related potentials (ERPs) in the context of scientific research. When applied to neural data, ML-based classification is commonly termed multivariate pattern analysis (MVPA) or decoding. The growing interest in ML-based classification is partly driven by the ability of this approach to detect subtle differences between stimulus classes that cannot be detected with traditional univariate approaches. For example, ML-based analyses have shown that scalp electrical signals can be used to distinguish between individual faces (Bae, 2021), between slightly different orientations being held in working memory (Bae & Luck, 2018), and between natural scenes with positive versus negative emotional content (Bo et al., 2022). Moreover, Carrasco et al. (2024) found that ML-based analyses often resulted in greater effect sizes than univariate analyses across a broad range of common ERP paradigms. Thus, ML-based analyses can make it possible to answer important scientific questions that would otherwise be intractable. The goal of the present study was to determine whether applying principal component analysis (PCA) during preprocessing can further increase the accuracy of ML-based classification in research designed to understand human brain activity [1].

In EEG/ERP studies, decoding is typically performed separately at each sample point following the onset of a stimulus (e.g., every 4 ms for a sampling rate of 250 Hz). The result is a waveform showing decoding accuracy as a function of time, leveraging the high temporal resolution of the EEG/ERP signal. For example, in a face recognition study, the decoder would attempt to use the pattern of voltage values across electrode sites at a given latency to determine whether a participant saw Face A or Face B. The decoder is trained on data from a subset of the trials so that it can “learn” how the scalp patterns differ between Face A and Face B, and then the decoder is tested with data from the trials that were left out from training. In most (but not all) MVPA EEG studies designed to understand human brain activity, multiple trials are averaged together for each training and testing case to increase the signal-to-noise ratio (Grootswagers et al., 2020). Decoding can also be performed in the frequency domain (Bae & Luck, 2018; Foster et al., 2016, 2017).

Figure 1 a illustrates the general logic behind EEG/ERP decoding in a simplified scenario with only two electrodes. The ML algorithm is given labeled training cases, which are essentially points in a 2-dimensional space defined by the amplitudes at the two electrodes. The algorithm then “learns” the best way to partition the 2-dimensional space to separate Face A and Face B. The decoder guesses the identity of unlabeled test cases by determining whether they fall into the part of the space corresponding to Face A or to Face B. In a real study, the number of dimensions would be far larger, with one dimension for each electrode site, but the general principle is the same.

[1]Note that the use of MVPA for understanding human brain activity has very different needs from the use of MVPA to perform real-world applications such as brain-computer interfaces and the prediction of psychopathology (Hebart & Baker, 2018). The present study focuses on the former, and the conclusions may not generalize to the latter. In addition, we have focused on maximizing decoding accuracy, not on calculation efficiency, because accuracy is the main concern in most current scientific applications of MVPA. Calculation efficiency may be important for some scientific applications (e.g., large datasets, closed-loop feedback studies) and would be a useful topic for future research.

There are many different ML algorithms that can be used to partition the space, including support vector machines (SVMs; (Boser et al., 1992), artificial neural networks (Basheer & Hajmeer, 2000 Yegnanarayana, 2009; Fisher, 1936), the random forest approach (Breiman, 2001), and linear discriminant analysis (Fisher, 1936). SVMs are particularly well suited to most EEG/ERP experiments because they are easy to implement and perform well with small training sets. Indeed, Trammel et al. (2023) found that SVMs yielded better performance than linear discriminant analysis and random forests in two language-related ERP paradigms. We therefore used SVMs in the present study.

PCA is commonly used for dimensionality reduction prior to SVM-based decoding (Bai et al., 2007; Fink et al., 2016; Hashmi et al., 2022; Kongwudhikunakorn et al., 2021; Rahman et al., 2020; Swarnkar et al., 2016; Zaree et al., 2023; Zhang et al., 2019). The component scores from the principal components (PCs) can be used as the input to the SVM decoder instead of the voltage values from the EEG channels (Figure 1b). In theory, PCs with small eigenvalues will mainly represent noise, and eliminating them might yield more robust decoding [2]. For example, the decoding could be based on the PCs that together explain 90% of the variance in the EEG (see Figure 1c). However, when the differences in voltage patterns between stimulus classes are subtle, these differences may not be present in the PCs that account for large amounts of variability (e.g., PCs with large eigenvalues). Thus, it is an empirical question whether PCA-based dimensionality reduction will improve ML-based decoding.

The present study tested whether PCA preprocessing improves SVM decoding performance across a broad range of experimental paradigms. There are many possible ways to implement PCA in the context of decoding, and we explored four common approaches. In some PCA implementations, the data from all participants are concatenated into a single dataset and a single set of PCs is derived (group-based PCA; Figure 2 a). These PCs are then used to obtain component scores separately from each individual participant for decoding. This approach has the advantage of maximizing the amount of data used to derive the PCs, but it has the disadvantage of disregarding potentially important differences between participants. An alternative approach is therefore to derive the PCs separately for each subject (subject-based PCA; Figure 2 b). We compared both of these approaches.

We also compared the application of Varimax rotation, a commonly used technique for spatial PCA (Dien, 2012; Kayser & Tenke, 2003), with no rotation. This gave us four basic PCA pipelines: group- versus subject-based PCA crossed with Varimax rotation versus no rotation. In addition, we varied the number of PCs that were retained for the decoding analysis. Each retained PC served as a feature (dimension) for training and testing the SVM. We compared all of these PCA-based decoding results with decoding that was performed on the basis of the untransformed voltages from each EEG channel, with each channel serving as a feature.

[2]Reducing the number of dimensions may also be beneficial by reducing the processing time needed for decoding. However, that is typically a concern with engineering applications and not with scientific applications.

To examine these different types of PCA in a systematic and generalizable manner, we examined the effects of PCA on decoding across a broad range of EEG/ERP experimental paradigms. This included six commonly used paradigms from the ERP CORE (Compendium of Open Resources and Experiments; Kappenman et al. (2021)) that yielded seven ERP components: N170, mismatch negativity (MMN), N2pc, P3b, N400, lateralized readiness potential (LRP), and error-related negativity (ERN). In these paradigms, we performed binary decoding of the primary experimental variable (e.g., faces versus cars for the N170 component, targets versus nontargets for the P3b component).

Those effects are easily detectable with conventional univariate methods, so we also examined three experimental paradigms in which more subtle stimulus classes must be distinguished. In one of these paradigms, the stimulus consisted of 16 different photographs of faces, combining 4 face identities with 4 emotional expressions (Bae, 2021). We decoded face identity independently of emotional expression and vice versa. The other two paradigms required participants to perceive and remember which of 16 different orientations was presented on each trial (Bae & Luck, 2018). We decoded which of the 16 orientations was presented, separately during perceptual processing and working memory maintenance. In one of these paradigms, the stimuli were presented at 16 different locations, and we also decoded the stimulus location. In the orientation experiments, we decoded frequency-domain data as well as time-domain data. Note that the paradigms also varied in the number and density of electrodes. To preview our findings, we found no evidence that any of the tested varieties of PCA was consistently helpful in increasing decoding accuracy.

2. Method

All analyses in the present study were performed on publicly available EEG datasets that had already been preprocessed to the point of epoching. The following sections will briefly describe the stimuli, tasks, recording methods, and preprocessing methods for each dataset. Complete descriptions can be found in the original papers describing these datasets, which we will refer to as ERP CORE (Kappenman et al. (2021); data available at https://doi.org/10.18115/D5JW4R), Orientations (Bae & Luck (2018); data available at https://osf.io/tgzew/), and Faces (Bae (2021); data available at https://osf.io/tgzew/). We will then describe the PCA and decoding analyses that were carried out for the present study. The studies were approved by the UC Davis Institutional Review Board and the Arizona State University Institutional Review Board.

The data analysis procedures were implemented with MATLAB 2021b (MathWorks Inc), using EEGLAB v2023.0 (Delorme & Makeig, 2004) combined with ERPLAB v10.00 (Lopez-Calderon & Luck 2014). The scripts are available at https://osf.io/tgzew/.

2.1. Participants

The participants in all datasets were college students who reported normal visual acuity and no history of major neurological disorders ($N = 40$ for ERP CORE; $N = 16$ for Orientations; $N = 22$ for Faces).

2.2. Experimental paradigms

The ERP CORE dataset contains data from six separate paradigms that can be used to isolate seven different ERP components. Each task was approximately 10 minutes long, and every participant received all six tasks in a single session. A face perception paradigm was used for the N170 component (Figure 3a). On each trial, one of four stimulus classes appeared in the center of the display (face, car, scrambled face, and scrambled car). Participants were instructed to press one of two buttons on each trial to indicate whether the stimulus was a "texture" (scrambled face or scrambled car) or an "object" (face or car). Only the face and car data are used in the present study. A passive auditory oddball paradigm was used for the mismatch negativity (MMN; Figure 3b). Participants heard a sequence of 80 dB standard stimuli (p = .8) and 70 dB deviant stimuli (p = .2). They watched a silent video and were instructed to ignore the auditory stimuli. An active visual oddball paradigm was used for the P3b component (Figure 3c). Participants viewed a randomized sequence of letters (A, B, C, D, and E; p = .2 for each). One of these letters was assigned as the target for a given block of trials. Participants were instructed to press one button if a given stimulus was the target (p = .2) and a different button if the stimulus was one of the nontargets (p = .8). A word pair judgment paradigm was used for the N400 component (Figure 3d). Each trial consisted of a red prime word followed by a green target word, and participants were instructed to press one of two buttons on each trial to indicate whether the green target word was semantically related or unrelated to the red prime word. An Eriksen flankers paradigm (Figure 3e) was used for the lateralized readiness potential (LRP) and the error-related negativity (ERN). On each trial, a central arrowhead that pointed either left or right was surrounded on both sides by arrowheads oriented in the same direction (congruent trials) or in the opposite direction (incongruent trials). Participants were instructed to report the direction of the central arrowhead on each trial by pressing a button with their left or right hand. A visual search paradigm was used for the N2pc component (Figure 3f). Each stimulus array consisted of one pink square, one blue square, and 22 black squares. In each trial block, either pink or blue was designated the target color. The side that contained the target color was unpredictable, but the target and non-target colors consistently appeared on opposite sides. Participants were instructed to press one of two buttons on each trial to indicate the location (top or bottom) of a gap in the square of the attended color. In all tasks, participants were instructed to maintain gaze on a central fixation point.

The Orientations dataset contains data from two separate paradigms, each of which was tested in a different group of participants. In the first paradigm (Orientations-1; Figure 3g), a teardrop shape (the sample stimulus) was presented for 200 ms in the center of the display at the beginning of each trial. There were 16 possible teardrop orientations (every 22.5°). After a 1300-ms blank delay period, participants were required to reproduce the orientation of the initial teardrop by using a mouse to adjust the orientation of a test teardrop. In the second paradigm (Orientations-2; Figure 3h), the sample teardrop was presented at one of 16 randomly selected orientations (every 22.5° around an invisible circle of locations). Participants were again required to reproduce the orientation at the end of the delay period, but the test teardrop was presented at a different random location. There was no need to remember the location of the teardrop across the delay period.

The Faces dataset contains data from a single experiment (Figure 3i). On each trial, a face was presented for 500 ms in the center of the display followed by a 1000-ms blank delay. The face was selected at random from a set of 16 faces created by combining four face identities (ID 1, ID 2, ID 3, and ID 4) with four emotional expressions (neutral, fearful, happy, and angry). On 89% of trials, the delay was followed by the next face stimulus. On the other 11% of trials, the delay was followed by a probe display containing four faces. The participant was instructed to indicate which face in the probe display matched the immediately preceding face in identity or expression. Identity probes and emotional expression probes were randomly intermixed, so participants needed to remember both the identity and the expression of a given face.

2.3. EEG recording and preprocessing

In the ERP CORE paradigms, the EEG was recorded at 1024 Hz with a Biosemi ActiveTwo recording system (Biosemi B.V.) using a fifth-order sinc antialiasing filter with a half-power cutoff at 204.8 Hz. The EEG was recorded from 30 scalp sites (FP1/2, Fz/3/4/7/8, FCz/3/4/, Cz/3/4/5/6, CPz, Pz/3/4/7/8, PO3/4//7/8/9/10, Oz/1/2) along with electrooculogram (EOG) electrodes adjacent to the eyes and below the right eye.

In the Faces and Orientations paradigms, the EEG was recorded at 500 Hz with a Brain Products actiCHamp recording system (Brain Products GmbH) using a cascaded integrator-comb antialiasing filter with a half-power cutoff at 130 Hz. The recordings included the left and right mastoids and the same EOG electrode locations used in the ERP CORE dataset, along with 59 scalp sites in the Faces paradigm (FP1/2, AFz/3/4/7/8, Fz/1/2/3/4/5/6/7/8, FCz/1/2/3/4/5/6, Cz/1/2/3/4/5/6, T7/8, CPz/1/2/3/4/5/6, TP7/8, Pz/1/2/3/4/5/6/7/9/10, POz/3/4/7/8, Oz/1/2) and 27 scalp sites in the Orientations paradigms (FP1/2, Fz/3/4/7/8, Cz/3/4, Pz/3/4/5/6/7/8/9/10, POz/3/4/7/8, Oz/1/2).

A detailed description of the ERP CORE preprocessing steps is provided by Kappenman et al. (2021). Key steps included the following. The continuous EEG signal was downsampled to 256 Hz. The EEG was then referenced to the average of the P9 and P10 electrodes (close to the left and right mastoids) except that the average of all scalp sites was used as the reference in the N170 paradigm. A noncausal Butterworth filter was applied (bandpass = 0.1 – 30 Hz, 12 dB/oct roll off). Independent component analysis (ICA; Jung et al. (2000a,b)) was used to correct the data for blinks and eye movements. The ICA-corrected data were epoched and baseline-corrected using the time windows shown in Table 1. Bad channels were interpolated using ERPLAB's spherical spline interpolation algorithm. Trials were rejected if they contained blinks that would have prevented perception of the stimulus, if they contained extreme values in any channel, or if they contained incorrect behavioral responses.

The preprocessing steps for the Faces and Orientations paradigms were identical to those for the ERP CORE except that the data were downsampled to 250 Hz and referenced to the average of the mastoids. In addition, no trials were rejected because of artifacts or behavioral errors (artifacts and errors were very rare in these dataset). See Bae (2021) and Bae & Luck (2018) for additional details and Table 1 for the epoch timing.

2.4. PCA approaches

As is typical when PCA is used for EEG dimensionality reduction, we applied a spatial PCA in which the EEG channels are the variables and the time samples are the observations. We applied four different PCA-based approaches. In the first approach (group-based PCA + no rotation), we first organized the single-trial EEG for all conditions and for all participants into a matrix of channels × time samples (see Figure 2a). The matrix was then decomposed into N principal components, where N is equal to the number of channels. The principal components were ordered from 1 to N according to the amount of variance in the EEG explained by each component, starting with the component that explained the most variance. For each participant, we created a component × time samples matrix of component scores. A subset of this matrix was then used for that participant's decoding analysis. That is, we used the first PC, the first three PCs, the first five PCs, and so on as the features for a given decoding run.

The second PCA approach (group-based PCA + rotation) was identical except that Varimax rotation was applied to the principal components prior to computing the component × time samples matrix. In the third PCA approach (subject-based PCA + no rotation), the PCA decomposition was performed on a single participant's matrix of channels × time samples and then used to create the components × time samples matrix for that participant prior to decoding (see Figure 2b). The fourth PCA approach (subject-based PCA + rotation) was identical to the third approach except that Varimax rotation was performed. We also performed decoding without PCA (using the channels × time samples matrix from a given participant as the input to the decoder).

Instead of varying the absolute number of PCs used for the decoding, it is possible instead to vary the number of PCs according to the percentage of EEG variance explained by the PCs. We therefore conducted a separate analysis in which we used the set of PCs that cumulatively explained 30%, 50%, 70%, 90%, 92%, 94%, 96%, 98% , and 100% of the total variance in the data matrix. The results for these analyses were similar to those obtained when we varied the absolute number of PCs used for decoding, so they are described only in the supplementary materials (Figures S1- S4).

As in most previous studies using PCA prior to EEG decoding, our approach was based on variance across all time points in the EEG epochs. This might not be optimal, because it includes time points that might not contain signals that are relevant for differentiating between the classes being decoded. We therefore attempted limiting the PCA to the time window of interest for a given ERP component (e.g., 300-500 ms for the P3b component). The results for this approach were very similar to the results using all time points, so we have not included this approach in the present manuscript.

2.5. SVM-based decoding methods

In the ERP CORE experiments, we decoded which of the two stimulus classes was present for a given paradigm (see Table 1). For example, we decoded whether the target word was semantically related or unrelated to the preceding prime word in the N400 paradigm. In the Faces paradigm, we performed two independent runs of decoding. In one, we decoded

which of the four face identities was present (collapsed across emotional expressions); in the other, we decoded which of the four emotional expressions was present (collapsed across identities). In the Orientations-1 and Orientations-2 paradigms, we decoded which of the 16 orientations was present (collapsing across stimulus locations in Orientations-2). In the Orientations-2 paradigm, we also decoded which of the 16 locations contained the stimulus (collapsing across stimulus orientations).Thus, the number of classes to be decoded was 2 in the ERP CORE paradigms, 4 in the Faces paradigm, and 16 in the Orientations paradigms.

Decoding was performed separately for each participant at each time sample. When 2 classes were being decoded, we used the Matlab function fitcsvm() to train the SVM. When more than 2 classes were being decoded, we used the Matlab function fitcecoc() to implement the error-correcting output codes approach for SVM training (Dietterich & Bakiri, 1994). In both cases, we used the Matlab function predict() to test the decoder.

A leave-one-out cross-validation approach was used to avoid overfitting (Bae, 2021; Bae & Luck, 2018; Carrasco et al., 2024). The procedure is illustrated in Figure 4. For each decoding analysis, we took the components × time points matrix (or the channels × time points matrix for the no-PCA analyses) for a given participant and divided it into separate matrices for each of the M stimulus classes. We created multiple averaged ERPs for each class from different random subsets of the trials and used these averaged data to train and test the SVM. From the total set of trials for a given class, we created J averages each containing L trials. In our experience, this method works best when L is between 10 and 20, and J is the largest number possible given that L and the number of available trials (see Table 2 for the values of M, J, and L used for each paradigm in the present study). In the Faces paradigm, for example, there were M = 4 face classes, with 160 trials per class, so we had J = 10 averages of L = 16 trials for each class. When the number of trials varied across classes (e.g., because of artifact rejection), we equated the number of trials used in the analysis for a given participant by subsampling from the available trials for that participant.

For each crossfold, we trained the decoder using J-1 of the averaged ERPs for each class, and then we tested the decoder by feeding the data from the one left-out averaged ERP from each class into the decoder. For example, in the Faces paradigm, in which the number of averages (J) was 10, we used 9 averaged ERPs for each of the four face classes to train the decoder, and then we tested the decoder with the remaining one average ERP for each class. This gave us four test cases. The process was then repeated nine more times, each time using a different set of J-1 training cases for each class and one test case for each class. Decoding accuracy was defined as the proportion of test cases that were correctly classified. Chance was 1/M (e.g., 1/4 or 0.25 for the 4-class decoding of face identity).

To increase the precision of the decoding accuracy, the process was repeated 100 times, with different random subsets of trials assigned to the averaged ERPs for each iteration. The final decoding accuracy was averaged across these 100 iterations. Note that this process was applied separately to each time point in the ERP waveforms, and each participant was decoded independently, so we obtained a separate decoding accuracy value at each time point for each participant.

The location and orientation information in the Orientations-1 and Orientations-2 datasets can be decoded from alpha-band EEG oscillations as well as from ERPs (Bae & Luck, 2018). To examine the effects of PCA on frequency-domain decoding, we filtered the EEG epochs using a bandpass filter from 8-12 Hz using the EEGLAB eegfilt() routine (which uses a two-way least-squares finite impulse response filter). The filtered epochs were then passed through the Hilbert transform and then squared to estimate the phase-independent power in the 8–12 Hz band at each time point. The set of power values across electrode sites at a given time point was then used as the input to the decoding procedure (in place of the set of voltage values across sites).

2.6. Time-window averaging and effect size calculation

Each decoding accuracy waveform contained hundreds of values, and we generated decoding accuracy waveforms for every PCA procedure and number of principal components retained. To distill the data down to a reasonable number of comparisons, we averaged the decoding accuracy across a predetermined time window in each paradigm (see Table 2). For the ERP CORE paradigms, these were simply the time windows used for the standard univariate analysis of a given ERP component. For the Faces and Orientations paradigms, we examined both a perceptual time window (100–500 ms) and a working memory time window (500–1000 ms).

In most scientific studies, a primary methodological challenge is to obtain decoding accuracy that is far enough away from chance to be consistent, so the best decoding method is typically the one that yields the greatest decoding accuracy. However, statistical power also depends on the amount of variability across participants, so it can also be useful to examine Cohen's d_z metric of effect size, which takes into account both the mean decoding accuracy and the variability across participants (Carrasco et al., 2024). We therefore computed Cohen's d_z for each analysis by taking the mean across participants (after time-window averaging) minus chance and dividing by the standard deviation across participants. Bootstrapping (10000 iterations) was used to estimate the standard error of the effect size. The pattern of results for Cohen's d_z was similar to that for mean decoding accuracy, so the d_z results are provided in supplementary materials.

3. Results

3.1. Easy, binary decoding cases

This section describes the results for the ERP CORE paradigms, all of which involve binary decoding of experimental conditions that produce very large differences in ERPs and are ordinarily analyzed with standard univariate analysis methods.

Figure 5 shows mean decoding accuracy as a function of the number of PCs that were retained for decoding (corresponding d_z values are shown in supplementary Figure S5). The horizontal lines show the results for the no-PCA analyses, which always included all of the scalp EEG channels. In all but one of the paradigms, we found that mean decoding accuracy increased as more PCs were retained. The one exception was the MMN paradigm, in which the number of PCs had relatively little effect. In all cases, the decoding accuracy

for the various PCA procedures was similar to or less than the decoding accuracy obtained without PCA. Thus, PCA produced no clear benefit in decoding accuracy, and in many cases PCA resulted in lower accuracy. When there were differences in performance among the four PCA approaches, the group-based PCA + no rotation approach was typically (but not always) the best. Similar results were obtained when the PCs were selected on the basis of the percentage of variance explained (supplementary Figure S1). Decoding was also not improved by performing decoding on ICs derived from ICA rather than on PCs derived from PCA (supplementary Figure S8).

Because the effects of the number of PCs was so different for the MMN paradigm, we took a close look at the PCA results for this paradigm. The factor loadings of the first PC for this paradigm had a scalp topography that closely resembled the topography of the MMN effect, and this may explain why adding more PCs had relatively little impact on decoding. However, this is probably a very unusual case.

Note that the decoding accuracy was below chance when a single PC was used for decoding in a few cases, leading to negative effect sizes. This is presumably because that one PC did not provide any information that could be used for decoding, and decoding accuracy was therefore equally likely to be above or below chance due to noise.

3.2. Difficult decoding cases

The remaining paradigms were much more challenging because the classes being decoded were very similar to each other (e.g., 4 different faces) and the number of classes was greater (4 or 16). These are the kinds of cases in which univariate analysis methods are typically insufficient and ML-based decoding methods are often necessary to distinguish between the classes. It is possible that PCA would be more helpful for such cases than for the relatively easy cases in the ERP CORE paradigms.

Figure 6 displays the mean decoding accuracy for each of these paradigms during the perceptual period (100–500 ms poststimulus) and the working memory maintenance period (500–1000 ms poststimulus). The corresponding d_z values are provided in supplementary Figure S6. In all cases, decoding accuracy increased as the number of PCs used for decoding increased. However, the maximum decoding accuracy was approximately the same as obtained without PCA. Similar results were obtained when the PCs were selected on the basis of the percentage of variance explained (supplementary Figures S2 and S3). The four different PCA approaches yielded similar results, although group-based PCA without rotation had a slight advantage when the number of PCs used for decoding was small. Decoding was also not improved in most cases by performing decoding on ICs derived from ICA rather than on PCs derived from PCA (supplementary Figure S9). The one exception was that subject-based ICA appeared to improve decoding accuracy in the 100-500 ms time period for the Orientations-2 dataset. However, given that this was an isolated case, it may be spurious.

Decoding accuracy was below chance in a few cases when a single PC was used for decoding, leading to negative effect sizes, presumably reflecting noise.

3.3. Time-frequency decoding

The analyses described so far were carried out on time-domain data. It is also possible to perform decoding after transforming the data into the time-frequency domain (Bae & Luck, 2018; Foster et al. 2016, 2017). For example, location and orientation can be decoded from the scalp distribution of alpha-band EEG oscillations (Bae & Luck, 2018; Foster et al., 2016, 2017). Alpha-band oscillations are large and therefore likely to be captured by PCs that account for significant variance, so we asked whether PCA might improve decoding performance for alpha-band signals in the Orientations-1 and Orientations-2 tasks.

Figure 7 displays the mean decoding accuracy for the alpha-band EEG signals (see supplementary Figure S7 for the corresponding d_z values). As was observed for the time-domain decoding in Figure 6, alpha-band decoding accuracy increased as the number of PCs used for decoding increased, but the maximum decoding accuracy was approximately the same as obtained without PCA. In addition, decoding accuracy was similar across the four PCA approaches. Similar results were obtained when the PCs were selected on the basis of the percentage of variance explained (supplementary Figure S4).

4. Discussion

We asked whether PCA (implemented in four different ways) improved EEG/ERP decoding performance across a broad range of experimental paradigms. To our knowledge, this is the first time that several different PCA approaches have been compared across a large number of paradigms, which is important for assessing the generality of the results. We found no evidence that any of these PCA approaches consistently improved decoding performance relative to decoding without PCA, either when decoding was relatively easy or when it was relatively challenging. PCA also produced no substantial benefit when time-frequency signals were decoded.

When PCA was used, decoding accuracy increased as the number of retained PCs increased (except for the MMN), but it was never substantially greater than the accuracy obtained without PCA. That is, PCA did not consistently yield decoding accuracy values more than one standard error beyond the decoding accuracy obtained without PCA. Similar results were obtained when ICA was used rather than PCA, except that performing the decoding on ICs did lead to improved decoding accuracy in one of the thirteen cases examined (see supplementary Figures S8 and S9).

The results were similar across multiple datasets that varied in the stimuli, the task, the number and nature of the classes to be decoded, and the number of electrode sites. Consequently, our finding that PCA provides minimal benefit likely generalizes across many different types of experiments. However, all of the datasets used similar participant populations (college students), similar recording systems, and similar preprocessing pipelines. It is possible that our PCA results would not generalize to datasets collected from other populations, such as infants and people with psychiatric disorders (Ashton et al. 2022; Masychev et al., 2021). In addition, they may not generalize to experimental paradigms or preprocessing pipelines that are very different from those used in the present study. For example, we applied ICA to correct for blinks and eye movements, and PCA might be useful

if artifacts are not corrected with ICA (although we found that the pattern of results was the same whether or not the preprocessing pipeline included artifact correction, see Figure S10). The effects of PCA might also be quite different for datasets that contain much higher levels of noise, such as those obtained outside the laboratory or without a gel-based conductor.

Nonetheless, the present results suggest that researchers should not assume that PCA will improve SVM-based decoding accuracy. Indeed, decoding accuracy might actually be reduced if too few PCs are retained. There are at least three possible reasons why PCA was ineffective in improving decoding accuracy in the present context. First, as noted by Mwangi et al. (2014), dimensionality reduction may be important primarily when the number of dimensions (predictors) available in the data is much larger than the number of observations (the small-n-large-p problem). For example, the number of predictors in fMRI research (e.g., 10,000 voxels) is often far larger than the number of observations (e.g., 20 trials). The ratio of predictors to observations is far smaller for the type of EEG decoding examined in the present study, where the number of predictors (channels) was comparable to or even smaller than the number of observations (trials). Second, we used a linear transformation to construct the principal components, and this transformation may fail to extract the optimal features for decoding. Nonlinear PCA algorithms (e.g., kernel PCA), which has been widely used in previous EEG studies (Guo et al. 2021; Patel et al., 2020; Widjaja et al., 2012), may be more effective. Third, the autocorrelation of the successive time points may reduce the information available to extract useful PCs (Vanhatalo & Kulahci, 2016).

It should be noted that we only used spatial PCA (i.e., the channels were the variables) before decoding in current study. Another alternative method, temporal PCA (i.e., the time samples are the variables), is often used to extract ERP components of interest (Scharf et al., 2022; Zhang et al., 2023). Temporal information is lost in temporal PCA, because the output is components × channels rather than components × time points. However, the temporal information can be reconstructed by projecting the factor scores onto electrode fields, that is, computing the outer product between factor loadings and factor scores (Zhang et al., 2023, 2020). Therefore, it is worth examining whether temporal PCA can yield better decoding performance in further research.

We found slightly better performance for unrotated PCs than for Varimax-rotated PCs in some cases (e.g., N170 and LRP), but the opposite result in other cases (the Orientations-2 dataset). This may simply reflect noise, but it is also possible that the effects of rotation depend on the nature of the dataset and experimental paradigm. It is also possible that other rotation methods would work better (see Dien (2010)), but that is beyond the scope of the present study.

There are also other PCA approaches that were not tested in the present study. For example, Gu et al. (2023) attempted to first extract PCs of interest using a sequence of temporal PCA followed by spatial PCA prior to decoding. Additional research is needed to determine whether other PCA approaches such as this might yield superior decoding.

It should be also noted that some researchers employed independent component analysis (ICA) prior to decoding analysis (Jin et al., 2020; Richhariya & Tanveer, 2018; Stewart

et al., 2014). A full examination of ICA would be beyond the scope of the present study, but we provided the preliminary ICA decoding results in the supplementary materials (see Figures S8-S9). We found that performing decoding on independent components yielded poorer decoding accuracy than performing decoding on the original channel voltages in most cases. However, it is possible that decoding from independent components would be advantageous in other populations (e.g., infants) and recording setups (e.g., dry electrodes).

Whereas PCA reduces the number of dimensions by creating new dimensions, another approach is to use data-driven approaches to directly reduce the number of electrodes used for the decoding. For example, regularization methods such as ridge regression and LASSO (Least Absolute Shrinkage and Selection Operator) regression can be used to perform shrinkage and feature selection in EEG decoding (Tibshirani, 1996; McDonald, 2009). Further research is needed to determine whether these regression-based approaches combined with regularization would be superior to SVMs for the types of decoding problems addressed in the present study.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was made possible by grants R01MH087450 and R01EY033329 from the National Institutes of Health to SJL.

References

Ashton K, Zinszer BD, Cichy RM, Nelson CA III, Aslin RN, & Bayet L (2022). Time-resolved multivariate pattern analysis of infant eeg data: A practical tutorial. Developmental Cognitive Neuroscience, 54, 101094. doi:10.1016/j.dcn.2022.101094. [PubMed: 35248819]

Bae G-Y (2021). The time course of face representations during perception and working memory maintenance. Cerebral Cortex Communications, 2, tgaa093. doi:10.1093/texcom/tgaa093. [PubMed: 34296148]

Bae G-Y, & Luck SJ (2018). Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. Journal of Neuroscience, 38, 409–422. doi:10.1523/JNEUROSCI.2860-17.2017. [PubMed: 29167407]

Bai O, Lin P, Vorbach S, Li J, Furlani S, & Hallett M (2007). Exploration of computational methods for classification of movement intention during human voluntary movement from single trial eeg. Clinical Neurophysiology, 118, 2637–2655. doi:10.1016/j.clinph.2007.08.025. [PubMed: 17967559]

Basheer IA, & Hajmeer M (2000). Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods, 43, 3–31. doi:10.1016/S0167-7012(00)00201-3. [PubMed: 11084225]

Bo K, Cui L, Yin S, Hu Z, Hong X, Kim S, Keil A, & Ding M (2022). Decoding the temporal dynamics of affective scene processing. NeuroImage, 261, 119532. doi:10.1016/j.neuroimage.2022.119532. [PubMed: 35931307]

Boser BE, Guyon IM, & Vapnik VN (1992). A training algorithm for optimal margin classifiers. In Proceedings of the fifth annual workshop on Computational learning theory (pp. 144–152). doi:10.1145/130385.130401.

Breiman L (2001). Random forests. Machine learning, 45, 5–32. doi:10.1023/A:1010933404324

Carrasco CD, Bahle B, Simmons AM, & Luck SJ (2024). Using multivariate pattern analysis to increase effect sizes for event-related potential analyses. Psychophysiology, (p. e14570). doi:10.1111/psyp.14570. [PubMed: 38516957]

Delorme A, & Makeig S (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods, 134, 9–21. doi:10.1016/j.jneumeth.2003.10.009. [PubMed: 15102499]

Dien J (2010). Evaluating two-step PCA of ERP data with geomin, infomax, oblimin, promax, and varimax rotations. Psychophysiology, 47, 170–183. doi:10.1111/j.1469-8986.2009.00885.x. [PubMed: 19761521]

Dien J (2012). Applying principal components analysis to event-related potentials: a tutorial. Developmental Neuropsychology, 37, 497–517. doi:10.1080/87565641.2012.697503. [PubMed: 22889342]

Dietterich TG, & Bakiri G (1994). Solving multiclass learning problems via error-correcting output codes. Journal of Artificial Intelligence Research, 2, 263–286. doi:10.1613/jair.105.

Fink BC, Steele VR, Maurer MJ, Fede SJ, Calhoun VD, & Kiehl KA (2016). Brain potentials predict substance abuse treatment completion in a prison sample. Brain and Behavior, 6, e00501. doi:10.1002/brb3.501. [PubMed: 27547503]

Fisher RA (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics, 7, 179–188. doi:10.1111/j.1469-1809.1936.tb02137.x.

Foster JJ, Bsales EM, Jaffe RJ, & Awh E (2017). Alpha-band activity reveals spontaneous representations of spatial position in visual working memory. Current Biology, 27, 3216–3223. doi:10.1016/j.cub.2017.09.031. [PubMed: 29033335]

Foster JJ, Sutterer DW, Serences JT, Vogel EK, & Awh E (2016). The topography of alpha-band activity tracks the content of spatial working memory. Journal of Neurophysiology, 115, 168–177. doi:10.1152/jn.00860.2015. [PubMed: 26467522]

Grootswagers T, Kennedy BL, Most SB, & Carlson TA (2020). Neural signatures of dynamic emotion constructs in the human brain. Neuropsychologia, 145, 106535. doi:10.1016/j.neuropsychologia.2017.10.016. [PubMed: 29037506]

Gu L, Jiang J, Han H, Gan JQ, & Wang H (2023). Recognition of unilateral lower limb movement based on EEG signals with ERP-PCA analysis. Neuroscience Letters, 800, 137133. doi:10.1016/j.neulet.2023.137133. [PubMed: 36801241]

Guo Y, Zhang Z, & Tang F (2021). Feature selection with kernelized multi-class support vector machine. Pattern Recognition, 117, 107988. doi:10.1016/j.patcog.2021.107988.

Hashmi MF, Kene JD, Kotambkar DM, Matte P, & Keskar AG (2022). An efficient P300 detection algorithm based on kernel principal component analysis-support vector machine. Computers & Electrical Engineering, 97, 107608. doi:10.1016/j.compeleceng.2021.107608.

Hebart MN, & Baker CI (2018). Deconstructing multivariate decoding for the study of brain function. Neuroimage, 180, 4–18. doi:10.1016/j.neuroimage.2017.08.005. [PubMed: 28782682]

Jin CY, Borst JP, & van Vugt MK (2020). Distinguishing vigilance decrement and low task demands from mind-wandering: A machine learning analysis of EEG. European Journal of Neuroscience, 52, 4147–4164. doi:10.1111/ejn.14863. [PubMed: 32538509]

Jung T-P, Makeig S, Humphries C, Lee T-W, Mckeown MJ, Iragui V, & Sejnowski TJ (2000a). Removing electroencephalographic artifacts by blind source separation. Psychophysiology, 37, 163–178. doi:10.1111/1469-8986.3720163. [PubMed: 10731767]

Jung T-P, Makeig S, Westerfield M, Townsend J, Courchesne E, & Sejnowski TJ (2000b). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clinical Neurophysiology, 111, 1745–1758. doi:10.1016/S1388-2457(00)00386-2. [PubMed: 11018488]

Kappenman ES, Farrens JL, Zhang W, Stewart AX, & Luck SJ (2021). ERP CORE: An open resource for human event-related potential research. NeuroImage, 225, 117465. doi:10.1016/j.neuroimage.2020.117465. [PubMed: 33099010]

Kayser J, & Tenke CE (2003). Optimizing PCA methodology for ERP component identification and measurement: theoretical rationale and empirical evaluation. Clinical neurophysiology, 114, 2307–2325. doi:10.1016/S1388-2457(03)00241-4. [PubMed: 14652090]

Kongwudhikunakorn S, Kiatthaveephong S, Thanontip K, Leelaarporn P, Piriyajitakonkij M, Charoenpattarawut T, Autthasan P, Chaisaen R, Dujada P, Sudhawiyangkul T et al. (2021). A pilot study on visually stimulated cognitive tasks for EEG-based dementia recognition. IEEE Transactions on Instrumentation and Measurement, 70, 1–10. doi:10.1109/TIM.2021.3120131. [PubMed: 33776080]

Lopez-Calderon J, & Luck SJ (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. Frontiers in Human Neuroscience, 8, 213. doi:10.3389/fnhum.2014.00213. [PubMed: 24782741]

Masychev K, Ciprian C, Ravan M, Reilly JP, & MacCrimmon D (2021). Advanced signal processing methods for characterization of schizophrenia. IEEE Transactions on Biomedical Engineering, 68, 1123–1130. doi:10.1109/TBME.2020.3011842. [PubMed: 33656984]

McDonald GC (2009). Ridge regression. Wiley Interdisciplinary Reviews: Computational Statistics, 1, 93–100. doi:10.1002/wics.14.

Mwangi B, Tian TS, & Soares JC (2014). A review of feature reduction techniques in neuroimaging. Neuroinformatics, 12, 229–244. doi:10.1007/s12021-013-9204-3. [PubMed: 24013948]

Patel R, Gireesan K, & Sengottuvel S (2020). Decoding non-linearity for effective extraction of the eye-blink artifact pattern from eeg recordings. Pattern Recognition Letters, 139, 42–49. doi:10.1016/j.patrec.2018.01.022.

Rahman MA, Hossain MF, Hossain M, & Ahmmed R (2020). Employing PCA and t-statistical approach for feature extraction and classification of emotion from multichannel EEG signal. Egyptian Informatics Journal, 21, 23–35. doi:10.1016/j.eij.2019.10.002.

Richhariya B, & Tanveer M (2018). EEG signal classification using universum support vector machine. Expert Systems with Applications, 106, 169–182. doi:10.1016/j.eswa.2018.03.053.

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. Developmental Cognitive Neuroscience, 54, 101072. doi:10.1016/j.dcn.2022.101072. [PubMed: 35123341]

Stewart AX, Nuthmann A, & Sanguinetti G (2014). Single-trial classification of EEG in a visual object task using ICA and machine learning. Journal of Neuroscience Methods, 228, 1–14. doi:10.1016/j.jneumeth.2014.02.014. [PubMed: 24613798]

Swarnkar R, Keskar A, Prasad PS, & Shivprakash N (2016). A new approach to detect P300 in a single trial based on PCA and SVM classifier. In 2016 IEEE Region 10 Symposium (TENSYMP) (pp. 355–360). IEEE. doi:10.1109/TENC0NSpring.2016.7519432

Tibshirani R (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology, 58, 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x.

Trammel T, Khodayari N, Luck SJ, Traxler MJ, & Swaab TY (2023). Decoding semantic relatedness and prediction from EEG: A classification method comparison. NeuroImage, 277, 120268. doi:10.1016/j.neuroimage.2023.120268. [PubMed: 37422278]

Vanhatalo E, & Kulahci M (2016). Impact of autocorrelation on principal components and their use in statistical process control. Quality and Reliability Engineering International, 32, 1483–1500. doi:10.1002/qre.1858.

Widjaja D, Varon C, Dorado A, Suykens JA, & Van Huffel S (2012). Application of kernel principal component analysis for single-lead-ECG-derived respiration. IEEE Transactions on Biomedical Engineering, 59, 1169–1176. doi:10.1109/TBME.2012.2186448. [PubMed: 22438200]

Yegnanarayana B (2009). Artificial neural networks. PHI Learning Pvt. Ltd.

Zaree M, Mohebbi M, & Rostami R (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. Biomedical Signal Processing and Control, 86, 105224. doi:10.1016/j.bspc.2023.105224.

Zhang G, Li X, Cong F et al. (2020). Objective extraction of evoked event-related oscillation from time-frequency representation of event-related potentials. Neural Plasticity, 2020. doi:10.1155/2020/8841354.

Zhang G, Li X, Lu Y, Tiihonen T, Chang Z, & Cong F (2023). Single-trial-based temporal principal component analysis on extracting event-related potentials of interest for an individual subject. Journal of neuroscience methods, 385, 109768. doi:10.1016/j.jneumeth.2022.109768. [PubMed: 36529386]

Zhang T, Chen W, & Li M (2019). Classification of inter-ictal and ictal EEGs using multi-basis MODWPT, dimensionality reduction algorithms and LS-SVM: A comparative study. Biomedical Signal Processing and Control, 47, 240–251. doi:10.1016/j.bspc.2018.08.038.

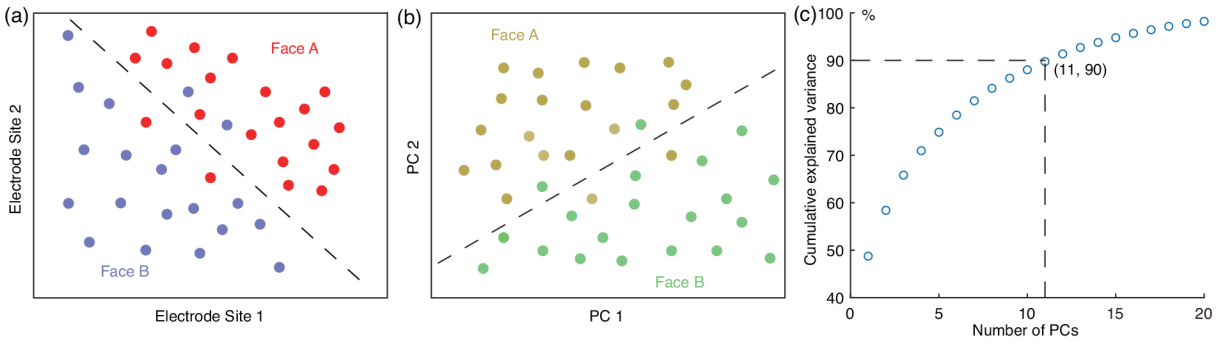

Figure 1:

(a) Example of EEG/ERP decoding with two electrodes. (b) Example of EEG/ERP decoding with two principal components (PCs). (c) Example of cumulative explained variance for principal component analysis.

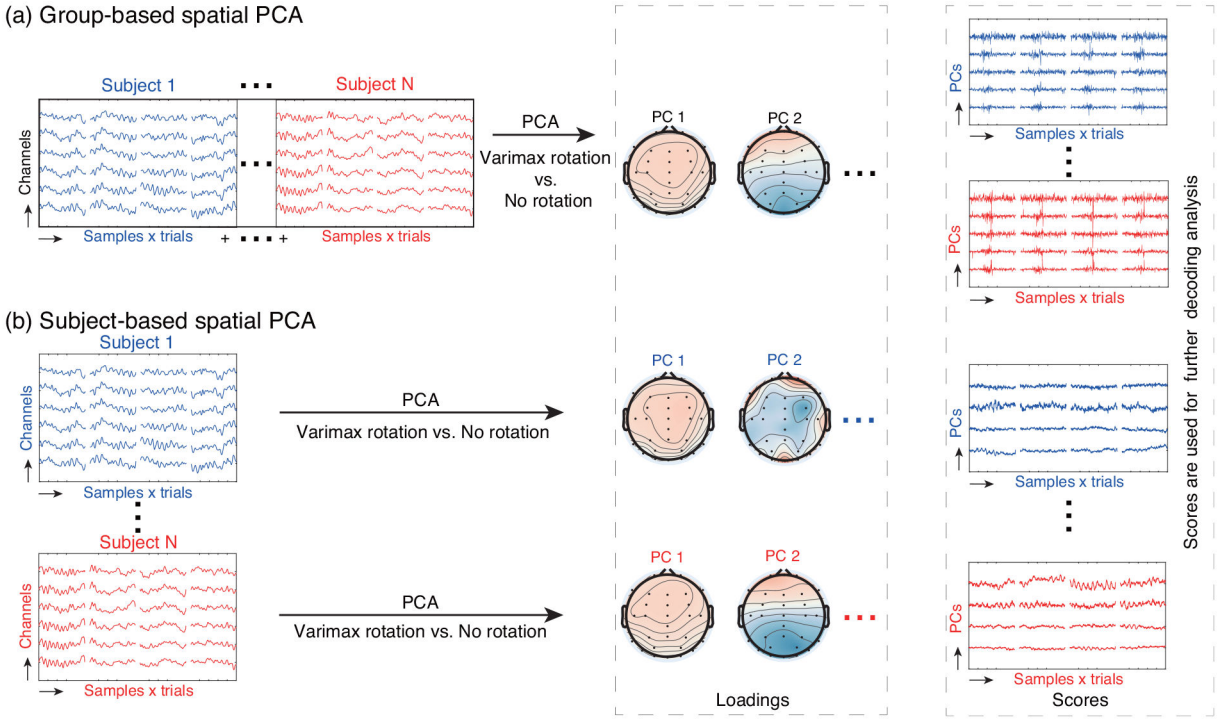

Figure 2:
Examples of group-based spatial-PCA and subject-based spatial-PCA for single-trial EEG data.

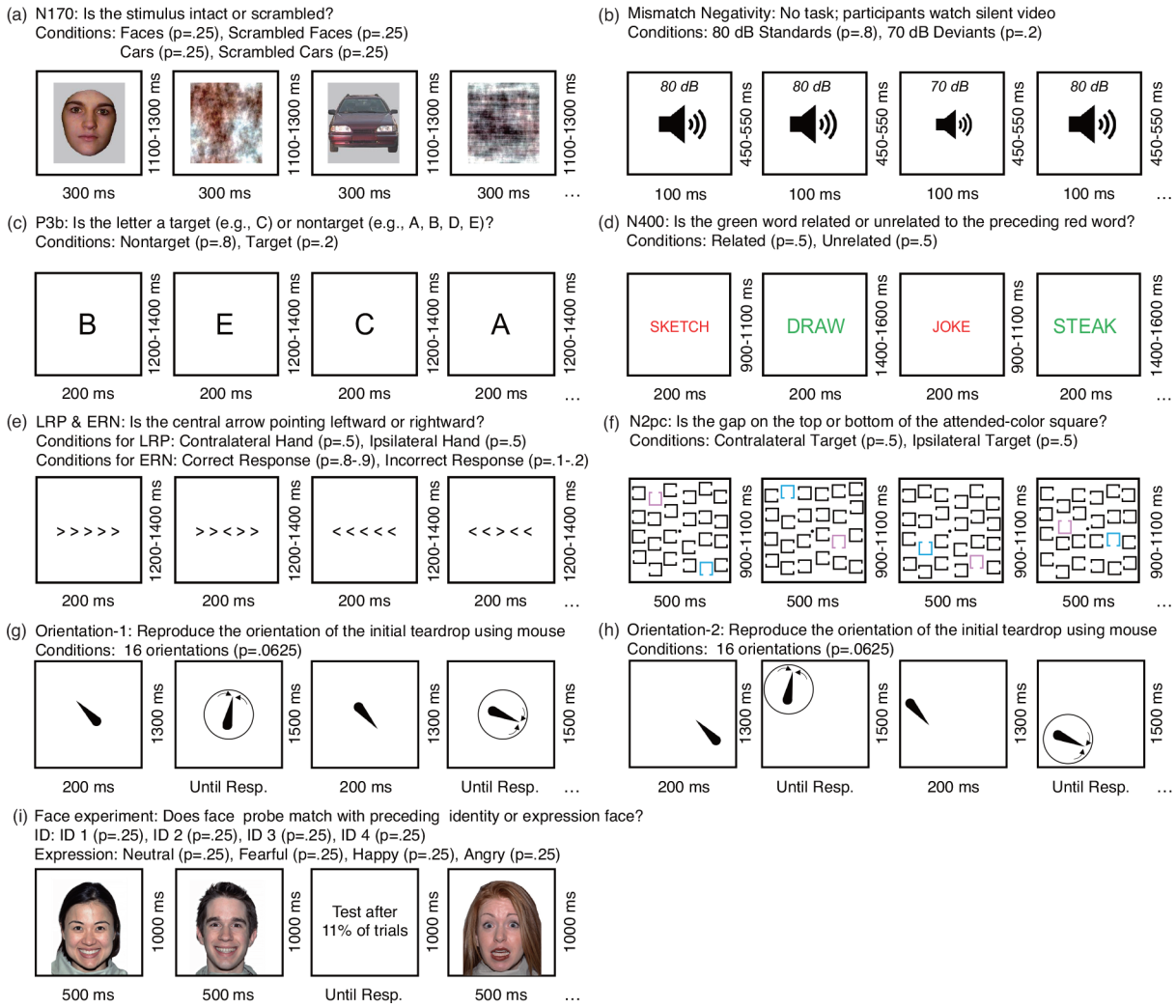

Figure 3:

Example trials from the experimental paradigms. (a) N170 paradigm. Participants reported whether a given stimulus was intact (faces or cars) or scrambled (scrambled faces or cars). The present study focused solely on the face and car trials. (b) Mismatch negativity paradigm. Participants viewed a silent video while task-irrelevant standard tones (80 dB, probability = .8) and deviant tones (70 dB, probability = .2) were presented. (c) P3b paradigm. The letters A, B, C, D, and E were presented in random order with a probability of .2 for each letter. One letter was defined as the target for a given trial block, and participants reported whether a given stimulus was the target or a nontarget. (d) N400 paradigm. Participants reported whether a green target word was semantically related or unrelated to a preceding red prime word. (e) Flankers task that was used to elicit the lateralized readiness potential (LRP) and error-related negativity (ERN). Participants reported the direction of the central arrow (leftward or rightward) while disregarding the surrounding arrows. (f) N2pc paradigm. Pink or blue was defined as the target for each trial block, and participants reported whether the gap on the target-colored square was at the top or bottom of the square. (g) Orientations-1 paradigm. Participants were required to store

the initial teardrop in memory, retain it over the 1300-ms delay period, and then report its orientation by using a mouse to adjust the orientation of a test teardrop. (g) Orientations-2 paradigm, which was identical to Orientations-1 except that the teardrop location varied randomly among 16 different locations. (i) Faces paradigm. On each trial, one of 16 faces was shown (one of four face identities expressing one of four emotions). On 11% of trials, memory for the most recent face was probed.

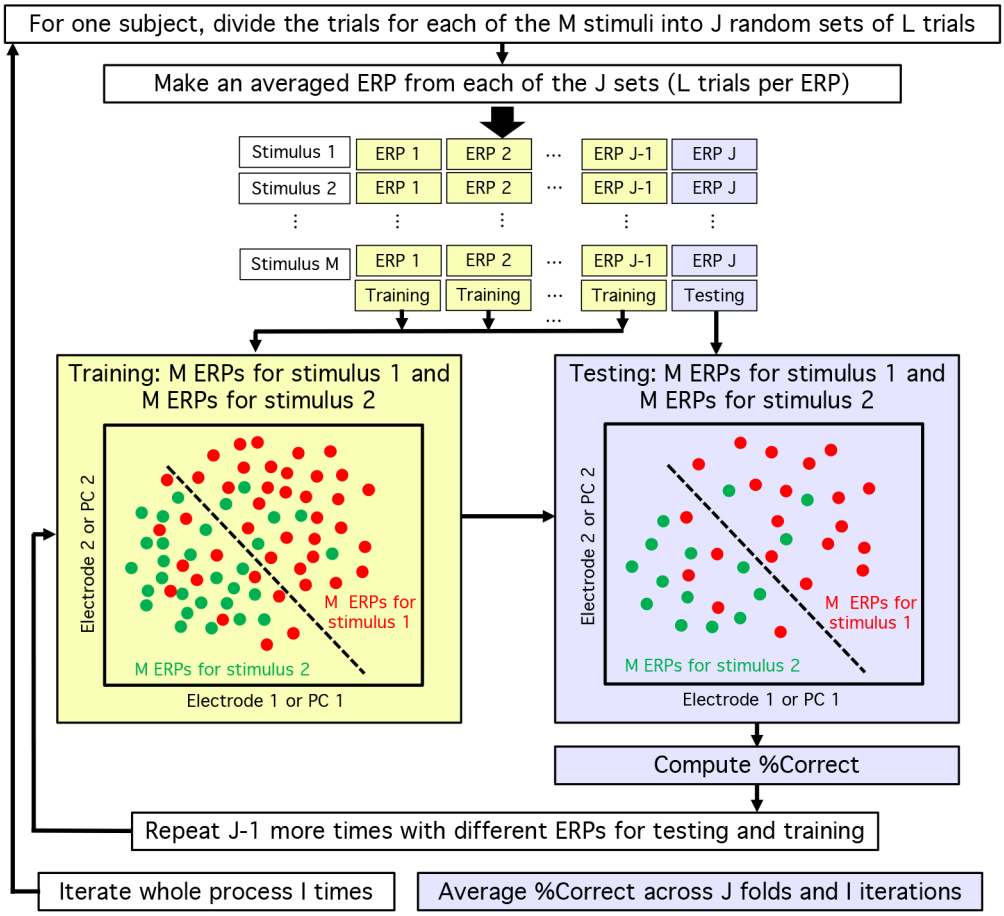

Figure 4:
Flowchart for decoding one subject at one time point.

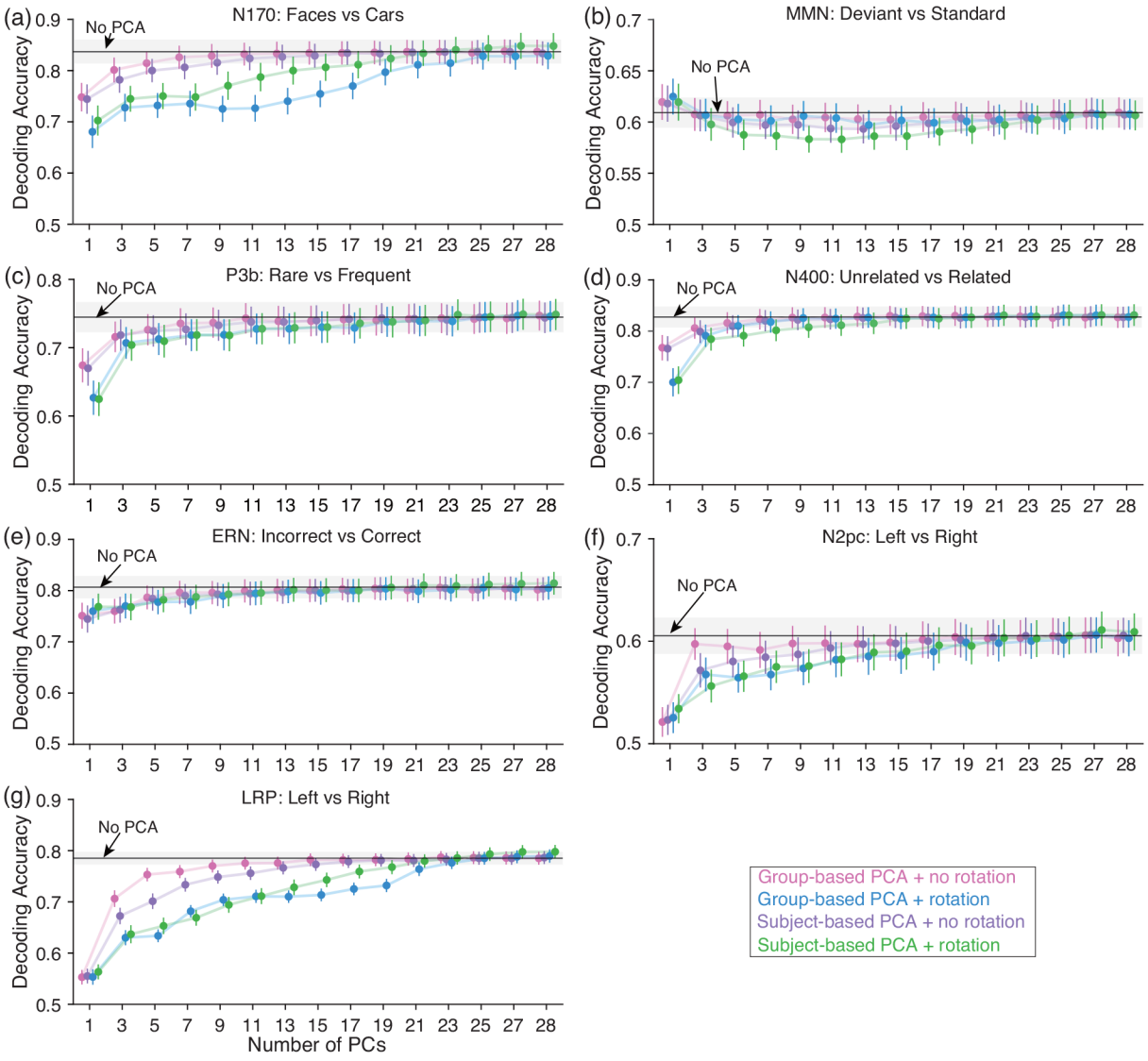

Figure 5:

Decoding performance for the ERP CORE paradigm, averaged across the time-window shown in Table 2. The black horizontal line indicates mean decoding accuracy without PCA, with the SEM shown as gray shading. Each point represents mean decoding accuracy for a given approach to principal component analysis (PCA), with the SEM shown as the error bar.

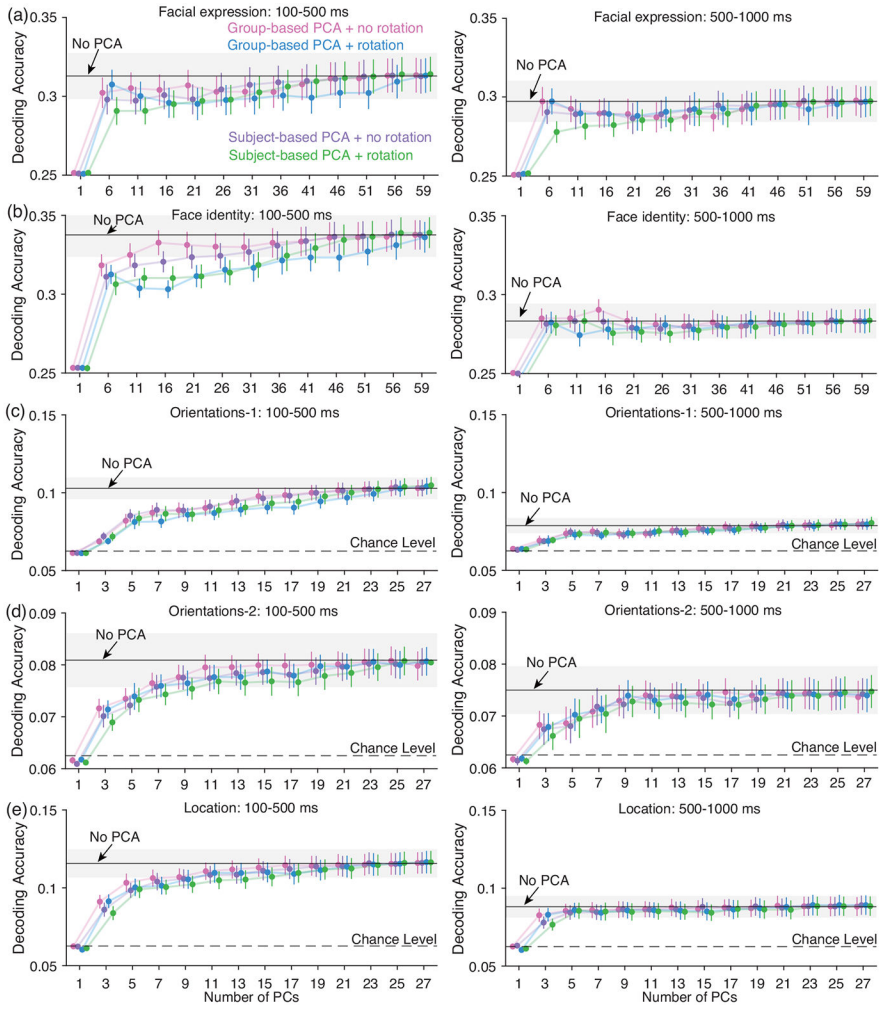

Figure 6:

Facial expression and facial identity decoding performance for the Faces dataset, orientation decoding performance for the Orientations-1 dataset, and orientation and location decoding performance for the Orientations-2 dataset. Decoding accuracy was averaged across two different time windows (100-500 ms to characterize perceptual processing and 500-1000 ms to characterize working memory maintenance). The black horizontal line indicates mean decoding accuracy without PCA, with the SEM shown as gray shading. Each point represents mean decoding accuracy for a given approach to principal component analysis (PCA), with the SEM shown as the error bar.

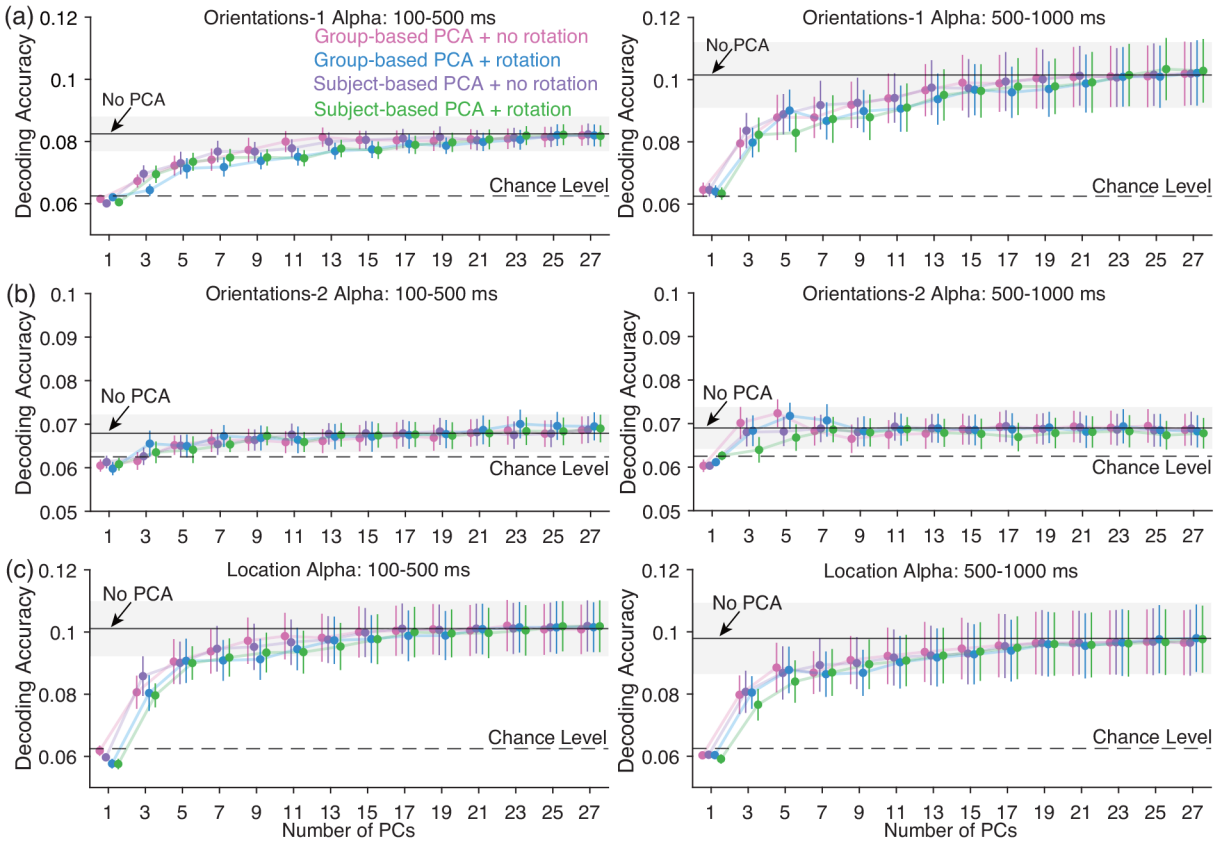

Figure 7:

Decoding performance using alpha-band data in the Orientations-1 and Orientations-2 datasets. Decoding accuracy was averaged across two different time windows (100-500 ms to characterize perceptual processing and 500-1000 ms to characterize working memory maintenance). The black horizontal line indicates mean decoding accuracy without PCA, with the SEM shown as gray shading. Each point represents mean decoding accuracy for a given approach to principal component analysis (PCA), with the SEM shown as the error bar.

Table 1:

Epoch window, baseline period,number of subjects, and number of classes used for each experiment.

Experiment	Epoch (ms)	Baseline period (ms)	Number of subjects	Number of classes (M)
N170	−200 to 800	−200 to 0	37	2 (faces vs. cars)
MMN	−200 to 800	−200 to 0	39	2 (deviants vs. standards)
P3b	−200 to 800	−200 to 0	34	2 (targets vs. nontargets)
N400	−200 to 800	−200 to 0	37	2 (related vs. unrelated)
ERN	−600 to 400	−600 to −400	36	2 (correct vs. error)
LRP	−800 to 200	−800 to −600	37	2 (left hand vs. right hand)
N2pc	−200 to 800	−200 to 0	35	2 (left target vs. right target)
Face experiment	−500 to 1500	−500 to 0	22	4 (4 face identities or 4 emotional expressions)
Orientation experiments	−500 to 1500	−500 to 0	16	16 (16 orientations or 16 locations)

Table 2:

Information about the decoding procedure for each case.

Decoding case	Measurement window (ms)	# of trials available for each class (mean±std)	# of crossfolds (J)	Min # of trials per crossfold (L)	# of channels
N170	110 to 150	68.41±6.95	5	13	28
MMN	125 to 225	185.77±13.94	10	18	28
P3b	300 to 600	51.05±5.08	3	10	28
N400	300 to 500	33.62±4.05	4	11	28
ERN	0 to 100	41.17±21.12	4	10	28
LRP	−100 to 0	168.86±15.84	10	16	28
N2pc	200 to 275	57.54±9.87	5	10	28
Facial expression	100 to 500 or 500 to 1000	160.77±3.95	10	16	59
Facial identity	100 to 500 or 500 to 1000	160.50±2.58	10	16	59
Orientation experiment 1	100 to 500 or 500 to 1000	38±0.37	3	12	27
Orientation experiment 2	100 to 500 or 500 to 1000	40.00±0.00	3	13	27