tvst

Article

Developmental Effects on Pattern Visual Evoked Potentials Characterized by Principal Component Analysis

Carlyn Patterson Gentile[1], Nabin R. Joshi[2], Kenneth J. Ciuffreda[2], Kristy B. Arbogast[1,3], Christina Master[1,3], and Geoffrey K. Aguirre[3]

[1] Children's Hospital of Philadelphia, Philadelphia, PA, USA
[2] State University of New York College of Optometry, New York, NY, USA
[3] University of Pennsylvania, Philadelphia, PA, USA

Correspondence: Carlyn Patterson Gentile, Children's Hospital of Philadelphia, University of Pennsylvania, 3401 Civic Center Boulevard, Philadelphia, PA 19104, USA. e-mail: pattersonc@email.chop.edu

Received: December 10, 2020
Accepted: March 2, 2021
Published: April 1, 2021

Keywords: principal component analysis; visual evoked potentials; pediatric; development

Citation: Patterson Gentile C, Joshi NR, Ciuffreda KJ, Arbogast KB, Master C, Aguirre GK. Developmental effects on pattern visual evoked potentials characterized by principal component analysis. Trans Vis Sci Tech. 2021;10(4):1, https://doi.org/10.1167/tvst.10.4.1

Purpose: Peak amplitude and peak latency in the pattern reversal visual evoked potential (prVEP) vary with maturation. We considered that principal component analysis (PCA) may be used to describe age-related variation over the entire prVEP time course and provide a means of modeling and removing variation due to developmental age.

Methods: PrVEP was recorded from 155 healthy subjects ages 11 to 19 years at two time points. We created a model of the prVEP by identifying principal components (PCs) that explained >95% of the variance in a "training" dataset of 40 subjects. We examined the ability of the PCs to explain variance in an age- and sex-matched "validation" dataset ($n = 40$) and calculated the intrasubject reliability of the PC coefficients between the two time points. We explored the effect of subject age and sex upon the PC coefficients.

Results: Seven PCs accounted for 96.0% of the variability of the training dataset and 90.5% of the variability in the validation dataset with good within-subject reliability across time points ($R > 0.7$ for all PCs). The PCA model revealed narrowing and amplitude reduction of the P100 peak with maturation, and a broader and smaller P100 peak in male subjects compared to female subjects.

Conclusions: PCA is a generalizable, reliable, and unbiased method of analyzing prVEP. The PCA model revealed changes across maturation and biological sex not fully described by standard peak analysis.

Translational Relevance: We describe a novel application of PCA to characterize developmental changes of prVEP in youths that can be used to compare healthy and pathologic pediatric cohorts.

Introduction

Visual evoked potentials (VEPs) are a relatively inexpensive and easily implemented method of probing the cortical visual system. Pattern reversal VEPs (prVEPs), which are generated by viewing reversing checkerboard patterns, have the additional desirable property of relatively high intrasubject reliability in adults.[1] The prVEP can be helpful for clinical diagnosis and monitoring in demyelinating conditions, such as multiple sclerosis,[2] identifying sources of visual acuity loss,[3] and shows differences in a broad range of neurologic conditions, including concussion[4–7] and migraine.[8]

Despite the high potential of VEP as a diagnostic tool, prVEP interpretation faces unique challenges in the pediatric population. Multiple studies have found the prVEP varies during maturation, effects of which can be seen into early adulthood.[9–13] Age variation may confound studies that compare healthy and pathologic groups, either because of unmatched age differences in the populations, or more subtly, from a nonspecific effect of the disease upon

maturation. This is especially true for longitudinal prVEP measurements. Few studies have examined the stability of VEP across time in the pediatric population. One study showed high intrasubject variability of flash VEP signals collected in children 10 months apart on average.[14] This variability was higher than adult studies.[14] The ability to define and account for developmental changes is critical for the interpretation of VEPs in the pediatric population.

Standard prVEP analysis relies on peak amplitude and peak latency measurements of predictable positive and negative peaks that occur in the first 150 ms of the prVEP.[15] Prior studies of developmental effects upon the prVEP have found differences at multiple prVEP peaks as a function of age, although these results have been somewhat inconsistent. Some studies have found general decreases in peak latency,[9,10] whereas others have seen increases[13] or no change.[11] Peak amplitude tends to decrease with increasing age,[10–12] however, some studies found that this was not the case for all peaks studied.[10] Inconsistencies in these results may be due, in part, to the peaks analyzed, how peaks were defined, and differences in nonlocal temporal VEP waveform shape that are not fully captured by restricting analysis to predefined maxima and minima.

Principal component analysis (PCA) offers a potential solution to these challenges. PCA partitions the variability of a dataset into orthogonal, uncorrelated dimensions called principal components (PCs). As biologically relevant variation may be restricted to a subset of possible outcomes, PCA, like peak analysis, may be used to reduce the dimensionality of the VEP, but with the additional advantage of preserving informative variability of the global temporal signal.[16] This approach does not require specific reference to predefined, predictable VEP peaks. Further, once characterized by weights on a set of components, the effects of a continuous biological variable like developmental age can be modeled and accounted for in future analysis.

We used PCA to analyze prVEP time series data from a single active electrode collected from a large youth cohort. Using separate training and validation datasets, we demonstrate that a particular set of PCA components provide for generalizable analysis of prVEP data. Using measurements from two separate sessions in the validation dataset, we further show that the coefficients measured for these components have good reproducibility for a given subject across a short time span (weeks to months). Finally, we use this method to characterize developmental differences in our pediatric cohort over a span of years.

Methods

Subjects

Subjects between the ages of 11 and 19 years were recruited from a local Philadelphia-area school as part of a study on youth concussion conducted as part of the Children's Hospital of Philadelphia Minds Matter Concussion Program. Specifically, subjects were recruited through the sports teams of their local school to collect preseason VEP data on healthy athletes. Subjects with a history of concussion were at least 30 days from their most recent concussion and had resolution of their concussion symptoms. Consent was obtained from subjects and guardians, and all studies were approved by the Children's Hospital of Philadelphia Institutional Review Board and were in accordance with the Declaration of Helsinki. PrVEPs were collected on 155 subjects between February 2018 and February 2020. One hundred five subjects had at least 2 recorded sessions separated by 0.7 to 17 months. For 90 of the subjects, the sessions were separated by <6 months. Subject demographics, including age, sex, race, ethnicity, and medical history were recorded. Any medications the subjects were using was recorded. As part of the broader study, all subjects filled out a post-concussion symptom inventory (PCSI). Subjects were prescreened to ensure normal or corrected normal visual acuity under binocular and monocular viewing conditions using the Snellen's visual acuity chart at 10 feet (5 subjects wore contacts, and 9 subjects wore glasses). Those with corrective lenses wore them during testing.

Study Paradigm

Viewing Conditions

Subjects were seated in a well-lit, quiet room approximately 1 meter distant from a visual stimulus monitor. The room was not darkened. The luminance of the white walls adjacent to the testing apparatus was 53.0 cd/m^2 as measured using a spectrophotometer (PR670 Jadak Inc., North Syracuse, NY). Data were collected under binocular viewing conditions. Visual stimuli were presented on a monitor as part of a vision testing system (NOVA; Diopsys, Pine Brook, NJ). The monitor was a 17 inch liquid crystal display (LCD) screen with 1280 × 1024 resolution, and a 75 Hz refresh rate. The screen was factory calibrated with gamma correction. The measured luminance of the stimulus background (half-on primaries) was 103.0 cd/m^2.

Visual Stimulus and Paradigm

Visual stimuli were International Society for Clinical Electrophysiology of Vision (ISCEV) Standard large checks[15] presented in a wide field 85% contrast checkerboard with a pattern reversal rate of 2 reversals per second. Checks were 0.97 degrees visual angle. Subjects were instructed to fixate on a central red "x" for the duration of the stimulus presentation. Stimuli were presented for a continuous 20 seconds for a total of 40 reversals in a single block. The block was repeated five times in a single session. There was approximately 30 to 60 seconds in between blocks.

VEP Recordings

VEPs were recorded with a vision testing system (NOVA; Diopsys). Two reference electrodes were placed at Fz, and laterally to Fz at Fp1, with the active electrode placed at Oz following the 10 to 20 international electroencephalogram (EEG) placement criteria[17] based on the Diopsys system recommendations. Data were recorded at a sampling rate of 1024 samples/second. Raw voltage-by-time data were exported for subsequent analysis. Visual stimulus presentation was time-locked to VEP recording.

Data Analysis

Preregistration for this study can be found at the following link: https://github.com/pattersongentilelab/preregistrations. Data analysis was performed using custom software written in a coding program (Matlab; MathWorks, Natick, MA).

VEP Preprocessing

Notch filters were applied at 60 Hz and 120 Hz to remove powerline noise. Data were parsed into 500 ms intervals corresponding to each pattern reversal. Pattern reversals were excluded if there was a voltage change of >1 mV indicating a large nonphysiologic change in voltage (approximately 2% of reversals were discarded by this criterion). Signal-to-noise ratios (SNRs) were calculated for each block by taking the mean squared of each time point divided by the standard deviation squared, then taking the mean SNR across all time points. Blocks with an SNR of <0.03 (level determined by distribution of SNR values across all blocks) were excluded (3% of blocks). All VEP trials were adjusted to a baseline of 0 by subtracting the mean of the first 50 ms.

Visual Evoked Potentials

The mean prVEP for each subject was calculated across all reversals collected in a given session. Each session consisted of 5, 20-second blocks. Representations of prVEPs show the mean across-subject VEP with 95% confidence intervals obtained by bootstrap analysis across subjects with replacement.

PCA Model

Twenty female subjects and 20 male subjects were randomly selected from a pool of 90 subjects who had at least 2 recorded sessions, less than 6 months apart (Supplementary Fig. S1). These subjects constituted the training dataset. Mean prVEP across all trials, and across the two sessions from the training dataset, were used for PCA calculations. PCA was calculated using singular value decomposition (SVD):

$$M_{m\,x\,n} = U_{m\,x\,m}\ \Sigma_{m\,x\,n}\ V_{n\,x\,n}$$

where Matrix M is a m x n matrix, Matrix U is a is a m x m orthogonal matrix, Σ is a m x n rectangular diagonal matrix, and V is a n x n orthogonal matrix. For our dataset, matrix M consisted of $m = 40$ subjects and $n = 512$ voltage points spanning 500 ms sampled at a rate of 1024 Hz. The PCA was decentered (i.e. the mean was not subtracted from the responses for the analysis). The initial PCs that described >95% of the total variability were retained for subsequent analysis. We refer below to this set of PCs as the "PCA model."

PCA Model Validation for Generalizability and Retest Reliability

Forty age- and sex-matched subjects were selected from the remaining pool. This validation dataset was projected onto the PCA model. The distribution of coefficients for each PC were compared between the training and validation datasets using a two-sample Kolmogorov-Smirnov (KS) test.

We examined the reproducibility of coefficients within subject. Each session from the validation dataset was projected on to the PCA model. We calculated the Pearson correlation coefficient between PC coefficients from session 1 and session 2 across subjects. We then calculated the distance between session 1 and session 2 within the 7 PC dimensions for all subjects to determine if a subject could be correctly matched between their first and second sessions. Rank order for the smallest Euclidean distance was determined for the subject's own match between session 1 and session 2, compared to combinations with all other subjects.

Sex and Age Comparisons

Categorical variation across the subject group in sex, and continuous variation in age, were each modeled by an ANOVA for which the covariates were created from the seven PC coefficients from each subject. We simulated prVEP signals for a 10-year-old, 15-year-old, and 20-year-old subject (regressing out the effect of

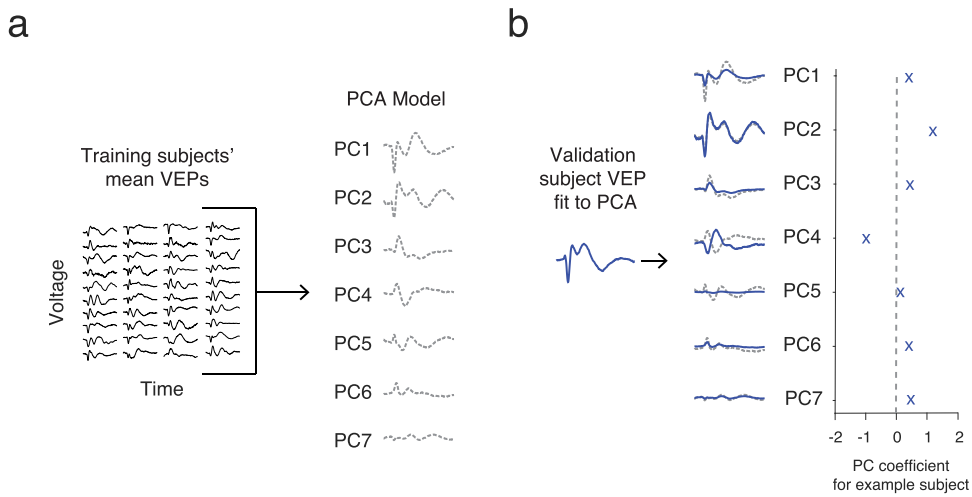

Figure 1. Schematic demonstrating generation and use of the PCA model. (**a**) Forty randomly selected subjects were used to generate the PCA model. Seven PCs (*gray dotted waveforms*) capture 96.0% of the variability across these 40 subjects. (**b**) An example subject from the validation dataset is fit to the PCA model. The subject was not used to generate the original model. The example subject's mean VEP waveform (*blue*) is fit to the PCA model (the PCs multiplied by the coefficients for the example subject are shown in *blue lines* overlapping the original PCs in *dotted gray*) and a PC coefficient is generated for the subject for each of the 7 PCs (*blue* "x").

sex) by obtaining the PC coefficients of the 7 PCs for this age group and fitting to the PCA model.

Peak Analysis

The N75 and P100 peaks were defined by identifying the local minimum in the 60 to 90 ms range, and the local maximum in the 90 to 130 ms range, respectively. This was done from the mean VEP waveform for each session for the comparison across sessions, and from the mean VEP waveform across sessions for the age correlation. Pearson Correlation Coefficients between session 1 to session 2, and correlation with age were calculated for these three variables.

Results

Seven Components are Sufficient to Capture prVEP Variability

The prVEP (averaged across sessions) from 40 randomly selected subjects (20 female subjects and 20 male subjects) was used to derive the PCA model (i.e. training dataset). Specifically, PCA was used to capture the variability in the VEP waveform across these 40 subjects (Fig. 1a). The mean prVEP for the training dataset is shown in Figure 2a. Seven PCs (Fig. 2b) accounted for 96.0% of the variability in this sample (Fig. 2c). We refer to these seven PCs generated by the training dataset as the PCA model. As the PCA was decentered, the first PC approximates the mean prVEP. PCs two through seven do not clearly map to particular peaks or points of time within the prVEP signal. However, certain peaks are represented in particular ways within these PCs. For instance, PC 2 has the effect of broadening the P100 peak. PC 3 has the effect of shortening the peak latency of P100 and increasing the P100 amplitude. PC 4 primarily broadens the P100 peak. These complex changes in peak shape would not be fully captured by traditional peak analysis techniques.

The PCA Model Yields Similar PC Coefficients in an Independent Validation Dataset

Forty age- and sex-matched subjects were selected to validation the generalizability of the PCA model. There were no significant differences in patient demographics, including race/ethnicity and medical history between the training and validation subjects (Table 1). There was a significant difference of the number of subjects on stimulant medications for attention deficit hyperactivity disorder (ADHD), with more subjects on these medications in the validation compared to the training group (Table 1).

Measurements from the validation subjects were projected onto the PCA model generated from the training subjects. An example of a validation subject having their VEP projected upon the PCA model with resultant coefficients for the 7 PCs is shown in Figure 1b. If the PCA model is generalizable, it should be able to capture variance in the validation dataset similar to its performance in the training dataset. We

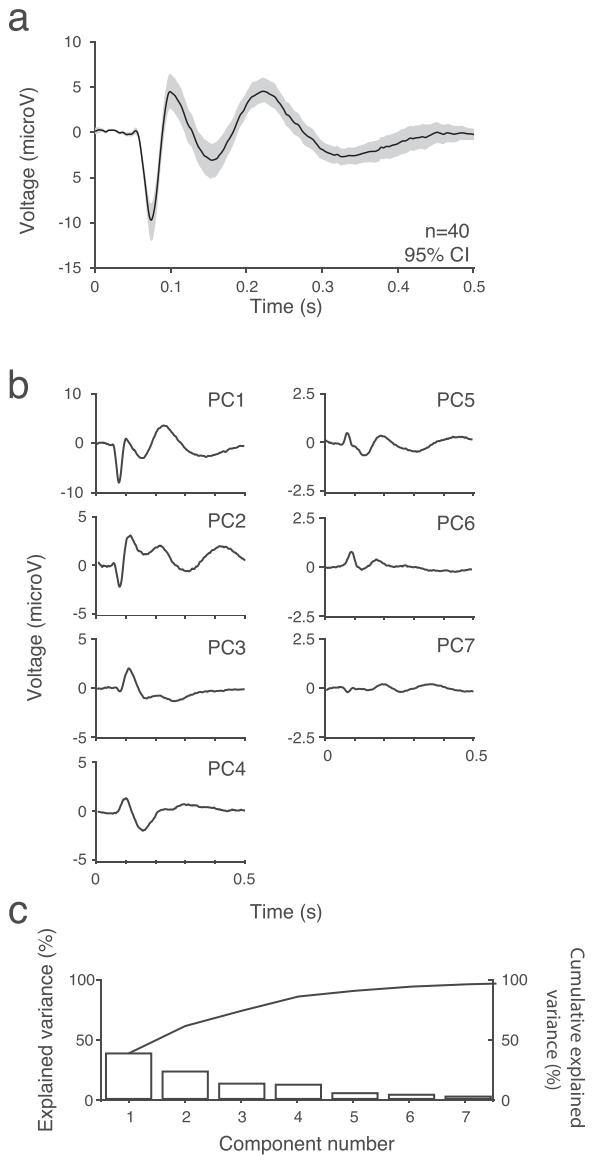

Figure 2. PCA model. PCA was generated with 40 randomly selected subjects (20 male subjects and 20 female subjects). (**a**) Mean VEP across all 40 training subjects; gray represents 95% confidence interval (CI) by bootstrap analysis. (**b**) PCs 1 to 7, which account for 96.0% of the variability in the data. (**c**) Scree plot showing percent explained variance of PCs 1 to 7.

found that the 7 PCs of the model accounted for 90.5% of the variance in the validation dataset.

As a second test of generalizability, we considered that the distributions of PC coefficients derived using the PCA model in the validation dataset should be similar to the distributions seen in the training dataset. We examined these distributions for each PC and found substantial overlap of the PC coefficients derived from the training (black) and validation (blue) datasets (Fig. 3a). Comparison of the mean and distribution of PC coefficients between the datasets revealed no significant differences (see Fig. 3a, $P > 0.3$ for all PCs; KS test). There was also no difference between training and validation datasets in the seven PC multidimensional space (ANOVA $F_{(1,78)} = 0.03$, $P = 0.86$).

Coefficients Derived Using PCA Model Have Good Within Subject Retest Reliability

The PCA model was created using the validation dataset. Specifically, the mean age and mean prVEP was calculated from two separate sessions for each subject. We next asked if the PC coefficients obtained using this model were similar within subject for the two testing sessions. For each PC, we examined the correlation of coefficients across subjects for data from the first and second testing session (Fig. 3b). The PCA model demonstrated high retest reliability across sessions (Pearson correlation coefficient, $R \geq 0.75$ for all PCs). We examined the correlation between session 1 and session 2 coefficients across the first 40 components of the PCA model (i.e. extending beyond the first 7 PCs that we retained; Supplementary Fig. S2a). The intrasubject reliability of the coefficients of the prVEP declines across the components, indicating that the higher dimensions of the model likely reflect nonreproducible noise in the measurement. Consequently, expressing prVEPs by projection onto the first seven PCs of the PCA model constitutes a noise-reduction technique.

The similarity of any two prVEPs may be expressed by projecting the waveforms upon the PCA model and measuring the Euclidean distance between the two sets of seven PC coefficients. The distance between session 1 and session 2 was calculated within the 7 PC dimensions for all subjects to determine if a subject could be correctly matched between their first and second sessions. This would indicate very high intrasubject compared to intersubject reliability and indicate stability of VEP signal for a single subject across sessions. Rank order for the subject's own match between session 1 and session 2, compared to combinations with all other subjects is shown (Fig. 3b, lower right). In 70% of the subjects, the best match to their session 1 prVEP was their session 2 prVEP. In over 85% of the subjects, the session 2 prVEP was in the top 5 ranking of prVEPs that were similar to session 1, compared to 40 possible subject session 1 and session 2 combinations. We examined the ability to match the prVEP from a subject across sessions as a function of the number of PCs used to describe the data. Accuracy plateaued with inclusion of approximately seven PCs further supporting the use of this dimensionality to describe the prVEP data (Supplementary Fig. S2b). We also looked at correlation between session 1 and session

Table 1. Subject Demographics for the Training and Validation Subject Groups Used to Generate and Validate the PCA Model, Respectively

	Training Subjects	Validation Subjects	
No. of subjects	40	40	
Male	20	20	
Female	20	20	
Median age at first VEP	15.3 y	15.3 y	*t*-test
[age range]	[11.2–19.1]	[11.8–18.1]	$P = 0.94$
Race/ethnicity			*t*-test
Non-Hispanic White	28	31	$P = 0.88$
Non-Hispanic Black	3	5	
Hispanic	3	0	
Non-Hispanic Asian	1	1	
Non-Hispanic mixed race	1	2	
Non-Hispanic other	0	0	
Unknown	4	1	
Medical history			z-test
Concussion	9	11	$P = 0.61$
Migraine	0	3	$P = 0.08$
Chronic headaches	0	1	$P = 0.32$
ADHD	1	5	$P = 0.09$
Motion sickness	3	2	$P = 0.64$
Sleep problem	0	2	$P = 0.15$
Anxiety	2	4	$P = 0.40$
Depression	1	4	$P = 0.17$
Other psychiatric disorder	0	2	$P = 0.15$
POTS	0	1	$P = 0.32$
Neuro-active medications			z-test
Stimulant (ADHD medication)	0	5	$P = 0.02$
SSRI or SNRI	1	3	$p = 0.31$

No subjects reported a medical history of dyslexia, bipolar, drug/alcohol use disorder, autism, epilepsy, tic disorder, or amplified musculoskeletal pain syndrome (AMPS).

POTS, postural orthostatic tachycardia syndrome; SNRI, serotonin-norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor.

2 for standard peak analysis metrics, including the N75 and P100 peak latencies and the N75 to P100 peak-to-peak amplitudes. N75 and P100 peak latency showed good correlation ($R = 0.45$, $P = 0.004$; $R = 0.64$, $P < 0.001$, respectively), and the N75 to P100 peak-to-peak amplitude showed excellent correlation ($R = 0.92$, $P < 0.001$) across sessions (Supplementary Fig. S3a). Given the high test-retest reliability of the measurements, the mean prVEP was calculated across two sessions for each subject (when available) for the remaining analyses.

Characteristics of the Full Cohort

The remaining analyses were performed on the full cohort of 155 subjects. Demographics are shown for the full cohort (Table 2). Two subjects had a medical history of a "lazy eye," two subjects had a history of wearing an eye patch, and one subject had a history of strabismus. Exclusion of these five subjects did not significantly change the outcome of the results. Seventeen (11%) of the subjects reported taking neuroactive medications, including stimulants for ADHD (9), selective serotonin reuptake inhibitors or serotonin and norepinephrine reuptake inhibitors (7), beta blockers (1), and antipsychotics (1). The other 34 subjects who reported medication use included as needed asthma inhalers, birth control, and allergy medications. Overall PCSI scores were low (median = 3.5 out of a possible total score of 126), and there was not a significant difference in the distribution of PCSI scores for subjects with versus

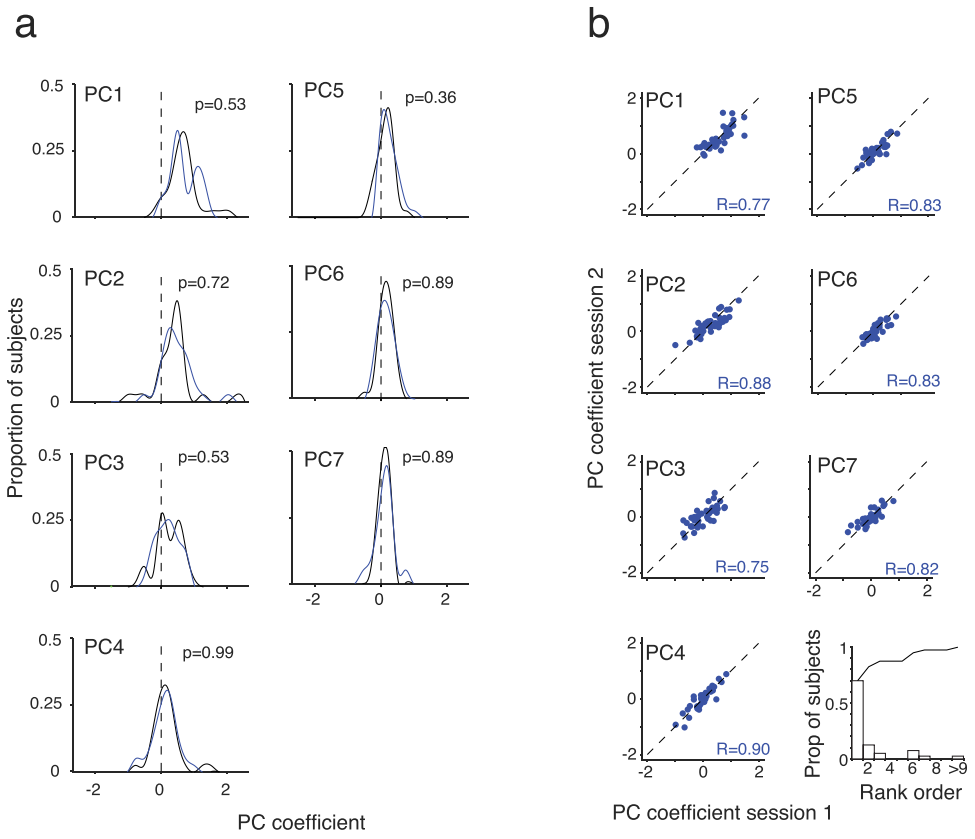

Figure 3. Generalizability and test-retest reliability of the PCA model. (**a**) PC coefficient comparison of 40 training subjects (*black*) and 40 validation subjects (*blue*) showed no significant differences in the mean or distribution of individual PCs (KS test $P > 0.3$ for all PCs) or multidimensional space across all PCs (ANOVA $F_{(1,78)} = 0.03$, $P = 0.87$). (**b**) Scatter plots present the coefficients for each PC derived from session 1 and session 2 for each subject. There was high correlation for these measurements between sessions across subjects (all $R > 0.7$). The bottom right panel shows the proportion of subjects with the rank score for a subject's Euclidean difference between session 1 and session 2 compared to session 2 of all validation subjects. *Black line* shows the cumulative proportion of subjects.

without a remote history of a concussion ($P = 0.53$, KS test).

There Was An Effect of Sex on prVEP in the PCA Model

A linear regression model was generated to examine if the seven PCs varied based on the sex of a subject. We did not observe a significant difference between male subjects and female subjects across all 155 subjects ($F_{(7,147)} = 1.69$, $P = 0.12$). However, when subjects on neuro-active medication were excluded (8 female subjects and 9 male subjects), there was a significant effect of sex on the PCA model ($F_{(7,130)} = 2.31$, $P = 0.03$). Coefficients that showed a significant difference between male and female subjects were PC1 ($t_{(1,130)} = -2.13$, $P = 0.03$) and PC6 ($t_{(1,130)} = -2.48$, $P = 0.01$). These differences capture an increased overall amplitude (PC1) and a narrower P100 peak (PC6) for female subjects compared to male subjects. Mean male and female prVEPs and the distribution of the seven PC coefficients with and without inclusion of subjects on neuro-active medications are shown (Supplementary Fig. S4).

The PCA Model Captures Variability Throughout Maturation in our Pediatric Cohort

A linear regression model was generated to examine the effect of the seven PCs upon the age of a subject. There was a significant, omnibus effect of age upon the seven PC coefficients ($F_{(7,147)} = 4.37$, $P = 0.0002$). The PC coefficients that showed a significant correlation with age were PC2 ($t_{(1,147)} = -4.06$, $P = 8e^{-5}$) and PC3 ($t_{(1,147)} = -3.21$, $P = 0.002$). PC coefficients for the seven PCs are shown as a function age (Supplementary Fig. S5). The prVEP from any one

Table 2. Subject Demographics of the Full Pediatric Cohort

Full Cohort	
No. of subjects	155
Male	68 (43.8%)
Female	87 (56.2%)
Median age at first VEP	15.2 y
[age range]	[11.2–19.1]
Race/ethnicity	
Non-Hispanic White	119 (76.7%)
Non-Hispanic Black	12 (7.7%)
Hispanic	7 (4.5%)
Non-Hispanic Asian	4 (2.6%)
Non-Hispanic mixed race	5 (3.2%)
Non-Hispanic other	2 (1.3%)
Unknown	6 (3.9%)
Medical history	
Concussion	42 (27.1%)
Migraine	5 (3.2%)
Chronic headaches	1 (0.6%)
Dyslexia	1 (0.6%)
ADHD	14 (9.0%)
Motion sickness	10 (6.5%)
Sleep problem	3 (1.9%)
Anxiety	10 (6.5%)
Depression	10 (6.5%)
Other psychiatric	2 (1.2%)
POTS	1 (0.6%)
Medications	
None	81
Any medication	51
Neuro-active medications	17
Stimulant (ADHD medication)	9
SSRI or SNRI	7
Beta blocker	1
Antipsychotic	1
Not reported	23

No subjects reported a medical history of bipolar, drug/alcohol use disorder, autism, epilepsy, tic disorder, or amplified musculoskeletal pain syndrome (AMPS).

POTS, postural orthostatic tachycardia syndrome; SNRI, serotonin-norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor.

subject can be described as a point in the 7-dimensional PC space. We plotted the prVEPs for our subjects along the subset of two dimensions that showed a significant change with maturation (Fig. 4a). A vector within the PC space describes the effect of age upon the coefficients. The PCA model may be used to synthesize representative prVEP waveforms that lie along this vector. The result is a synthetic waveform that expresses the central tendency of the prVEP corresponding to subjects of different ages (Fig. 4b). These waveforms capture multiple, nonlocal changes as a function of increasing subject age, including a progressively smaller P100 amplitude, narrowing of the P100 peak, and an increasing N135 amplitude. In agreement with these findings, peak analysis metrics showed a negative correlation of the N75 to P100 peak-to-peak amplitude with age ($R = -30$, $P = 0.002$), but there was no significant correlation between age and peak latency (see Supplementary Fig. S3b). Removal of subjects on neuro-active medications did not have a significant impact on the outcome of the PCA model across age.

We considered if a non-neural developmental change, such as head circumference, could account for age effects in the prVEP. We examined the correlation between age and head circumference in the 124 subjects with this measure and found a nonsignificant correlation ($R = -0.04$, $P = 0.66$). There was no relationship between individual differences in head circumference and the set of seven PC coefficients across subjects ($F_{(7,116)} = 0.70$, $P = 0.67$). It seems therefore that variation in head circumference in this cohort has minimal effect upon our prVEP measures.

Discussion

We found that a PCA model using 7 PCs accounted for over 90% of the intersubject variability of prVEP across 40 pediatric subjects for both a training and validation dataset. We demonstrate that this PCA model is a generalizable and reliable means of analyzing prVEP in a large pediatric sample. PCA offers a method of adjusting for differences across maturation to remove this confounding variability, which is important for the interpretation of VEP measurements in developing youth. For these reasons, PCA offers a promising complement to standard peak analysis for interpretation of VEP.

Comparison to Previous Studies

Prior studies that have relied on peak analysis have generally found decreasing peak amplitude and decreasing peak latency with maturation, although some conflicting results have been reported. Snyder and colleagues (1981) found that the amplitudes of the P50, N64, P100, and N150 peaks decreased with increasing age during adolescence.[12] In related work, Shearer and Dustman (1980) found that latencies of

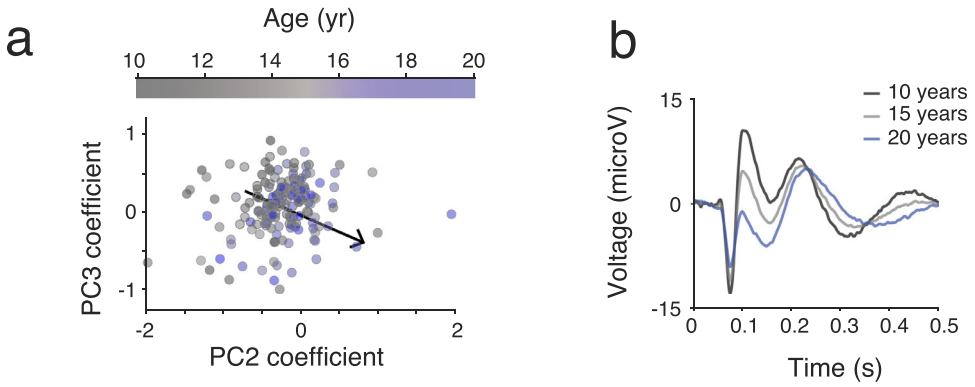

Figure 4. PC coefficients change systematically with age. (**a**) The VEP from each subject may be described by their PC2 and PC3 coefficients. These dimensions of the PCA model had a significant association with subject age. Plot points are colored from *black* to *blue* indicating increasing subject age, and the *red arrow* indicates the vector direction of the effect of subject age in this space. (**b**) Simulated VEPs for a 10-year-old, 15-year-old, and 20-year-old subject.

these peaks gradually increased over time.[13] Wright and colleagues (1985) found overall amplitudes were higher in their youngest age group (10–19 years) compared to the older age groups, and found no differences in peak latencies across age.[11] One limitation of this study was that prVEPs were averaged across subjects in 10-year age groups, which would have failed to capture differences of maturation during adolescence. Emmerson-Hanover and colleagues (1994), who examined a large cohort of 406 subjects ages 6 to 80 years, found somewhat different results. They observed that the P50 to N70 amplitude increased until about age 13 years and then decreased, whereas the N70 to P100 amplitude decreased with maturation.[10] They also found the P50, N70, and P100 peak latencies decreased during maturation.[10] Brecelj and colleagues (2002) also found in children 7 to 18 years old that P100 peak latency decreased with maturation.[18] Allison and colleagues (1983) similarly found the VEP P100 peak latency decreased between 4 and 19 years of age, although the peak latency of other peaks and peak amplitude were not reported.[9] We also observed that prVEP amplitude decreases with maturation using both the PCA model and peak analysis (see Fig. 4 and Supplementary Fig. S3b). Our PCA model also revealed a narrowing of the P100 peak with maturation that cannot be easily captured by standard peak analysis (see Fig. 4b). Multiple PCs show a prominent change in the vicinity of the P100 peak. This reflects that there is substantial individual variation in the shape and the amplitude of the waveform in this temporal window. Age-related variability was predominantly described by PC2 and PC3. PrVEP changes with age in youth may be due to maturation of neural circuits within the visual pathway.[18] Head circumference did not appear to account for the differences we observed across age, although we cannot rule out the possibility that other physical characteristics, such as skull thickness[19] or age-related closing of cranial sutures,[20] could have played a role. Of note, our model of age-related effects over adolescence may not generalize to other age groups. The use of PCA in younger and older age groups would be an interesting expansion of the work presented here.

We find good generalizability of the PCA model between training and validation subjects. Indeed, there were no differences in PC coefficients for the training and validation datasets even though there were a few subjects in the validation set on stimulant medications for ADHD. We also find good intrasubject reliability for prVEPs in our pediatric cohort using both the PCA model and standard peak analysis metrics. An advantage with the PCA model is it includes 7-dimensions for comparison, which we show improves the performance of matching a subject's session 1 to their own session 2 (see Supplementary Fig. S2b). We are unaware of a similar prior measurement in children, although the prVEP has also been found to have good reliability in adults.[21–23] Very few studies have examined the stability of VEP signals across sessions in the pediatric population. Schellberg and colleagues (1987) found substantial intrasubject variation of flash VEP between sessions in 26 children 10 to 13 years old spaced 10 months apart on average. The variability they reported was higher than what was observed in adult studies.[14] Their use of flash VEP, which can show higher intrasubject variability,[15] may at least partially account for the high intrasubject variability. Differences may also be in part due to reliance on peak analysis compared to PCA.

Prior work indicates that female subjects tend to have increased peak amplitudes and shorter peak latencies compared to male subjects.[24–26] The PCA model revealed a broader and lower amplitude P100 peak in male subjects, which is consistent with these prior results whereas providing a richer account of the differences in VEP waveforms between sexes. This effect achieved significance when subjects on neuroactive medications were excluded from the analysis (see Supplementary Fig. S4).

Global Temporal Analysis of VEP

Standard peak analysis of VEP remains an important tool for measuring the integrity of the visual system in many neurologic and ophthalmologic conditions. However, a focus upon the peak amplitude and peak latency in the prVEP necessarily limits the ability of an analysis to detect subtle, nonlocal changes in the shape of the prVEP waveform.[22] Sarnthein and colleagues (2009) used a combination of a metric of VEP shape and traditional peak analysis to address this limitation. They similarly demonstrated high test-retest reliability of prVEPs over 8 months in 10 healthy adult women using a combination of N75 and P100 peak-to-peak amplitude and peak time with pairwise regression of VEP waveforms to account for VEP shape.[22] Here, we describe a method using PCA that takes this one step further in eliminating focus on predefined peaks and quantifying instead the most informative individual variability in the temporal waveform.

Although there are many advantages, the PCA method has the limitation that the components are not temporally constrained so they cannot be as easily linked to discrete temporal physiologic events.[27] Many studies have focused on linking the N75, P100, and N135 peaks of the prVEP to different stages of processing in the visual hierarchy.[28] Compromise between the unconstrained decomposition of PCA and a rigid focus on individual peaks may be found in the use of informed basis functions, as has been used in the spatial analysis of functional magnetic resonance imaging (fMRI)[29] and EEG[30] data. In such an approach, the prVEP is projected onto a set of components that are crafted to reflect both temporally local features and global variation. The development and use of informed basis sets offers a promising direction for future prVEP analysis.

Although PCA has been applied to spatial localization in multifocal VEP studies,[27,31] we are unable to find prior examples of PCA or other dimensionality reduction approaches being used to characterize individual differences in the time series of prVEP or other EEG data. Indeed, there have been calls to increase the use of model-based inference in EEG.[32] We have demonstrated that modeling the global temporal variability of prVEP using PCA is highly generalizable and repeatable in a pediatric cohort. These results are promising for the application of this approach to the study longitudinal pediatric data. Further, our modeling approach may be used to capture and remove variability associated with maturation. This offers a means of addressing confounds of unmatched age differences in the populations, and testing for the subtle effects of a disease upon maturation. Finally, this model has the potential to identify differences between normal and pathologic groups that are not captured by focusing on specific predefined peaks.

Conclusion

PCA is a highly repeatable and generalizable method of analyzing prVEP data and offers a useful means of addressing variability during maturation in youth. These features are especially advantageous for longitudinal study designs. PCA offers a complimentary approach to standard peak analysis that can account for the global temporal variability in the prVEP waveform, which warrants further study in both healthy and diseased states.

Acknowledgments

The authors thank Fairuz Mohammed, Anne Mozel, and Olivia Podolak for their support in collecting and organizing the data in these experiments. The authors wish to thank the two anonymous reviewers, who provided patient, helpful, and constructive advice that markedly improved the manuscript and saved us from errors.

Supported by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health R01NS0975NIH, the Pennsylvania Department of Health SAP100077078, the American Academy of Neurology Clinical Research Scholarship (to C.P.G.), the International Headache Academy Research Fellowship (to C.P.G.), the Minds Matter Frontiers grant from the Children's Hospital of Philadelphia, and Department of Defense Grant W81XWH-151-0447 (to G.K.A.).

Disclosure: **C. Patterson Gentile,** (N); **N.R. Joshi,** (N); **K.J. Ciuffreda,** (N); **K.B. Arbogast,** (N); **C. Master,** (N); **G.K. Aguirre,** (N)

References

1. Brigell M, Kaufman DI, Bobak P, Beydoun A. The pattern visual evoked potential – a multicenter study using standardized techniques. *Doc Ophthalmol*. 1994;4:308.
2. Barton JL, Garber JY, Klistorner A, Barnett MH. The electrophysiological assessment of visual function in multiple sclerosis. *Clin Neurophysiol Pract*. 2019;4:90–96.
3. Bobak P, Khanna P, Goodwin J, Brigell M. Pattern visual evoked potentials in cases of ambiguous acuity loss. *Doc Ophthalmol*. 1993;85:185–192.
4. Moore RD, Broglio SP, Hillman CH. Sport-related concussion and sensory function in young adults. *J Athl Train*. 2014;49:36–41.
5. Baillargeon A, Lassonde M, Leclerc S, Ellemberg D. Neuropsychological and neurophysiological assessment of sport concussion in children, adolescents and adults. *Brain Inj*. 2012;26:211–220.
6. Fimreite V, Ciuffreda KJ, Yadav NK. Effect of luminance on the visually-evoked potential in visually-normal individuals and in mTBI/concussion. *Brain Inj*. 2015;29:1199–1210.
7. Yadav NK, Ciuffreda KJ. Optimization of the pattern visual evoked potential (VEP) in the visually-normal and mild traumatic brain injury (mTBI) populations. *Brain Inj*. 2013;27:1631–1642.
8. Áfra J, Proietti Cecchini A, Sándor PS, Schoenen J. Comparison of visual and auditory evoked cortical potentials in migraine patients between attacks. *Clin Neurophysiol*. 2000;111:1124–1129.
9. Allison T, Wood CC, Goff WR. Brain stem auditory, pattern-reversal visual, and short-latency somatosensory evoked potentials: latencies in relation to age, sex, and brain and body size. *Electroencephalogr Clin Neurophysiol*. 1983;55:619–636.
10. Emmerson-Hanover R, Shearer DE, Creel DJ, Dustman RE. Pattern reversal evoked potentials: gender differences and age-related changes in amplitude and latency. *Electroencephalogr Clin Neurophysiol Evoked Potentials*. 1994;92:93–101.
11. Wright CE, Williams DE, Drasdo N, Harding GFA. The influence of age on the electroretinogram and visual evoked potential. *Doc Ophthalmol*. 1985;59:365–384.
12. Snyder EW, Dustman RE, Shearer DE. Pattern reversal evoked potential amplitudes: life span changes. *Electroencephalogr Clin Neurophysiol*. 1981;52:429–434.
13. Shearer DE, Dustman RE. The pattern reversal evoked potential: the need for laboratory norms. *Am J EEG Technol*. 1980;20:185–200.
14. Schellberg D, Gasser T, Köhler W. The intra-individual reproducibility of flash-evoked potentials in a sample of children. *Int J Psychophysiol*. 1987;5:135–143.
15. Odom JV, Bach M, Brigell M, et al. ISCEV standard for clinical visual evoked potentials: (2016 update). *Doc Ophthalmol*. 2016;133:1–9.
16. Mulaik SA. *Foundations of Factor Analysis*. Second Edition. Boca Raton, FL: CRC Press; 2009:217–226.
17. Homan RW, Herman J, Purdy P. Cerebral location of international 10-20 system electrode placement. *Electroencephalogr Clin Neurophysiol*. 1987;66:376–382.
18. Brecelj J, Štrucl M, Zidar I, Tekavčič-Pompe M. Pattern ERG and VEP maturation in schoolchildren. *Clin Neurophysiol*. 2002;113:1764–1770.
19. Antonakakis M, Schrader S, Aydin Ü, et al. Inter-subject variability of skull conductivity and thickness in calibrated realistic head models. *Neuroimage*. 2020;223:117353.
20. Sabini RC, Elkowitz DE. Significance of differences in patency among cranial sutures. *J Am Osteopath Assoc*. 2006;106(10):600–604.
21. Oken BS, Chiappa KH, Gill E. Normal temporal variability of the P100. *Electroencephalogr Clin Neurophysiol Evoked Potentials*. 1987;68:153–156.
22. Sarnthein J, Andersson M, Zimmermann MB, Zumsteg D. High test-retest reliability of checkerboard reversal visual evoked potentials (VEP) over 8 months. *Clin Neurophysiol*. 2009;120:1835–1840.
23. Mellow TB, Liasis A, Lyons R, Thompson DA. The reproducibility of binocular pattern reversal visual evoked potentials: a single subject design. *Doc Ophthalmol*. 2011;122:133–139.
24. Sokol S, Moskowitz A, Towle VL. Age-related changes in the latency of the visual evoked potential: influences of check size. *Electroencephalogr Clin Neurophysiol*. 1981;51:559–562.
25. Tobimatsu S, Kurita-Tashima S, Nakayama-Hiromatsu M, Akazawa K, Kato M. Age-related changes in pattern visual evoked potentials: differential effects of luminance, contrast and check size. *Electroencephalogr Clin Neurophysiol Evoked Potentials*. 1993;88:12–19.
26. Celesia GG, Kaufman D, Cone S. Effects of age and sex on pattern electroretinograms and visual evoked potentials. *Electroencephalogr Clin Neurophysiol Evoked Potentials*. 1987;68:161–171.

27. Maier J, Dagnelie G, Spekreijse H, van Dijk BW. Principal components analysis for source localization of VEPs in man. *Vision Res*. 1987;27:165–177.
28. Tobimatsu S, Celesia GG. Studies of human visual pathophysiology with visual evoked potentials. *Clin Neurophysiol*. 2006;117:1414–1433.
29. Kiebel SJ, Goebel R, Friston KJ. Anatomically informed basis functions. *Neuroimage*. 2000;11:656–667.
30. Phillips C, Rugg MD, Friston KJ. Anatomically informed basis functions for EEG source localization: combining functional and anatomical constraints. *Neuroimage*. 2002;16:678–695.
31. Zhang X, Hood DC. A principal component analysis of multifocal pattern reversal VEP. *J Vis*. 2004;4:4.
32. Bridwell DA, et al. Moving beyond ERP components: a selective review of approaches to integrate EEG and behavior. *Front Hum Neurosci*. 2018;12:106.