A Multi-Dimensional Functional Principal Components Analysis of EEG Data

Kyle Hasenstab[1,†], **Aaron Scheffler**[2,†], **Donatello Telesca**[2], **Catherine A. Sugar**[1,2,3], **Shafali Jeste**[3], **Charlotte DiStefano**[3], and **Damla Şentürk**[1,2,*]

[1]Department of Statistics, University of California, Los Angeles, CA 90095, U.S.A

[2]Department of Biostatistics, University of California, Los Angeles, CA 90095, U.S.A

[3]Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, CA 90095, U.S.A

Summary

The electroencephalography (EEG) data created in event-related potential (ERP) experiments have a complex high-dimensional structure. Each stimulus presentation, or trial, generates an ERP waveform which is an instance of functional data. The experiments are made up of sequences of multiple trials, resulting in longitudinal functional data and moreover, responses are recorded at multiple electrodes on the scalp, adding an electrode dimension. Traditional EEG analyses involve multiple simplifications of this structure to increase the signal-to-noise ratio, effectively collapsing the functional and longitudinal components by identifying key features of the ERPs and averaging them across trials. Motivated by an implicit learning paradigm used in autism research in which the functional, longitudinal and electrode components all have critical interpretations, we propose a multidimensional functional principal components analysis (MD-FPCA) technique which does not collapse any of the dimensions of the ERP data. The proposed decomposition is based on separation of the total variation into subject and subunit level variation which are further decomposed in a two-stage functional principal components analysis. The proposed methodology is shown to be useful for modeling longitudinal trends in the ERP functions, leading to novel insights into the learning patterns of children with Autism Spectrum Disorder (ASD) and their typically developing peers as well as comparisons between the two groups. Finite sample properties of MD-FPCA are further studied via extensive simulations.

Keywords

Electroencephalography; Event-related potentials data; Functional data analysis; Multilevel functional principal components

---

*dsenturk@ucla.edu.

†Equal contribution by the first and second authors.

1. Introduction

Electroencephalography (EEG) is a well-established noninvasive method for measuring spontaneous electrical activity across brain regions to identify neural function and cognitive states. Our motivating data is from a visual implicit learning study on young children with autism spectrum disorder (ASD) (Jeste et al., 2015). The experiment involved event-related potentials (ERP) in which EEG signals were time locked to the presentation of a continuous sequence of colored shapes (visual stimuli) recorded in age-matched 2 to 5 year old typically developing and ASD children (Figure 1(a)). The six colored shapes, grouped into three shape pairs, were presented in random order. Transitions within a shape pair were labeled 'expected' since they could be learned (shape ordering within a pair was fixed) and transitions between shape pairs were labeled 'unexpected' since they could not be predicted. The goal of the study was to characterize implicit learning, defined as the detection of regular patterns in one's environment without a conscious awareness or intention to learn, by contrasting brain response to expected and unexpected transitions.

The data created in typical ERP studies as the one described above are rich and multidimensional. Each stimulus, corresponding to the presentation of a single shape, referred to as a trial, results in an ERP function with paradigm-specific phasic components. The P3 peak and N1 dip phasic components typically studied in this paradigm and thought to be related to cognitive processes and early category recognition are given in Figure 1(b) (Jeste et al., 2015). Hence the experiment creates functional data (ERP curves) for each subject, collected longitudinally over trials (presentation of each shape) at multiple electrodes placed on the scalp. Due to the richness and multifaceted nature of the data along functional, longitudinal and electrode dimensions (repetitions over electrodes), typical practice involves multiple simplifications of the data before analysis. To increase the low signal-to-noise ratio (SNR) in raw ERP data, data are first collapsed in the longitudinal dimension, in which ERP functions observed over trials are averaged for each subject (Jeste et al., 2015; Gasser and Molinari, 1996). In addition, aside from a few works on functional mixed effects modeling for the analysis of ERP data (Bugli and Lambert, 2006; Davidson, 2009), the functional dimension is typically summarized by the amplitude (magnitude of the peak or dip) or latency (time when the peak or dip occurs) of the phasic components of the averaged ERP functions. Hence, once both functional and longitudinal dimensions have been collapsed into one-dimensional data summaries, a repeated measures ANOVA can be used for analysis.

In this paper we propose, for the first time in the literature, a longitudinal principal components decomposition for the EEG data, which we refer to as multi-dimensional functional principal components analysis (MD-FPCA). MD-FPCA embodies all three dimensions (functional, longitudinal and electrode) of the ERP data, preserving the full complexity without stringent assumptions or data reduction. In order to increase the SNR without collapsing the longitudinal dimension over trials, we adopt MAP-ERP, a meta-preprocessing step based on a moving average of the ERP functions over trials in a sliding window (Hasenstab et al., 2015). Capturing the longitudinal dimension is especially important in settings such as our motivating example, where patterns of learning correspond by definition to changes in ERP functions across trials. Previous studies in neuroscience and

biomedical engineering have acknowledged that ERP function morphology may change over the course of a task. However, most prior work has focused on controlling for longitudinal trends (Gasser et al., 1983; Turetsky et al., 1989) rather than modeling them; the few works on modeling longitudinal trends have been limited to parametric forms (Rossi et al., 2007; De Silva et al., 2012). We will build our proposed MD-FPCA on data produced by the novel meta-preprocessing step, MAP-ERP, capturing the continuum of longitudinal dynamics.

The literature on functional data analysis (FDA) (Ramsay and Silverman, 2005) has grown rapidly over the past two decades, with a considerable fraction of the work involving applications to longitudinal data (James et al., 2000; Müller, 2008; Sentürk and Müller, 2010). More recently, there has been interest in analyzing multiple trajectories, with dependencies among the repeatedly measured functional data (Crainiceanu et al., 2009; Di et al., 2009; Kundu et al., 2016; Morris and Carrol, 2009; Morris et al., 2003; Zupinnikov et al., 2011). Functional principal components decompositions for multilevel or longitudinal functional processes have been a major modeling theme in the FDA literature. Di et al. (2009) suggested decomposing sources of functional variation in an additive fashion via multilevel ANOVA, which we refer to as the ANOVA functional principal components decomposition (ANOVA-FPCA). Greven et al. (2010) proposed a decomposition based on a functional random intercept and slope to capture longitudinal variations, which we refer to as linear FPCA (LFPCA). Chen and Müller (2012) suggested a double decomposition (DFPCA) to capture potential nonlinear and nonparametric longitudinal trends within repeatedly observed functional data; parsimonious extensions of DFPCA have recently been proposed by Park and Staicu (2015) and Chen et al. (2016). While ANOVA-FPCA models longitudinal repetitions as repeated measurements without a particular time ordering, similar to an ANOVA, LFPCA models longitudinal trends linearly, and DFPCA does not assume a parametric form. For spatially correlated functional data, Delicado et al. (2010) summarize limited works in three categories, analysis of geostatistical functional data (Baladandayuthapani et al., 2008; Giraldo et al., 2010; Zhou et al., 2010; Staicu et al., 2010; Liu et al., 2016) (mostly involving distance-based parametric correlation structures), point processes with associated functional data and functional areal data.

Our proposed MD-FPCA combines the flexible DFPCA modeling of longitudinal trends, especially important for modeling learning trajectories in the motivating implicit learning experiment, with the decomposition of the total variation into subject and electrode level components as in ANOVA-FPCA, to embody all the dimensions of the multidimensional ERP data. MD-FPCA induces correlation between the electrode repetitions via random effects and utilizes multilevel random effects for extensions that involve data from multiple scalp regions. Following the initial ANOVA decomposition of the total variation, the proposed MD-FPCA involves a two-stage functional principal components decomposition of the subject and electrode level variations across functional and longitudinal time, leading to highly interpretable components contributing to the principal surfaces in a multiplicative fashion. Hence, even though multiple decompositions have been proposed for longitudinally observed or spatially correlated functional data in the literature, MD-FPCA is the first decomposition proposed for repeatedly measured longitudinal functional data which is tailored to model the specific features of the EEG data produced in ERP studies.

The remainder of the paper is organized as follows. Section 2 introduces the proposed MD-FPCA approach, compares it with other recently proposed functional principal components decompositions for longitudinally observed functional data, and outlines the extension of the methodology to analysis of data from multiple scalp regions. Section 3 provides insights gained from the implicit learning application, including comparisons of learning patterns in ASD and TD groups, summarized via the longitudinal trends in ERP functions across the experiment. We study the performance of the proposed decomposition in extensive simulation studies summarized in Section 4 and conclude with a brief discussion in Section 5.

2. Multi-Dimensional Functional Principal Components Analysis (MD-FPCA)

2.1 The Proposed MD-FPCA Decomposition

Denote by $X_{ij}(t|s)$ a multilevel square integrable random function observed across continuous functional time t, $t \in \mathcal{T}$, at longitudinal time s, $s \in \mathcal{S}$, for subunit j, $j = 1, \ldots, J$, and subject i, $i = 1, \ldots, n$. In applications to the EEG data, data collected over electrodes represent the subunits within subjects, functional time is the time scale of the ERP function and longitudinal time corresponds to trials. The notation, $X_{ij}(t|s)$, for the multilevel longitudinal functional process is used to stress the two-stage nature of the Karhunen-Loève decompositions in MD-FPCA, where the first-stage expansions are conditional on a particular longitudinal time s, and second-stage decompositions describe the variations along longitudinal time s. The function $X_{ij}(t|s)$ is decomposed using a multilevel random effects model at each longitudinal time s,

$$X_{ij}(t|s) = \mu(t,s) + \eta_j(t,s) + Z_i(t|s) + W_{ij}(t|s) + \varepsilon_{ij}(t|s), \quad (1)$$

where $\mu(t, s)$ and $\eta_j(t, s)$ are fixed functional effects that represent the overall mean function and subunit-specific shifts, respectively; $Z_i(t|s)$ and $W_{ij}(t|s)$ are the random subject- and subunit-specific deviations, respectively; and $\varepsilon_{ij}(t|s)$ is measurement error with mean zero and variance σ_s^2. Denote the total variation of $X_{ij}(t|s)$ at a fixed longitudinal time s by $\Sigma_T(t, t'|s) = \text{cov}\{X_{ij}(t|s), X_{ij}(t'|s)\}$ and let $\tilde{\sum}_T(t, t'|s) = cov\{X_{ij}(t|s), X_{ij}(t'|s)\} - \sigma_s^2 1_{\{t=t'\}}$ be the total variation without the measurement error with $1_{\{A\}}$ denoting the indicator function for event A. Assuming the subject and subunit-specific deviations, $Z_i(t|s)$ and $W_{ij}(t|s)$, are uncorrelated mean zero stochastic processes, (1) implies separation of the total variation $\tilde{\sum}_T(t, t'|s)$ at each longitudinal time s into subject level

$\sum^{(1)}(t, t'|s) = cov\{X_{ij}(t|s), X_{ij'}(t'|s)\} = \sum_k \lambda_k^{(1)}(s)\phi_k^{(1)}(t|s)\phi_k^{(1)}(t'|s)$ and electrode level

$\sum^{(2)}(t, t'|s) = \tilde{\sum}_T(t, t'|s) - \sum^{(1)}(t, t'|s) = \sum_p \lambda_p^{(2)}(s)\phi_p^{(2)}(t|s)\phi_p^{(2)}(t'|s)$ variation. Note that $\Sigma^{(1)}(t, t'|s)$ captures variation between electrodes within a subject and $\Sigma^{(2)}(t, t'|s)$ represents the remaining second level variance; we refer to these quantities as subject and electrode level variations, respectively, to build intuition that this separation is analogous to an ANOVA decomposition. In this formulation, $\phi_k^{(1)}(t|s)$ and $\phi_k^{(2)}(t|s)$ are the first and

second level eigenfunctions, describing modes of variation across functional time t at each longitudinal time s, and $\lambda_k^{(1)}(s)$ and $\lambda_p^{(2)}(s)$ are the first and second level eigenvalues.

Note that while the terminology 'level' is used to refer to the separation of the variation into the subject and electrode components, 'stage' is going to be used to refer to the two subsequent functional principal components decompositions applied to the deviation/variance defined at each level, first conditional on a particular longitudinal time s, and second describing variations along s. The first-stage Karhunen-Loève decompositions for $Z_i(t|s)=\sum_k \xi_{ik}(s)\phi_k^{(1)}(t|s)$ and $W_{ij}(t|s)=\sum_p \xi_{ijp}(s)\phi_k^{(2)}(t|s)$ are carried out at each s, yielding

$$X_{ij}(t|s)=\mu(t,s)+\eta_j(t,s)+\sum_{k=1}^{\infty}\xi_{ik}(s)\phi_k^{(1)}(t|s)+\sum_{p=1}^{\infty}\zeta_{ijp}(s)\phi_p^{(2)}(t|s)+\varepsilon_{ij}(t|s),$$

where $\xi_{ik}(s)$ and $\zeta_{ijp}(s)$ are the first and second level eigenscores, respectively, such that var $\{\xi_{ik}(s)\}=\lambda_k^{(1)}(s)$ and var $\{\zeta_{ijp}(s)\}=\lambda_p^{(2)}(s)$. In practice, the decompositions are truncated at only a small number of eigencomponents K and P (Web Appendix A). Note that the subunit repetitions within a subject, $W_{ij}(t|s)$ and hence $\zeta_{ijp}(s)$, are assumed to be independent across j, since the subunit dependency is modeled by the subject-specific random component $Z_i(t|s)$. Next, the second-stage Karhunen-Loève decompositions for the first and second level eigenscores, $\xi_{ik}(s)=\sum_{k'=1}^{\infty}\xi'_{ikk'}\psi_{kk'}^{(1)}(s), \zeta_{ijp}(s)=\sum_{p'=1}^{\infty}\zeta'_{ijpp'}\psi_{pp'}^{(2)}(s)$, yield

$$X_{ij}(t|s)=\mu(t,s)+\eta_j(t,s)+\sum_{k=1}^{\infty}\sum_{k'=1}^{\infty}\xi'_{ikk'}\psi_{kk'}^{(1)}(s)\phi_k^{(1)}(t|s)+\sum_{p=1}^{\infty}\sum_{p'=1}^{\infty}\zeta'_{ijpp'}\psi_{pp'}^{(2)}(s)\phi_p^{(2)}(t|s)+\varepsilon_{ij}(t|s).$$

$$(2)$$

In (2), $\xi'_{ikk'}$ and $\zeta'_{ijpp'}$ are the eigenscores, $\lambda_{kk'}^{(1)}=var(\xi'_{ikk'})$ and $\lambda_{pp'}^{(2)}=var(\zeta'_{ijpp'})$ are the eigenvalues and $\psi_{kk'}^{(1)}(s)$ and $\psi_{pp'}^{(2)}(s)$ are the eigenfunctions describing the modes of variation across longitudinal time of the first-stage eigenscores.

We propose two decomposition summaries for MD-FPCA that are important in identifying the contributions of different sources to the total variation in the analysis of the multilevel stochastic process $X_{ij}(t|s)$. While the first quantity, $\rho(s)=\{\sum_k \lambda_k^{(1)}(s)\}/\{\sum_k \lambda_k^{(1)}(S)+\sum_p \lambda_p^{(2)}(s)\}$, summarizes the proportion of variability explained by the subject level (first level) variation in the first-stage of MD-FPCA conditional on longitudinal time s, the second summary measure, $\rho=\{\sum_k \sum_{k'} \lambda_{kk'}^{(1)}\}/\{\sum_k \sum_{k'} \lambda_{kk'}^{(1)}+\sum_p \sum_{p'} \lambda_{pp'}^{(2)}\}$, captures the overall proportion of variability explained by the subject level variation in both stages of the MD-FPCA across

longitudinal time s. The two summaries can be viewed as extensions of the intra-cluster correlation of the linear mixed effects framework to the decomposition of multilevel longitudinally observed functional processes. The intraclass correlations can also be interpreted as the average correlation between two subunits from the same subject, conditional on and across longitudinal time s, respectively. In applications to ERP data, repetitions over electrodes are considered subunits; within-subject correlations between these repetitions provide insight into the similarity of the trends across electrodes.

2.2 Comparison to other Karhunen-Loève Decompositions

We briefly review three recently proposed principal components decompositions for functional processes, which are special cases of the proposed MD-FPCA, and highlight differences of MD-FPCA from a multilevel two-dimensional Karhunen-Loève decomposition. The three special cases are given without additive measurement error for simplicity. The ANOVA-FPCA of Di et al. (2009),

$$X_{ij}(t) = \mu(t) + \eta_j(t) + Z_i(t) + W_{ij}(t) = \mu(t,s) + \sum_{k=1}^{\infty} \xi_{ik}\phi_k^{(1)}(t) + \sum_{p=1}^{\infty} \zeta_{ijp}\phi_p^{(2)}(t), \tag{3}$$

is a special case of MD-FPCA, where the decomposition does not have a longitudinal component. The repeatedly observed functional process $X_{ij}(t)$ is decomposed into an overall mean $\mu(t)$, a visit-specific mean $\eta_j(t)$, a subject-specific deviation from the visit-specific mean $Z_i(t)$ and a subject-visit-specific deviation $W_{ij}(t)$. Similar to MD-FPCA, $Z_i(t)$ and $W_{ij}(t)$, which are called the first and second level deviations, respectively, are expanded using Karhunen-Loève decompositions with first and second level eigenscores ξ_{ik} and ζ_{ijp} and eigenfunctions $\phi_k^{(1)}(t)$ and $\phi_p^{(2)}(t)$, respectively. Note that repetitions of the functional process are modeled without a particular time ordering, similar to an ANOVA.

In contrast, the linear FPCA (LFPCA) of Greven et al. (2010) models longitudinal trends in a repeatedly observed functional process linearly. The double FPCA (DFPCA) decomposition of Chen and Müller (2012),

$$X_i(t|s) = \mu(t,s) + Z_i(t|s) = \mu(t|s) + \sum_{k=1}^{\infty} \xi_{ik}(s)\phi_k(t|s) = \mu(t,s) + \sum_{k=1}^{\infty}\sum_{p=1}^{\infty} \zeta_{ikp}\psi_{kp}(s)\phi_k(t|s), \tag{4}$$

does not assume a parametric form for the longitudinal time trend and thus can capture very flexible dynamics, similar to MD-FPCA. Note that the decomposition is for a longitudinally observed functional process that is not repeatedly observed, hence is a special case of MD-FPCA which considers the repeatedly observed longitudinal functional process $X_{ij}(t|s)$ with subject- and subunit-specific deviations. In (4), $X_i(t|s)$ is decomposed at a grid of longitudinal times s, yielding the mean function $\mu(t,s)$, eigenfunctions $\phi_k(t|s)$ and subject-specific random eigenscores $\xi_{ik}(s)$ at the first-stage. The eigenscores at each longitudinal time s are then decomposed further at the second-stage to yield eigenscores ζ_{ikp} and eigenfunctions $\psi_{kp}(s)$. Only decompositions LFPCA and DFPCA are suitable for situations in which

interest centers on detecting changes in the ERP functions across longitudinal time. Moreover, in applications to the implicit learning paradigm, even LFPCA may be quite restrictive, since it requires these trends to be linear. Hence the proposed MD-FPCA combines the more flexible DFPCA to model longitudinal trends with ANOVA-FPCA to model repetitions over electrodes (the electrode dimension), enabling us to compare the nature of the implicit learning processes of children with ASD and their typically developing peers.

In comparing MD-FPCA to a multilevel two-dimensional Karhunen-Loève decomposition, note that the proposed MD-FPCA in (2) implies principal surfaces across both the functional and longitudinal domains,

$$X_{ij}(t|s)=\mu(t,s)+\eta_j(t,s)+\sum_{k=1}^{\infty}\sum_{k'=1}^{\infty}\xi'_{ikk'}\varphi^{(1)}_{kk'}(t,s)+\sum_{p=1}^{\infty}\sum_{p'=1}^{\infty}\zeta'_{ijpp'}\varphi^{(2)}_{pp'}(t,s)+\varepsilon_{ij}(t|s),$$ (5)

such that $\varphi^{(1)}_{kk'}(t,s)=\phi^{(1)}_k(t|s)\psi^{(1)}_{kk'}(s)$ is a product of the first- and second-stage eigenfunctions, describing the variation conditional on longitudinal time s and along longitudinal time s, respectively; similarly $\varphi^{(2)}_{pp'}(t,s)=\psi^{(2)}_{pp'}(s)\phi^{(2)}_p(t|s)$. However, the principal surfaces $\varphi^{(1)}_{kk'}(t,s)$ and $\varphi^{(2)}_{pp'}(t,s)$ in (5) are not the eigenfunctions of the unconditional subject and subunit level covariance operators. In other words, the proposed MD-FPCA is distinct from a multilevel two-dimensional Karhunen-Loève decomposition with eigenscores θ_{ik} and ν_{ijp} and two-dimensional orthogonal eigenfunctions $\omega^{(1)}_k(t,s)$ and $\omega^{(2)}_p(t,s)$,

$$X_{ij}(t,s)=\mu(t,s)+\eta_j(t,s)+\sum_{k=1}^{\infty}\theta_{ik}\omega^{(1)}_k(t,s)+\sum_{p=1}^{\infty}\nu_{ijp}\omega^{(2)}_p(t,s)+\varepsilon_{ij}(t,s).$$

For the multilevel two-dimensional Karhunen-Loève decomposition, the unconditional total covariance of $X_{ij}(t, s)$ minus measurement error,

$\tilde{\sum}_T(t,s,t',s')=cov\{X_{ij}(t,s),X_{ij}(t',s')\}-\sigma_s^2 1_{\{t=t',s=s'\}}$, would be decomposed into subject $\sum^{(1)}(t,s,t',s')=cov\{X_{ij}(t,s),X_{ij'}(t',s')\}$ and subunit level $\sum^{(2)}(t,s,t',s')=\tilde{\sum}_T(t,s,t',s')-\sum^{(1)}(t,s,t',s')$ covariances. Then the covariation at both the subject and subunit levels would be expanded with two-dimensional functional principal component expansions,

$\sum^{(1)}(t,s,t',s')=\sum_k\tau^{(1)}_k\omega^{(1)}_k(t,s)\omega^{(1)}_k(t',s'),\sum^{(2)}(t,s,t',s')=\sum_p\tau^{(2)}_p\omega^{(2)}_p(t,s)\omega^{(2)}_p(t',s')$ with eigenvalues $\tau^{(1)}_k$ and $\tau^{(2)}_k$ and eigenfunctions $\omega^{(1)}_k(t,s)$ and $\omega^{(2)}_p(t,s)$.

A major advantage of the proposed MD-FPCA is that while the multilevel two-dimensional Karhunen-Loève decomposition would require decomposition and hence smoothing of

multiple four-dimensional covariance surfaces at the subject and subunit levels, the proposed MD-FPCA involves decompositions of only two-dimensional covariance surfaces due to the two-stage structure, leading to ease in implementation and savings in computational costs. Conditioning on longitudinal time s at the first-stage lowers the dimension of the covariance surface considered at both stages of the proposed algorithm. A second major advantage, as will be demonstrated in our application to the implicit learning study, is in interpretations of the decomposition components. The two-stage structure leads to additional decomposition components, such as the first- and second-stage eigenfunctions, which help with the interpretation of complex variation patterns in higher dimensions.

2.3 Extension to Data From Multiple Scalp Regions

In applications to the ERP data from the implicit learning paradigm, we consider four electrodes in the right frontal region of the scalp where maximal condition differentiation is detected. Motivated by the exchangeable correlation structure among electrodes within the same scalp region observed in the longitudinal functional ERP data, MD-FPCA models electrode repetitions by a random effect similar to an ANOVA approach. However MD-FPCA can be extended within the same ANOVA framework for analysis of data from multiple regions of interest on the scalp by an additional level of random effects at the scalp region level to account for differences between electrodes from different regions.

Denote by $X_{irj}(t|s)$ a multilevel square integrable random function observed across continuous functional time $t, t \in \mathcal{T}$, at longitudinal time $s, s \in \mathcal{S}$, for electrode $j, j = 1, \ldots, J$, within scalp region $r, r = 1, \ldots, R$, and subject $i, i = 1, \ldots, n$. Separating the total variation into variability at the subject, region and electrode levels at each longitudinal time s leads to

$$X_{irj}(t|s) = \mu(t,s) + \eta_r(t,s) + \alpha_{rj}(t,s) + Z_i(t|s) + W_{ir}(t|s) U_{irj}(t|s) + \varepsilon_{irj}(t|s), \quad (6)$$

where $\mu(t, s)$, $\eta_r(t, s)$ and $\alpha_{rj}(t, s)$ are the overall mean function, region- and electrode-specific shifts, respectively; $Z_i(t|s)$, $W_{ir}(t|s)$ and $U_{irj}(t|s)$ are the random subject-, region- and electrode-specific deviations, respectively; and $\varepsilon_{irj}(t|s)$ is measurement error with mean zero and variance σ_s^2. Denote the total variation of $X_{ij}(t|s)$ at a fixed longitudinal time s by $\sum_T(t,t'|s) = cov\{X_{irj}(t|s), X_{irj}(t'|s)\}$ and let $\tilde{\sum}_T(t,t'|s) = cov\{X_{irj}(t|s), X_{irj}(t'|s)\} - \sigma_s^2 1_{\{t=t'\}}$. Decomposition (6) implies separation of the total variation $\tilde{\sum}_T(t,t'|s)$ at each longitudinal time s into subject level

$$\sum{}^{(1)}(t,t'|s) = cov\{X_{irj}(t|s), X_{ir'j'}(t'|s)\} = \sum_k \lambda_k^{(1)}(s) \phi_k^{(1)}(t|s) \phi_k^{(1)}(t'|s)$$ region level

$$\sum{}^{(2)}(t,t'|s) = cov\{X_{irj}(t|s), X_{irj'}(t'|s)\} = \sum_p \lambda_p^{(2)}(s) \phi_p^{(2)}(t|s) \phi_p^{(2)}(t'|s)$$ and electrode level

$$\sum{}^{(3)}(t,t'|s) = \tilde{\sum}_T(t,t'|s) - \sum{}^{(1)}(t,t'|s) - \sum{}^{(2)}(t,t'|s) = \sum_v \lambda_v^{(3)}(s) \phi_v^{(3)}(t|s) \phi_v^{(3)}(t'|s)$$ variation. The second-stage Karhunen-Loève decompositions applied to the eigenscores $\xi_{ik}(s) = \sum_{k'=1}^{\infty} \xi'_{ikk'} \psi_{kk'}^{(1)}(s)$, $\zeta_{irp}(s) = \sum_{p'=1}^{\infty} \zeta'_{irpp'} \psi_{pp'}^{(2)}(s)$ and

$\beta_{irjv}(s) = \sum_{v'=1}^{\infty} \beta'_{irjvv'} \psi_{vv'}^{(3)}(s)$ from the first-stage Karhunen-Loève decompositions,

$Z_i(t|s) = \sum_k \xi_{ik}(s) \phi_k^{(1)}(t|s)$, $W_{ir}(t|s) = \sum_p \zeta_{irp}(s) \phi_p^{(2)}(t|s)$ and

$U_{irj}(t|s) = \sum_v \beta_{irjv}(s) \phi_v^{(3)}(t|s)$, yield the extended MD-FPCA decomposition

$$X_{ij}(t|s) = \mu(t,s) + \eta_r(t,s) + \alpha_{rj}(t,s) + \sum_{k=1}^{\infty} \sum_{k'=1}^{\infty} \xi'_{ikk'} \psi_{kk'}^{(1)}(s) \phi_k^{(1)}(t|s)$$
$$+ \sum_{p=1}^{\infty} \sum_{p'=1}^{\infty} \zeta'_{irpp'} \psi_{pp'}^{(2)}(s) \phi_p^{(2)}(t|s) + \sum_{v=1}^{\infty} \sum_{v'=1}^{\infty} \beta'_{irjvv'} \psi_{vv'}^{(3)}(s) \phi_v^{(3)}(t|s) + \varepsilon_{ij}(t|s).$$

Similar to MD-FPCA, $\xi'_{ikk'}$, $\zeta'_{irpp'}$ and $\beta'_{irjvv'}$ denote the second-stage eigenscores, $\lambda_k^{(1)}(s) = var\{\xi_{ik}(s)\}$, $\lambda_p^{(2)}(s) = var\{\zeta_{irp}(s)\}$ and $\lambda_v^{(3)}(s) = var\{\beta_{irjv}(s)\}$ are the first-stage eigenvalues, $\lambda_{kk'}^{(1)} = var(\xi'_{ikk'})$, $\lambda_{pp'}^{(2)} = var(\zeta'_{ijpp'})$, $\lambda_{vv'}^{(3)} = var(\beta'_{irjvv'})$ are the second-stage eigenvalues and $\psi_{kk'}^{(1)}(s)$, $\psi_{pp'}^{(2)}(s)$ and $\psi_{vv'}^{(3)}(s)$ are the second-stage eigenfunctions.

Note that the above outlined extension models correlations between electrodes within a scalp region and correlations between electrodes across different scalp regions as exchangeable. While in our experience, the first assumption is easier to verify for EEG data, the second assumption may be relaxed by modeling spatial correlations between electrodes across different scalp regions based on anatomical distances between the scalp regions. Distance-based correlations can be modeled by the addition of a spatial process to the expansion in (6) similar to the approach taken by (Staicu et al., 2010). Authors model multilevel functional data where spatial correlations are modeled at the lowest level of the hierarchy via a random spatial process. For EEG applications, the random spatial process would be added at the region level rather than the lowest level of the hierarchy which is the electrode level. This extension requires further developments and is identified as a topic for future research. Finally note that both the mean and the random effects structures can include additional terms such as diagnostic group (e.g. ASD, TD) or condition (e.g. expected, unexpected) for incorporation of different experiment specific factors into MD-FPCA. For illustration of the methodology, we model condition difference trajectories directly in applications to the implicit learning paradigm in the next section, where the diagnostic groups are modeled separately since we expect the ASD and TD groups to be different in mean trends as well as covariation.

3. Application to the Implicit Learning Study

3.1 Description of the Data Structure

In our motivating implicit learning study, EEG data are recorded on 37 ASD and 34 TD children for 120 trials per condition (expected and unexpected) for each subject at 128 electrodes. The EEG signals are sampled at 250Hz, producing 250 within-trial time points per waveform spanning 1000ms. The standard preprocessing steps of the data include artifact detection, bad channel replacement, referencing and baseline corrections. Next, the

meta-preprocessing of Hasenstab et al. (2015) is applied to the data, following the preprocessing steps, to increase the SNR to a level at which P3 peak locations can be identified without collapsing the entire longitudinal dimension (via averaging the ERP curves over all trials). The meta-preprocessing step averages ERP functions separately for each subject, electrode and condition, in a moving window of 30 trials to identify the P3 peak location in the averaged ERPs (see Web Appendix B for more details). We capture the shape of the entire P3 peak from these averaged ERPs by examining a 140ms window around the P3 peak identified in the meta-preprocessing step (i.e. functional time domain around the P3 peak location is $t \in [-70\text{ms}, 70\text{ms}]$ in 4ms increments). The length of the functional time domain of 140ms is determined by scientific practice and our own observation of the length of the entire P3 peak across trials. Note that the ERP curves in the described functional domain are already aligned across subjects, trials and electrodes, since we consider a symmetric window around each P3 peak. Since the interest lies in condition differentiation characterizing implicit learning, we focus on ERP difference functions obtained by subtracting the meta-preprocessed ERP corresponding to the unexpected condition from the expected condition. For illustration of the proposed methods, MD-FPCA is applied to meta-preprocessed difference ERPs from four electrodes in the right frontal region of the scalp, observed in trials $s \in [5, 60]$, where maximal condition differentiation is detected. Note that the longitudinal time domain starts at trial 5 to avoid boundary effects.

The proposed MD-FPCA algorithm is applied to the ASD and TD groups separately. Five subjects are removed as outliers prior to analysis. One of the removed subjects in the ASD group did not have available data until trial 20 and the remaining four subjects (two in each diagnostic group) had ERP difference functions more than 2 standard deviations away from their respective group means for most of the observations across both functional and longitudinal time domains. In addition, a single electrode is omitted from two subjects in the TD group due to highly nonhomogeneous trends compared to the other electrodes. The bandwidths of the mean functions and covariance smooths are selected using GCV and visual inspection of the one- and two-dimensional smooths. The selected bandwidths for the two-dimensional smoothing of the overall and subunit mean functions and total and within covariances in the first-stage decompositions are (30ms, 30 trials), (30ms, 30 trials), (15ms, 15ms), (15ms, 15ms) in the ASD group and (30ms, 30 trials), (30ms, 30 trials), (5ms, 5ms), (5ms, 5ms) in the TD group. The selected bandwidths for the mean functions and covariances in the second-stage decompositions are (15 trials, 15 trials), (5 trials, 5 trials) in the ASD group and (10 trials, 10 trials), (15 trials, 15 trials) in the TD group.

3.2 Data Analysis Results

Overall mean surface estimates $\mu(t, s)$ of ERP difference trajectories for both diagnostic groups are given in Figure 2 (a–b). The ASD mean surface displays a trend of positive concave condition differentiation across trials that is uniform across ERP time. The mean surface peaks around trial 35 where there is a slight differential increase around the P3 peak location (indexed by functional time $t = 0$). In contrast to the ASD mean surface, the TD group exhibits a trend of negative differentiation across trials with much smaller magnitude, including a prominent dip of negative differentiation around trial 25. Since the mean surfaces represent condition differentiation, the opposing mean trends between diagnostic

groups imply that children with ASD have higher EEG values in the expected condition while those in the TD group have higher values in the unexpected condition. This is consistent with our previous findings in Hasenstab et al. (2015) and Hasenstab et al. (2016). Another difference between diagnostic groups is in the timing of maximal condition differentiation (trial 35 for ASD and trial 25 in TD). This implies that while both diagnostic groups differentiate between the conditions, implying implicit learning, the children in the TD group are learning at higher speeds. Finally, while the entire P3 peak trajectory is increasing until trial 35 in the ASD group, there is a narrower window around the P3 peak that is minimized at the time of maximal condition differentiation in TD youth. The electrode specific means have similar patterns to the overall mean surfaces and are deferred to Web Figures 2 and 3.

Estimated leading subject level first-stage eigenfunctions $\phi_k^{(1)}(t|s)$ and second-stage eigenfunctions $\psi_{kk'}^{(1)}(s)$ are given in Figures 3 and 4 for the ASD and TD groups, respectively. Recall that while the eigenfunctions $\phi_k^{(1)}(t|s)$ display modes of variation in the functional dimension at a fixed longitudinal time s, the eigenfunctions $\phi_{kk'}^{(1)}(s)$ display modes of variation of the first-stage eigenscores in the longitudinal dimension. The products of these two quantities create subject level principal surfaces $\varphi_{kk'}^{(1)}(t, s)$ in (5) capturing the variation along both dimensions. Note that the model components $\phi_k^{(1)}(t|s)$ and $\phi_{kk'}^{(1)}(s)$ within themselves are quantities of interest and viewing them together provides an easily interpretable summary of the total variation conditional on and along longitudinal time.

In the ASD group, the uniform variation across ERP time in the leading component $\phi_1^{(1)}(t|s)$ (Figure 3 (a)), coupled with $\psi_{1k'}^{(1)}(s), k' = 1, 2$ (Figure 3 (b)) displaying variation along trials, indicate that majority of the variation is in the longitudinal/trial dimension at intermediate and later trials, $k' = 1$ (solid, corresponding to 27% of total variation explained), and at the boundary trials, $k' = 2$ (dashed, corresponding to 18.4% of variation explained). The resulting product principal surfaces and surface intervals $\mu(t, s) \pm \sqrt{\lambda_{1k'}^{(1)}}\varphi_{1k'}^{(1)}(t, s)$ for $k' = 1$ and 2, leading to the same interpretations, are given in Web Figures 6–7, Figure 2 (c) and Web Figure 5 (a), respectively. The second component $\phi_2^{(1)}(t|s)$ (Figure 3 (c)) of the subject level variation conditional on longitudinal time, captures a uniformly concave mode of variation in ERP time maximized at the P3 peak location $t = 0$. Estimated $\psi_{21}^{(1)}(s)$ (Figure 3 (d)) (solid, 1.8%), capturing modes of variation in the trial direction, indicates that the variation around the P3 peak is maximized at trial 35, the trial of maximum positive condition differentiation in the overall mean surface in the ASD group. There is additional variation in the boundary and intermediate trials in component $\psi_{22}^{(1)}(s)$ (Figure 3 (d)) (dashed, 1%).

As in the flat contour of the ASD leading eigenfunction, the leading component $\phi_1^{(1)}(t|s)$ (Figure 4 (a)) for the TD group is also fairly flat with majority of the variation still in the

longitudinal/trial dimension at the early and intermediate trials as captured by $\psi_{11}^{(1)}(s)$ (Figure 4 (b)) (solid, 36.9%) and later trials as captured by $\psi_{12}^{(1)}(s)$ (Figure 4 (b)) (dashed, 18.6%). (See Figure 2 (d) and Web Figure 5 (b) for surface intervals $\mu(t,s) \pm \sqrt{\lambda_{1k'}^{(1)}} \varphi_{1k'}^{(1)}(t,s) K' = 1$ and 2, respectively.) The estimated $\phi_2^{(1)}(t|s)$ (Figure 4 (c)) captures leftover variation around the peak location and in the boundaries of the functional/ERP time domain, with variation in the trial direction maximized at boundary trials as reflected in $\psi_{21}^{(1)}(s)$ (Figure 4 (d)) (solid, 1.3%) and intermediate trials as reflected in $\psi_{22}^{(1)}(s)$ (Figure 4 (d)) (dashed, 0.9%). In summary, while the majority of the variation is in the longitudinal/trial dimension for both ASD and TD groups, most of the variation is observed at intermediate and later trials in the ASD group and at early and intermediate trials in the TD group. For interpretations on the electrode level variation and subject-specific eigenscores, see Web Appendix C.

The number of principal components at both stages of the MD-FPCA fit are selected to explain at least 90% of the variation. The breakdown of the total variation explained by each component of the subject and electrode level decompositions are given in Table 1 for the ASD and TD groups. While two components are selected in the first-stage decompositions uniformly across levels and diagnostic groups, three to four components are needed in the second-stage decompositions. The variability explained by the subject level variation in both stages of the MD-FPCA across longitudinal time (ρ) is estimated to be 62% and 72%, respectively, in the ASD and TD groups. This indicates that the longitudinal functional trajectories observed at the four electrodes in the right frontal region within a subject behave similarly, as expected, and the majority of the total variation is explained at the subject level for both diagnostic groups. Nevertheless, the similarity among electrodes seems to be larger in the TD group, although the difference between diagnostic groups is not found to be significant (90% percentile CI's for ρ for ASD and TD groups are (0.53, 0.71) and (0.67, 0.82), respectively, based on a bootstrap procedure on the meta-preprocessing and MD-FPCA using 200 data sets sampled with replacement from subjects). Figure 5 (a–b) display the estimated proportion of variability $\rho(s)$ explained at the subject level in the first-stage of MD-FPCA for the ASD and TD groups, respectively. The average estimated $\rho(s)$ values again are higher in the TD group, where locations (along s) of maximum condition differentiation (trials 20 to 30) correspond to higher estimated $\rho(s)$ values in both groups. The similarity among electrodes within a subject seems to get stronger as the children start differentiating between the conditions, especially in the TD group.

To conclude, we briefly highlight the additional insights gained by utilizing all the dimensions of the data in the analysis without collapsing across longitudinal, functional or electrode repetitions. Note that the analysis of the longitudinal functional condition differentiation trajectories averaged over the electrodes, collapsing the electrode dimension, can be carried out by DFPCA of Chen and Müller (2012) which is a special case of MD-FPCA. Modeling the electrode dimension allowed us to study the electrode level variability, including comparisons to the variability at the subject level and the direction of the variation at the electrode level. More specifically, we learned that the majority of the variability was explained at the subject level in both groups (62% and 72% in ASD and TD, respectively).

In addition, the proposed index $\rho(s)$ provided a more detailed depiction of the proportion of total variability at the subject level as a function of longitudinal time.

Similarly, the analysis of the data collapsed over either the functional or longitudinal dimensions can be carried out by ANOVA-FPCA of Di et al. (2009), another special case of MD-FPCA. Not collapsing the longitudinal dimension (enabled by the meta-preprocessing and MD-FPCA) revealed critical information in the application to the implicit learning study. We were able to characterize the entire learning process as well as its speed in addition to comparisons between groups. There is exploratory evidence that the TD group starts differentiating between the two conditions of the experiment earlier (trial 25) than the ASD group (trial 35). Modeling the P3 waveform instead of just the P3 peak amplitude (see our previous work (Hasenstab et al., 2015, 2016)) allowed us to compare the variation in the longitudinal and functional dimensions. Variations/changes over longitudinal time (trials) explain more of the total variation in the data than variation in the functional dimension. We also observed that longitudinal changes in the P3 waveform morphology were different between the two groups. While the entire P3 peak trajectory increased until trial 35 in the ASD group, the TD group showed condition differentiation in a much narrower functional time window around the P3 peak at the time of maximal condition differentiation.

4. Simulation

We study the finite sample properties of MD-FPCA through extensive simulations outlined in Web Appendix D. MD-FPCA recovers the true first- and second-stage model components for both small ($N = 30$) and moderate ($N = 100$) sample sizes, under varying SNRs (between 1 and 100) and sparsity levels in the longitudinal time domain, with up to 40% of data missing at random longitudinal time points per subject. The median relative squared errors (RSEs) for all model components decrease with a denser design, increasing sample size and a higher SNR with the exception of the RSEs of the second-stage eigenfunctions which do not change with increasing SNR. This may be due to the fact that these quantities do not directly depend on data observed with measurement error.

5. Discussion

The proposed MD-FPCA has been presented under general settings without stringent assumptions on the separability of the longitudinal, functional and electrode covariances. Note that under the additional assumptions that modes of variability in the functional dimension stay the same across longitudinal times and electrode locations, or that modes of variability in the longitudinal dimension stay the same across functional times and electrode locations, more parsimonious versions of MD-FPCA can be derived using the marginal and product FPCA ideas of Park and Staicu (2015) and Chen et al. (2016). These extensions would lead to a common set of eigenfunctions in functional time across longitudinal times and electrode locations and/or a common set of eigenfunctions in longitudinal time across functional time and electrode locations. Finally, while we focused on modeling the P3 peak curves in the current application, MD-FPCA can be extended to model the entire ERP waveform in the functional dimension. This extension would require warping of the ERP waveforms after meta-preprocessing according to data features (e.g. N1, P3) while

simultaneously carrying out the multi-dimensional functional principal components decompositions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Baladandayuthapani V, Mallick B, Hong M, Lupton J, Turner N, Caroll R. Bayesian hierarchical spatially correlated functional data analysis with application to colon carcinogenesis. Biometrics. 2008; 64:64–73. [PubMed: 17608780]

Bugli C, Lambert P. Functional ANOVA with random functional effects: an application to event-related potentials modelling for electroencephalograms analysis. Statistics in Medicine. 2006; 25(21):3718–3739. [PubMed: 16372388]

Chen K, Müller HG. Modeling repeated longitudinal observations. Journal of the American Statistical Association. 2012; 107:1599–1609.

Chen K, Delicate P, Müller HG. Modeling function-valued stochastic processes, with applications to fertility dynamics. Journal of the Royal Statistical Society, Series B. 2016; doi: 10.1111/rssb.12160

Crainiceanu CM, Staicu AM, Di CZ. Generalized multilevel functional regression. Journal of The American Statistical Association. 2009; 104:1550–1561. [PubMed: 20625442]

Davidson DJ. Functional mixed-effects models for electrophysiological responses. Neurophysiology. 2009; 41(1):79–87.

De Silva AC, Sinclair NC, Liley DTJ. Limitations in the rapid extraction of evoked potentials using parametric modeling. Transactions on Biomedical Engineering. 2012; 59(5):1462–1471.

Delicado P, Giraldo R, Comas C, Mateu J. Statistics for spatial functional data: some recent contributions. Environmetrics. 2010; 21:224–239.

Di C, Crainiceanu CM, Caffo BS, Punjabi NM. Multilevel functional principal component analysis. The Annals of Applied Statistics. 2009; 3:458–488. [PubMed: 20221415]

Gasser T, Mocks J, Verleger R. SELAVCO: A method to deal with trial-to-trial variability of evoked potentials. Electroencephalography and Clinical Neurophysiology. 1983; 55(6):717–723. [PubMed: 6189703]

Gasser T, Molinari L. The analysis of the EEG. Statistical Methods in Medical Research. 1996; 5(1): 67–99. [PubMed: 8743079]

Giraldo R, Delicado P, Mateu J. Continuous time-varying kriging for spatial prediction of functional data: An environmental application. Journal of Agricultural, Biological, and Environmental Statistics. 2010; 15(1):66–82.

Greven S, Crainiceanu C, Caffo B, Reich D. Longitudinal functional principal component analysis. Electronic Journal of Statistics. 2010; 4:1022–1054. [PubMed: 21743825]

Hasenstab K, Sugar C, Telesca D, Jeste S, McEvoy K, entürk D. Identifying longitudinal trends within EEG experiments. Biometrics. 2015; 71(4):1090–1100. [PubMed: 26195327]

Hasenstab K, Sugar C, Telesca D, Jeste S, S̨entürk D. Robust functional clustering of longitudinal ERP trends. Biostatistics. 2016; 17(3):484–498. [PubMed: 26846337]

James G, Hastie TJ, Sugar CA. Functional modeling and classification of longitudinal data. Biometrika. 2000; 87:587–602.

Jeste SS, Kirkham N, S̨entürk D, Hasenstab K, Sugar C, Kupelian, et al. Electrophysiological evidence of heterogeneity in visual statistical learning in young children with ASD. Developmental Science. 2015; 18:90–105. [PubMed: 24824992]

Kundu MG, Harezlak J, Randolph TW. Longitudinal functional models with structured penalties. Statistical Modelling. 2016; 16(2):114–139. [PubMed: 28316508]

Liu C, Ray S, Hooker G. Functional principal components analysis of spatially correlated data. Statistics and Computing. 2016; doi: 10.1007/s11222-016-9708-4

Morris JS, Carrol RJ. Wavelet-based functional mixed models. Statistical Methodology. 2009; 68:179–199.

Morris JS, Vannucci M, Brown PJ, Carroll RJ. Wavelet-based nonparametric modeling of hierarchical functions in colon carcinogenesis. Journal of The American Statistical Association. 2003; 98:573–597.

Müller, HG., editor. Functional modeling of longitudinal data Longitudinal Data Analysis (Handbook of Modern Statistical Methods). Chapman and Hall/CRC; New York: 2008.

Park SY, Staicu AM. Longitudinal functional data analysis. Stat (Int Stat Inst). 2015; 41(1):212–226.

Ramsay, JO., Silverman, BW. Functional Data Analysis. New York: Springer; 2005.

Rossi L, Bianchi AM, Merzagora A, Gaggiani, et al. Single trial somatosensory evoked potential extraction with ARX filtering for a combined spinal cord intraoperative neuromonitoring technique. Biomedical Engineering Online. 2007; 6(2):1–12. [PubMed: 17199890]

entürk D, Müller HG. Functional varying coefficient models for longitudinal data. Journal of the American Statistical Association. 2010; 105:1256–1264.

Staicu A, Crainiceanu CM, Carroll RJ. Fast methods for spatially correlated multilevel functional data. Biostatistics. 2010; 11(2):177–194. [PubMed: 20089508]

Turetsky B, Raz J, Fein G. Estimation of trial-to-trial variation in evoked potential signals by smoothing across trials. Psychophysiology. 1989; 26(6):700–712. [PubMed: 2629018]

Zhou L, Huang JZ, Martinez JG, Maity A, Baladandayuthapani V, Carroll RJ. Reduced rank mixed effects models for spatially correlated hierarchical functional data Journal of the American Statistical Association. 2010; 105(489):390–400. [PubMed: 20396628]

Zupinnikov V, Caffo B, Yousem DM, Davatzikos C, Schwatz BS, Crainiceanu C. Multilevel functional principal component analysis for high-dimensional data. Journal of Computational and Graphical Statistics. 2011; 20:852–873. [PubMed: 25960627]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript
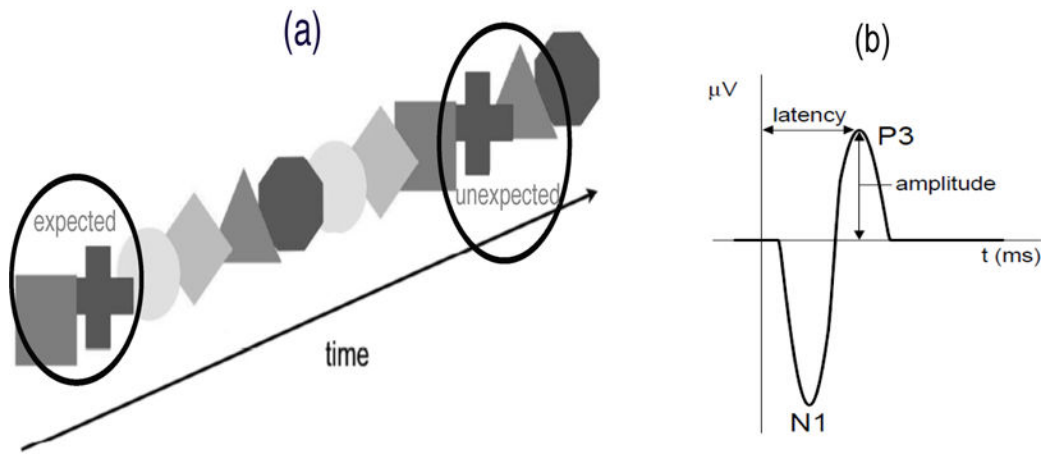
Figure 1.
(a) The sequence of shape pairs in the implicit learning study. Transitions within a shape pair are labelled 'expected' (square and cross are a shape pair, so that the cross always follows the square); transitions between shape pairs are labelled 'unexpected'. (b) The ERP waveform containing the P3 and N1 phasic components from the implicit learning study.
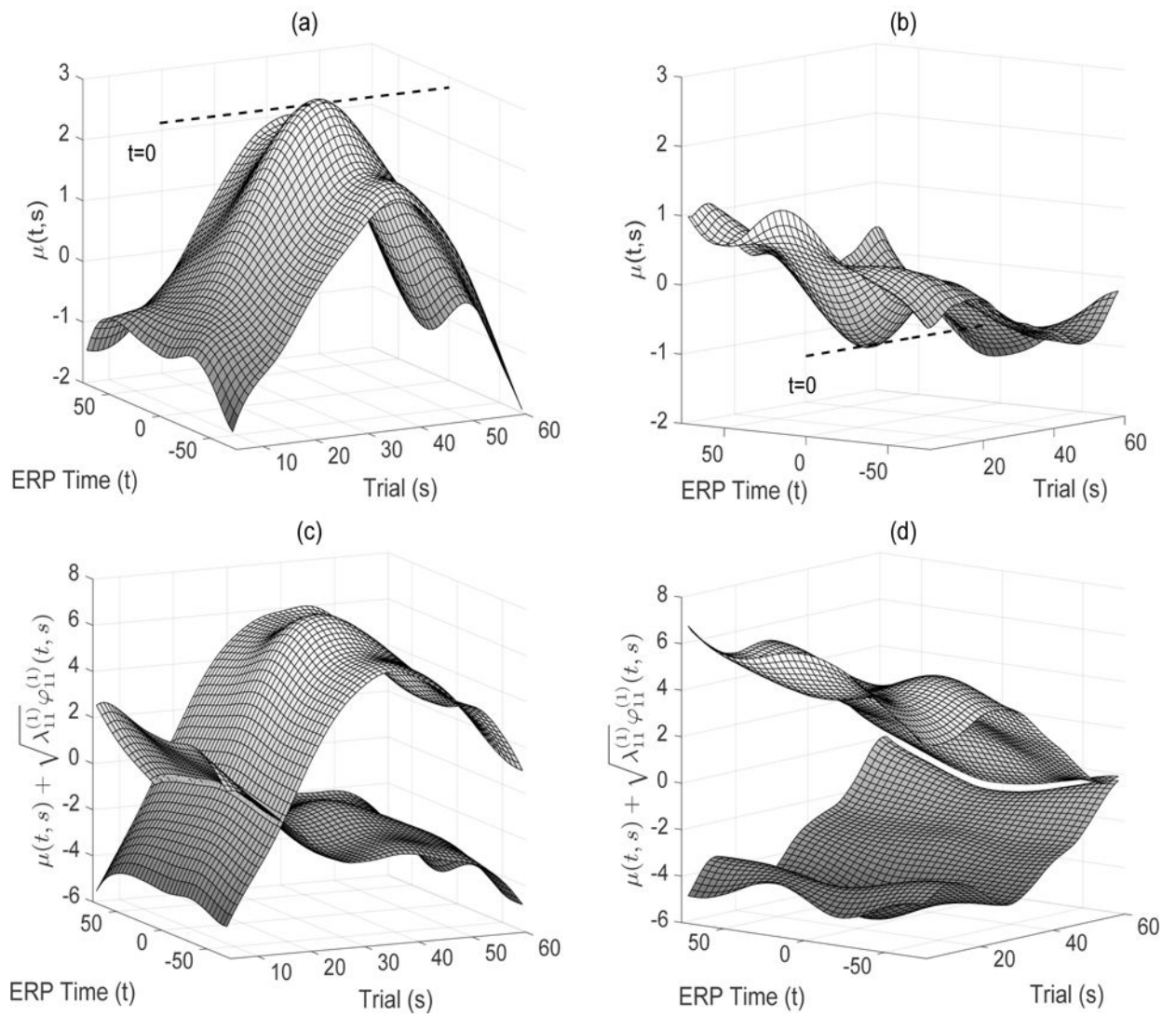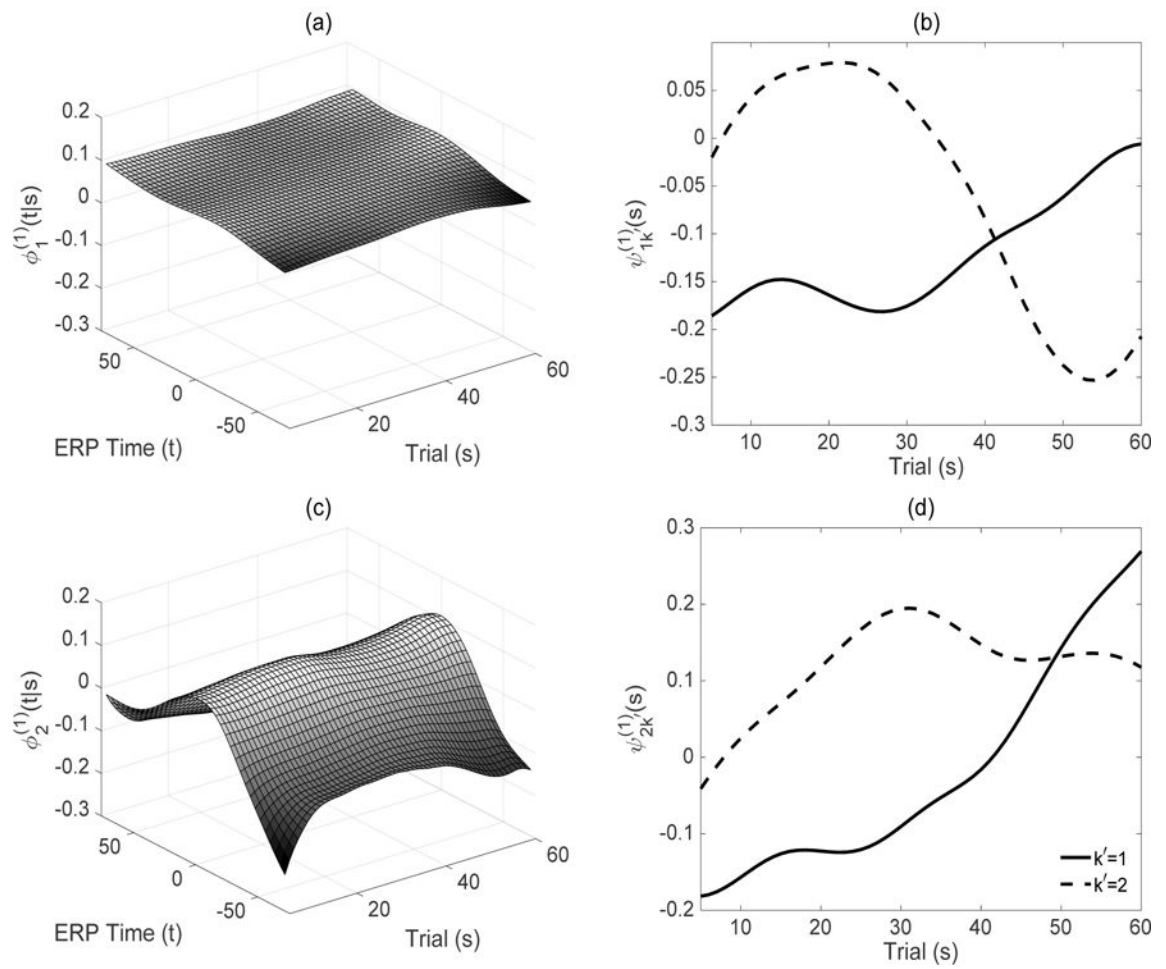
Figure 2.
(a)–(b) Estimated mean surfaces, $\mu(t, s)$, for the ASD and TD groups, respectively. (c)–(d)

Estimated surface intervals, $\mu(t, s) \pm \sqrt{\lambda_{11}^{(1)}} \varphi_{11}^{(1)}(t, s)$, for the ASD and TD groups, respectively.
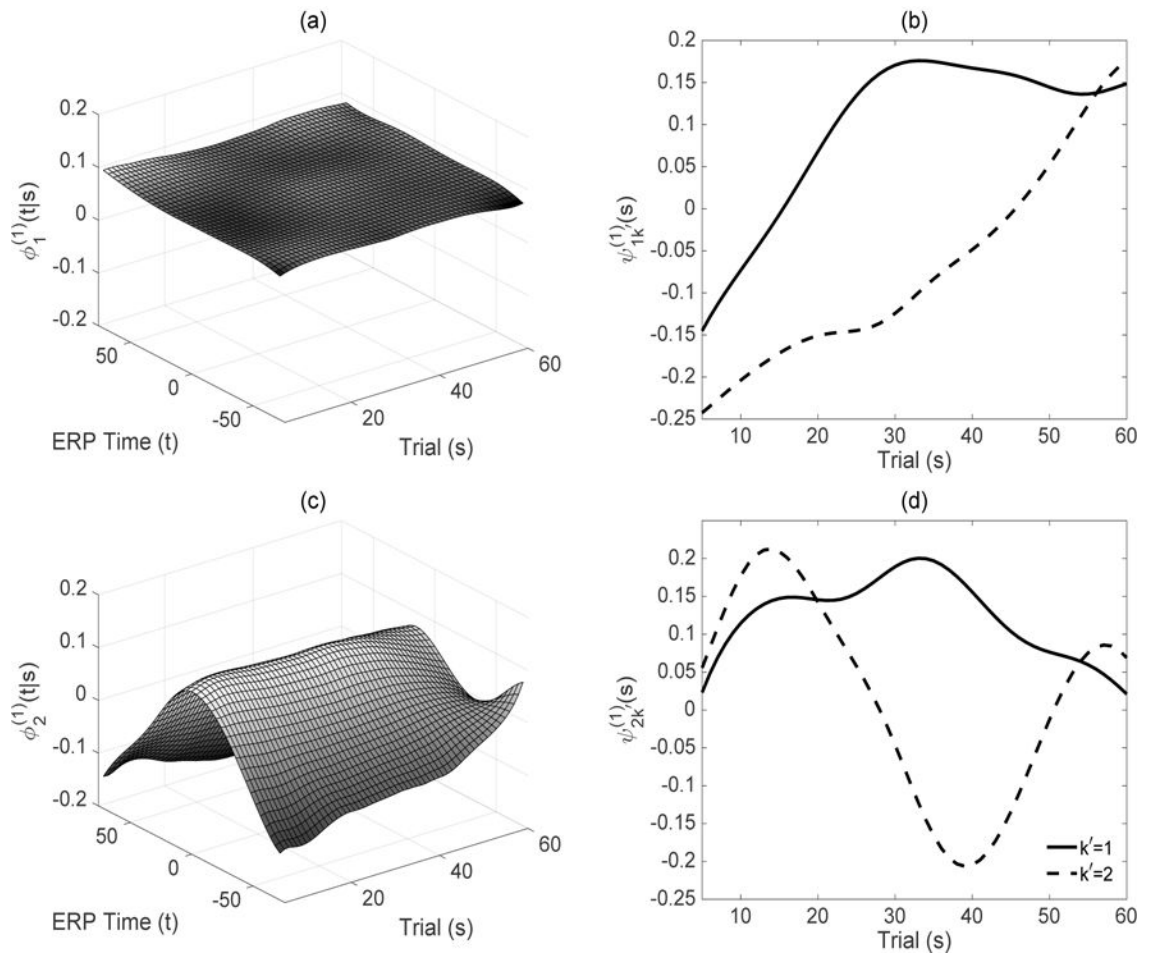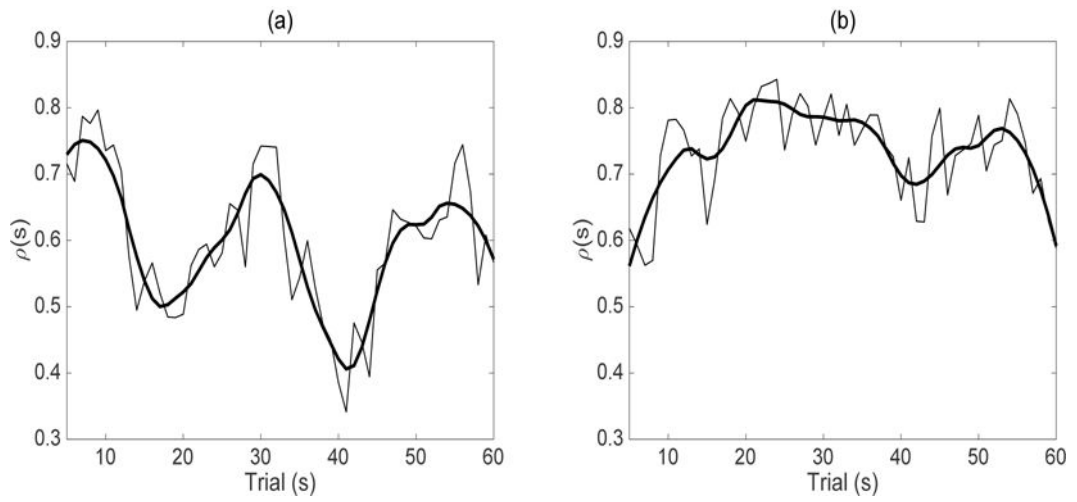
Figure 3.

(a), (c) Estimated leading subject level first-stage eigenfunctions, $\{\phi_k^{(1)}(t|s)\}, k=1, 2,$ respectively, for the ASD group. (b), (d) Estimated leading subject level second-stage eigenfunctions, $\{\psi_{kk'}^{(1)}(s)\}, k, k'=1, 2.$

Figure 4.

(a), (c) Estimated leading subject level first-stage eigenfunctions, $\{\phi_k^{(1)}(t|s)\}, k=1, 2,$ respectively, for the TD group. (b), (d) Estimated leading subject level second-secondstage eigenfunctions $\{\psi_{kk'}^{(1)}(s)\}, k, k'=1, 2.$

Figure 5.
(a)–(b) Estimated proportion of variability explained at the subject level in the first-stage of MD-FPCA for the ASD and TD groups, respectively. The thin black line corresponds to the raw proportion of variability explained while the thick line corresponds to its smooth.

Table 1

The number of principal components at both stages of the MD-FPCA fit selected to explain at least 90% of the variation.

| | Level 1 | | | | | Level 2 | | | |
| | ASD | | TD | | | ASD | | TD | |
	$k = 1$	$k = 2$	$k = 1$	$k = 2$		$p = 1$	$p = 2$	$p = 1$	$p = 2$
$k' = 1$	.270	.018	.369	.013	$p' = 1$	.217	.006	.118	.007
$k' = 2$	.184	.010	.186	.009	$p' = 2$	.085	.004	.074	.005
$k' = 3$	.133	.005	.141	.002	$p' = 3$	.062	.002	.055	.003
$k' = 4$	/	.004	/	/	$p' = 4$	/	/	.014	.001