**TITLE**: Developmental effects on pattern visual evoked potentials characterized by principal component analysis.

**Authors and Affiliations:**
Carlyn Patterson Gentile MD, PhD (corresponding author)
Children's Hospital of Philadelphia
University of Pennsylvania

Nabin R. Joshi B.Optom, PhD
State University of New York College of Optometry

Kenneth J. Ciuffreda OD, PhD
State University of New York College of Optometry

Kristy B. Arbogast PhD
Children's Hospital of Philadelphia
University of Pennsylvania

Christina Master MD
Children's Hospital of Philadelphia

Geoffrey K. Aguirre MD, PhD
University of Pennsylvania

**Commercial Relationships Disclosures:**
C.P.G.: none
N.R.B.: none
K.J.C.: none
K.B.A.: none
C.M.: none
G.K.A.: none

**Funding:** This work was 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.).

**Acknowledgements:** The authors would like to thank Fairuz Mohammed, Anne Mozel, and Olivia Podolak for their support in collecting and organizing the data in these experiments.

**Keywords:** principal component analysis, visual evoked potentials, pediatric, development

**Abbreviations:**

ADHD = Attention Deficit Hyperactivity Disorder
AMPS = Amplified Pain Syndrome
fMRI = functional magnetic resonance imaging
EEG = electroencephalogram
PCA = Principal Component Analysis
PC = Principal Component
PMH = Past Medical History
POTS = Postural orthostatic tachycardia syndrome
prVEP = Pattern reversal visual evoked potential
SVD = Singular value decomposition

**ABSTRACT**

**Purpose**: Peak amplitude and 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 11-19 years during two sessions (spaced 0.7 to 17 months apart). 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 "test" subject group (n=40) and calculated the intra-subject reliability of the PC coefficients between the two sessions. We then explored the effect of subject age and sex upon the PC coefficients.

**Results:** Seven PCs accounted for 96.0% of the variability. The model was generalizable (training vs. test coefficient distributions $p>0.36$ for all PCs) with good within-subject reliability ($R>0.7$ for all PCs). The PCA model did not show a significant difference between males and females ($F_{(7,147)}=1.69$, $p=0.12$), but showed a significant effect of subject age ($F_{(7,147)}=4.37$, $p=0.0002$).

**Conclusions:** PCA is a generalizable, reliable, and unbiased method of analyzing prVEP that can quantify and remove developmental variability present in the global temporal VEP signal.

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

## INTRODUCTION:

Visual evoked potentials (VEP) are a relatively inexpensive and easily implemented method of probing the cortical visual system. Pattern reversal (pr) VEPs, which are generated by viewing reversing checkerboard patterns, have the additional desirable property of relatively high intra-subject 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 non-specific 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 intra-subject 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 amplitude and latency measurements of predictable positive and negative peaks that occur in the first 150ms of the response[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], while 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 non-local temporal VEP waveform shape that are not fully captured by restricting analysis to pre-defined 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 meaningful natural variation may be restricted to a subset of possible responses, PCA may be used to reduce the dimensionality of data while preserving informative variability[16]. This approach does not require specific reference to pre-defined, predictable peaks in the response. 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 prVEPs collected from 155 subjects who were between 11 and 19 years old. Using separate training and test datasets comprised of 40 subjects each, we demonstrate that a particular set of PCA components provide for generalizable analysis of prVEP data. Using measurements from two separate sessions in the test 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 pre-season VEP data on healthy athletes. 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. prVEP data were collected on 155 subjects between February 2018 and February 2020. One-hundred and five subjects had at least two 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 past medical history (PMH) were 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). Testing was done with best corrected vision. Two subjects had a past medical history of a 'lazy eye', two subjects had a history of eye patching, and one subject had a history of strabismus. Exclusion of these 5 subjects did not significantly change the outcome of the results.

### Study paradigm

*Viewing conditions:* Subjects were seated in a well-lit, quiet room approximately 39 cm 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" LCD screen with 1280 x 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 largely reflected ISCEV recommendations for VEP[15]. Visual stimuli consisted of a wide field (51 by 51-degrees visual angle), 85% contrast checkerboard with a pattern reversal rate of 2 reversals per second. Checks were 1.6-degree 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 5 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®, Pine Brook, NJ). Two reference electrodes were placed in the center of the participant's forehead and on either the right or left temple, with the active electrode placed at Oz following the 10-20 international EEG placement criteria[17]. Data were recorded at a sampling rate of 1024 samples/second. Raw voltage-by-time data were exported for subsequent analysis.

**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 500ms intervals corresponding to each pattern reversal. Pattern reversals were excluded if there was a voltage change of >1 mV indicating a large non-physiologic change in voltage (~2% of reversals were discarded by this criterion). Signal-to-noise ratios (SNR) 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 normalizing the mean of the first 50ms to 0.

*VEP responses:* The mean prVEP response for each subject was calculated across all reversals collected in a given session. Each session consisted of 5, 20-second blocks, yielding a total of 200 reversals. Representations of VEP responses show the mean across-subject VEP with 95% confidence intervals obtained by bootstrap analysis across subjects with replacement.

*PCA model*: Twenty females and 20 males were randomly selected from a pool of 90 subjects who had at least two recorded sessions, less than 6 months apart (Figure S1). These subjects constituted the training dataset. Mean VEP responses 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, ∑ 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 500ms 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 test dataset was projected onto the PCA model. The distribution of coefficients for each PC were compared between the training and test datasets using a two-sample Kolmogorov-Smirnov test.

We examined the reproducibility of coefficients within subject. Each session from the test 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. The Euclidean distance between the set of 7 PC coefficients derived from sessions 1 and 2 was computed for all subject combinations. The Euclidean distance between a subject's matched session 1 and session 2 were then ranked out of the 40 possible combinations, with a rank of 1 indicating that the closest value from session 2 in Euclidean space to a given subject's session 1 value was their own matched session 2.

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

## RESULTS:

### Seven components are sufficient to capture prVEP variability

The VEP response (averaged across sessions) from 40 randomly selected subjects (20 females, 20 males) was used to derive the PCA model (i.e., training dataset). The mean VEP response for the training dataset is shown in Figure 1a. Seven PCs (Figure 1b) accounted for 96.0% of the variability in this sample (Figure 1c). We refer to these 7 PCs generated by the training dataset as the PCA model. As the PCA was decentered, the first PC approximates the mean prVEP

response. PCs 2 though 7 do not clearly map to particular peaks or points of time within the VEP 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 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 test dataset**

Forty age- and sex-matched subjects were selected to test the generalizability of the PCA model. There were no significant differences in patient demographics including race/ethnicity and PMH between the training and test subjects (Table 1).

Measurements from the test subjects were projected onto the PCA model generated from the training subjects. If the PCA model is generalizable, it should be able to capture variance in the test dataset similar to its performance in the training dataset. We found that the 7 PCs of the model accounted for 90.5% of the variance in the test dataset.

As a second test of generalizability, we considered that the distributions of PC coefficients derived using the PCA model in the test 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 test (blue) datasets (Figure 2a). Comparison of the mean and distribution of PC coefficients between the datasets revealed

no significant differences (Figure 2a, $p>0.3$ for all PCs; Kolmogorov–Smirnov test). There was also no difference between training and test datasets in the 7 PC multidimensional space (ANOVA $F_{(1,78)}=0.03$, $p=0.86$).

**Coefficients derived using PCA model have good within subject re-test reliability**

The PCA model was created using the test dataset and after averaging the data 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 (Figure 2b). 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) (Figure S2a). The intra-subject reliability of the coefficients of the prVEP declines across the components, indicating that the higher dimensions of the model likely reflect non-reproducible noise in the measurement. Consequently, expressing prVEP responses by projection onto the first 7 PCs of the PCA model constitutes a noise-reduction technique.

The similarity of any two prVEP measurements may be expressed by projecting the responses upon the PCA model and measuring the Euclidean distance between the two sets of 7 PC coefficients. We 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. This would indicate very high intra-subject compared to inter-subject reliability and indicate stability of VEP signal for a single subject across 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 (Figure 2b, lower right). In 70% of the subjects, the best match to their session 1 response was their session 2 response. In over 85% of the subjects, the session 2 response was in the top 5 ranking of responses that were similar to session 1, compared to 40 possible subject session 1 and session 2 combinations. We examined the ability to match the response from a subject across sessions as a function of the number of PCs used to describe the data. Accuracy plateaued with inclusion of approximately 7 PCs further supporting the use of this dimensionality to describe the prVEP data (Figure S2b). Given the high test-retest reliability of the measurements, the mean prVEP response was calculated across two sessions for each subject (when available) for the remaining analyses.

The remaining analyses were performed on the full cohort of 155 subjects. Demographics are shown for the full cohort (Table 2).

**There was not a significant effect of sex on prVEP in the PCA model**

A linear regression model was generated to examine the effect of the 7 PCs upon the sex of a subject. We did not observe a significant difference between males and females ($F_{(7,147)}=1.69$, $p=0.12$). Comparison of mean male and female prVEPs (Figure S3a) and the distribution of the 7PC components (Figure S3b) are shown. There were no significant differences of PC coefficients between males and females after correction for multiple comparisons.

**The PCA model captures variability throughout maturation in our pediatric cohort**

A linear regression model was generated to examine the effect of the 7 PCs upon the age of a subject. There was a significant, omnibus effect of age upon the 7 PC coefficients ($F_{(7,147)}$=4.37, p=0.0002). The PC coefficients that showed the highest correlation with age were PC2 ($t_{(1,147)}$=-4.06, p=$8e^{-5}$), PC3 ($t_{(1,147)}$=-3.21, p=0.002) and PC6 ($t_{(1,147)}$=-1.76, p=0.08). Only effects of PC2 and PC3 were significant. PC coefficients for the 7 PCs are shown as a function age (Figure S4). The prVEP response from any one subject can be described as a point in the 7-dimensional PC space. We plotted the responses for our subjects along the sub-set of three dimensions that best described the effect of age (Figure 3a). 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 response corresponding to subjects of different ages (Figure 3b). These waveforms capture multiple, non-local changes as a function of increasing subject age, including a progressively smaller P100 amplitude, shortening latency, and an increasing N135 amplitude.

We considered if a non-neural developmental change like head circumference could account for age effects in the prVEP response. We examined the correlation between age and head circumference in the 124 subjects with this measure and found a non-significant correlation (R=-0.04, p=0.66). There was no relationship between individual differences in head circumference and the set of 7 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 inter-subject variability of prVEP across 40 pediatric subjects for both a training and test 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.

### Comparison to previous studies

Prior studies that have relied on peak analysis have generally found decreasing peak amplitude and decreasing peak latency with maturation, though 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 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-N70 amplitude increased until about age 13 years and then decreased, while the N70-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 7-18 years old that P100 latency decreased with maturation[18]. Allison and colleagues (1983) similarly found the VEP P100 latency decreased between 4-19 years of age, although the latency of other peaks and peak amplitude were not reported[9]. While our study examines a more global representation of temporal variability, we also observe that prVEP amplitude and P100 peak latency decreases with maturation (see Figure 3b). 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, though we cannot rule out the possibility that other physical characteristics like skull thickness[19] could have played a role.

We find good intra-subject reliability for the prVEP in our pediatric cohort. We are unaware of a similar prior measurement in children, although the prVEP has also been found to have good reliability in in adults[20–22]. Very few studies have examined the stability of VEP signals across sessions in the pediatric population. Schellberg and colleagues (1987) found substantial intra-subject variation of flash VEP between sessions in 26 children 10 to 13 years old. The variability they reported was higher than what was observed in adult studies[14]. Their use of flash VEP, which can show higher intra-subject variability[15], may at least partially account for the high intra-subject variability. Differences may also be in part to reliance on peak analysis compared to PCA.

We did not observe significant sex differences. Prior work indicates that females tend to have increased peak amplitudes and shorter peak latencies compared to males[23–25]. We see a similar trend in the mean prVEP responses of males and females in our data (figure S3), but these effects are not significant in our analyses.

**Global temporal analysis of VEP**

A focus upon the amplitude and latency of specific peaks in the prVEP necessarily limits the ability of an analysis to detect subtle, non-local changes in the shape of the response waveform[21]. 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[21]. Our method using PCA takes this one step further in eliminating focus on predefined peaks and quantifying instead the most informative individual variability in the temporal waveform.

Though there are many advantages, the PCA method has the limitation that the components are not temporally constrained so they cannot be as easily be linked to discrete temporal events[26]. Many studies have focused on linking the N75, P100, and N135 peaks of the prVEP to different stages of processing in the visual hierarchy[27]. 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 fMRI[28] and EEG[29] data. In such an approach, the response 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.

While PCA has been applied to spatial localization in multi-focal VEP studies[26,30], we are unable to find prior examples of PCA or other dimensionality reduction approaches being used to characterize individual differences in the timeseries of prVEP or other EEG data. Indeed, there have been calls to increase the use of model-based inference in EEG[31]. 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 pre-defined 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. PCA is a promising approach to

prVEP data analysis comparing healthy and pathologic pediatric cohorts, particularly for longitudinal study designs.

**TABLES**

| | Training subjects | Test subjects | |
|---|---|---|---|
| No. of subjects | 40 | 40 | |
| Male | 20 | 20 | |
| Female | 20 | 20 | |
| Median age at first VEP [age range] | 15.3 years [11.2 – 19.1] | 15.3 years [11.8 – 18.1] | t-test 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 | |
| PMH | | | 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 |
| Postural orthostatic tachycardia syndrome (POTS) | 0 | 1 | p=0.32 |

Table 1: Subject demographics for the training and test subject groups used to generate and validate the PCA model, respectively. No subjects reported a PMH of dyslexia, bipolar, drug/alcohol use disorder, autism, epilepsy, tic disorder, or amplified musculoskeletal pain syndrome (AMPS).

| Full Cohort | |
|---|---|
| No. of subjects | 155 |
| Male | 68 (43.8%) |
| Female | 87 (56.2%) |
| Median age at first VEP [age range] | 15.2 years [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%) |
| PMH | |
| 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%) |

Table 2: Subject demographics of the full pediatric cohort. No subjects reported a PMH of bipolar, drug/alcohol use disorder, autism, epilepsy, tic disorder, or AMPS.

1. Brigell, M., Kaufman, D. I., Bobak, P. & Beydoun, A. The pattern visual evoked potential - A multicenter study using standardized techniques. *Doc. Ophthalmol.* **4**, 308 (1994).
2. Barton, J. L., Garber, J. Y., Klistorner, A. & Barnett, M. H. The electrophysiological assessment of visual function in Multiple Sclerosis. *Clin. Neurophysiol. Pract.* **4**, 90–96 (2019).
3. Bobak, P., Khanna, P., Goodwin, J. & Brigell, M. Pattern visual evoked potentials in cases of ambiguous acuity loss. *Doc. Ophthalmol.* **85**, 185–192 (1993).
4. Moore, R. D., Broglio, S. P. & Hillman, C. H. Sport-related concussion and sensory function in young adults. *J. Athl. Train.* **49**, 36–41 (2014).
5. Baillargeon, A., Lassonde, M., Leclerc, S. & Ellemberg, D. Neuropsychological and neurophysiological assessment of sport concussion in children, adolescents and adults. *Brain Inj.* **26**, 211–220 (2012).
6. Fimreite, V., Ciuffreda, K. J. & Yadav, N. K. Effect of luminance on the visually-evoked potential in visually-normal individuals and in mTBI/concussion. *Brain Inj.* **29**, 1199–1210 (2015).
7. Yadav, N. K. & Ciuffreda, K. J. Optimization of the pattern visual evoked potential (VEP) in the visually-normal and mild traumatic brain injury (mTBI) populations. *Brain Inj.* **27**, 1631–1642 (2013).
8. Ófra, J., Proietti Cecchini, A., Sándor, P. S. & Schoenen, J. Comparison of visual and auditory evoked cortical potentials in migraine patients between attacks. *Clin. Neurophysiol.* **111**, 1124–1129 (2000).
9. Allison, T., Wood, C. C. & Goff, W. R. 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.* **55**, 619–636 (1983).
10. Emmerson-Hanover, R., Shearer, D. E., Creel, D. J. & Dustman, R. E. Pattern reversal evoked potentials: gender differences and age-related changes in amplitude and latency. *Electroencephalogr. Clin. Neurophysiol. Evoked Potentials* **92**, 93–101 (1994).
11. Wright, C. E., Williams, D. E., Drasdo, N. & Harding, G. F. A. The influence of age on the electroretinogram and visual evoked potential. *Doc. Ophthalmol.* **59**, 365–384 (1985).
12. Snyder, E. W., Dustman, R. E. & Shearer, D. E. Pattern reversal evoked potential amplitudes: Life span changes. *Electroencephalogr. Clin. Neurophysiol.* **52**, 429–434 (1981).
13. Shearer, D. E. & Dustman, R. E. The pattern reversal evoked potential: The need for laboratory norms. *Am. J. EEG Technol.* **20**, 185–200 (1980).
14. Schellberg, D., Gasser, T. & Köhler, W. The intra-individual reproducibility of flash-evoked potentials in a sample of children. *Int. J. Psychophysiol.* **5**, 135–143 (1987).
15. Odom, J. V. *et al.* ISCEV standard for clinical visual evoked potentials: (2016 update). *Doc. Ophthalmol.* **133**, 1–9 (2016).
16. Mulaik, S. A. *Foundations of Factor Analysis*. *Foundations of Factor Analysis* (CRC press, 2009). doi:10.1201/b15851
17. Homan, R. W., Herman, J. & Purdy, P. Cerebral location of international 10-20 system electrode placement. *Electroencephalogr. Clin. Neurophysiol.* **66**, 376–382 (1987).
18. Brecelj, J., Štrucl, M., Zidar, I. & Tekavčič-Pompe, M. Pattern ERG and VEP maturation in schoolchildren. *Clin. Neurophysiol.* **113**, 1764–1770 (2002).

19. Antonakakis, M. *et al.* Inter-Subject Variability of Skull Conductivity and Thickness in Calibrated Realistic Head Models. *Neuroimage* **223**, 117353 (2020).
20. Oken, B. S., Chiappa, K. H. & Gill, E. Normal temporal variability of the P100. *Electroencephalogr. Clin. Neurophysiol. Evoked Potentials* **68**, 153–156 (1987).
21. Sarnthein, J., Andersson, M., Zimmermann, M. B. & Zumsteg, D. High test-retest reliability of checkerboard reversal visual evoked potentials (VEP) over 8 months. *Clin. Neurophysiol.* **120**, 1835–1840 (2009).
22. Mellow, T. B., Liasis, A., Lyons, R. & Thompson, D. A. The reproducibility of binocular pattern reversal visual evoked potentials: A single subject design. *Doc. Ophthalmol.* **122**, 133–139 (2011).
23. Sokol, S., Moskowitz, A. & Towle, V. L. Age-related changes in the latency of the visual evoked potential: Influences of check size. *Electroencephalogr. Clin. Neurophysiol.* **51**, 559–562 (1981).
24. 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* **88**, 12–19 (1993).
25. Celesia, G. G., Kaufman, D. & Cone, S. Effects of age and sex on pattern electroretinograms and visual evoked potentials. *Electroencephalogr. Clin. Neurophysiol. Evoked Potentials* **68**, 161–171 (1987).
26. Maier, J., Dagnelie, G., Spekreijse, H. & van Dijk, B. W. Principal components analysis for source localization of VEPs in man. *Vision Res.* **27**, 165–177 (1987).
27. Tobimatsu, S. & Celesia, G. G. Studies of human visual pathophysiology with visual evoked potentials. *Clin. Neurophysiol.* **117**, 1414–1433 (2006).
28. Kiebel, S. J., Goebel, R. & Friston, K. J. Anatomically informed basis functions. *Neuroimage* **11**, 656–667 (2000).
29. Phillips, C., Rugg, M. D. & Friston, K. J. Anatomically informed basis functions for EEG source localization: Combining functional and anatomical constraints. *Neuroimage* **16**, 678–695 (2002).
30. Zhang, X. & Hood, D. C. A principal component analysis of multifocal pattern reversal VEP. *J. Vis.* **4**, 4 (2004).
31. Bridwell, D. A. *et al.* Moving beyond ERP components: A selective review of approaches to integrate EEG and behavior. *Front. Hum. Neurosci.* **12**, 106 (2018).

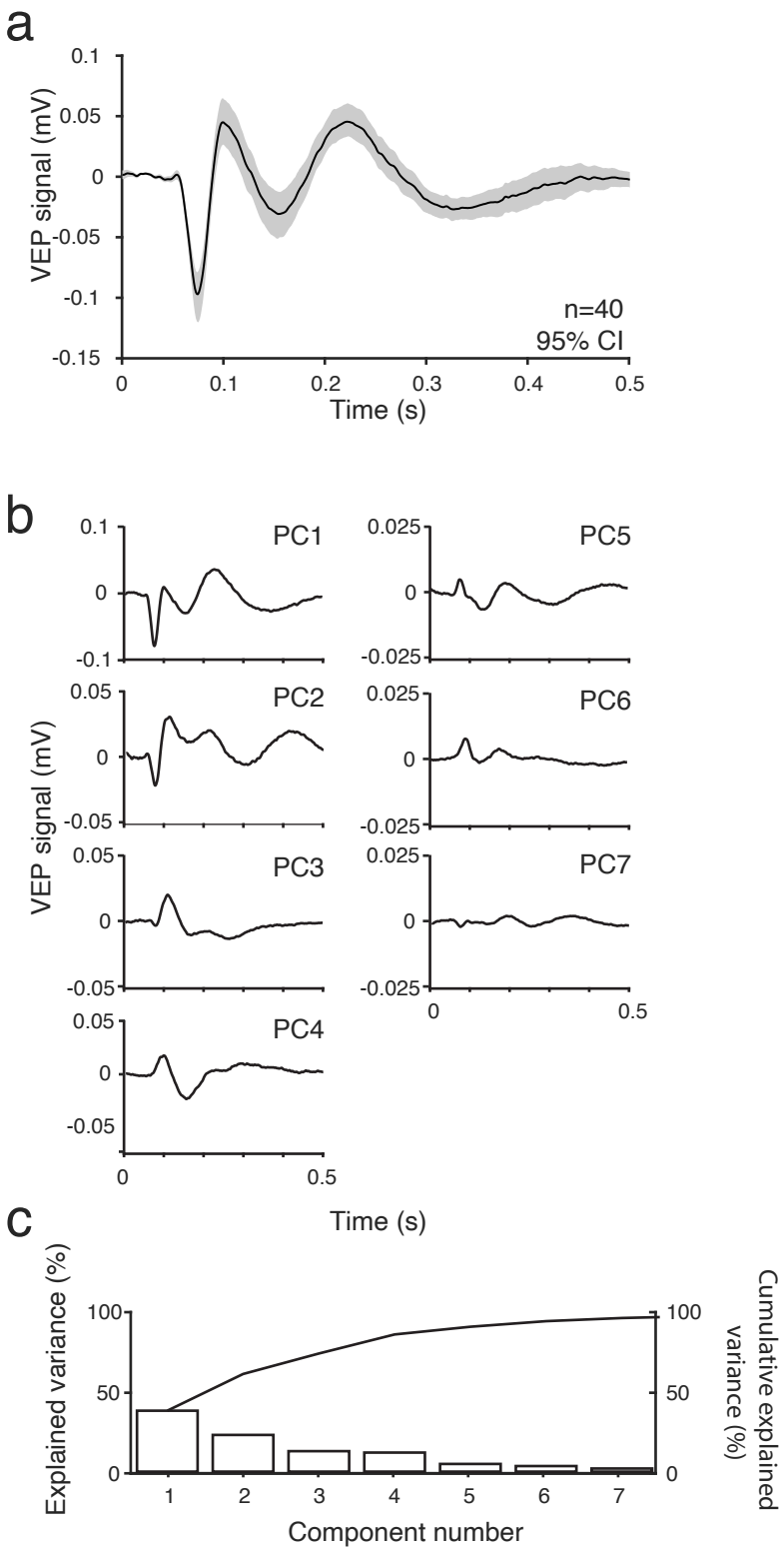


Figure 1: PCA Model. PCA was generated with 40 randomly selected subjects (20 males, 20 females). (a) mean VEP response across all 40 training subjects; gray represents 95% confidence interval (CI) by bootstrap analysis. (b) Principal components (PCs) 1 to 7, which account for 96.0% of the variability in the data. (c) Scree plot showing percent explained variance of PCs 1 to 7.

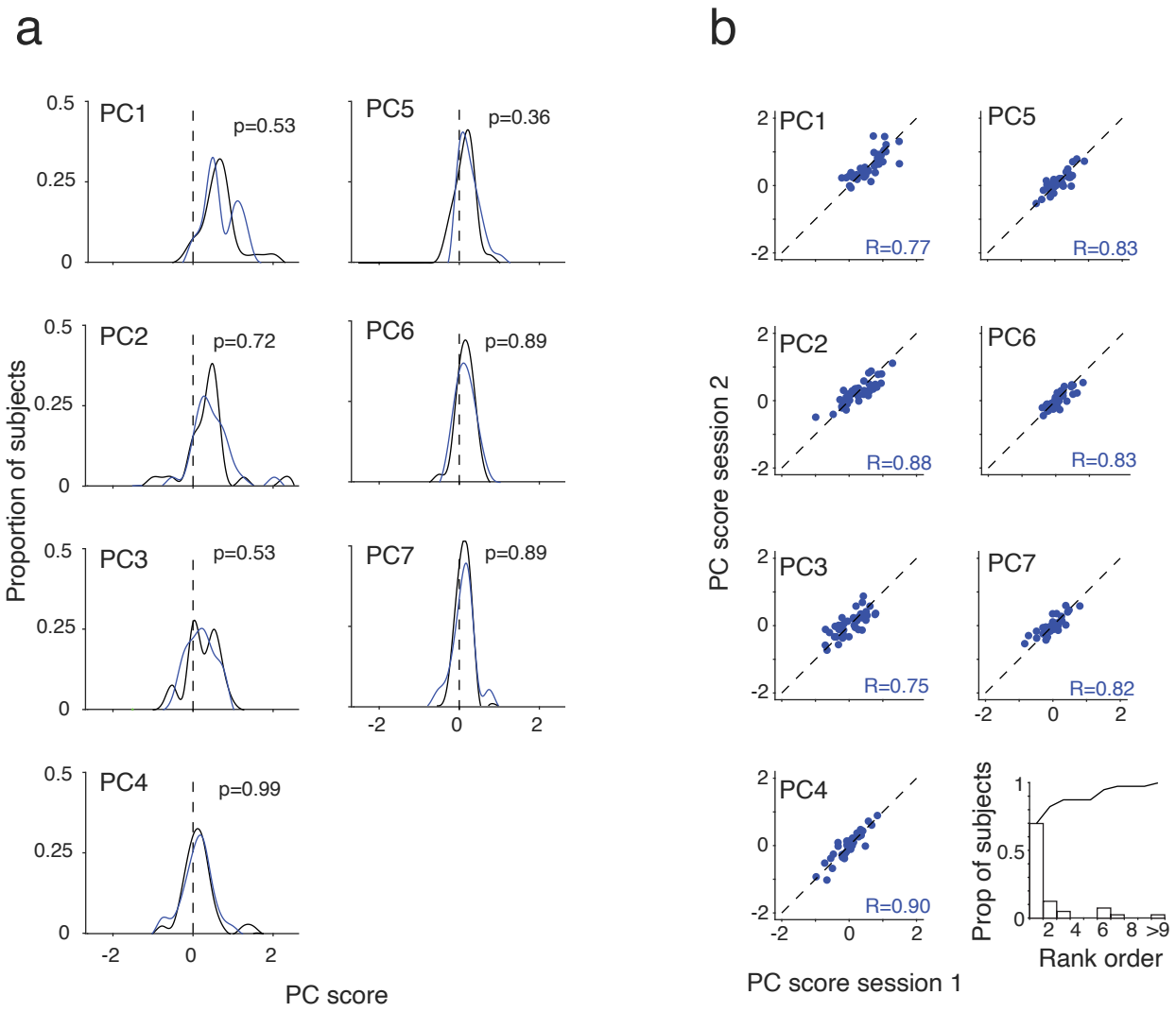


Figure 2: Generalizability and test-retest reliability of the PCA Model. (a) PC coefficent comparison of 40 training subjects (black) and 40 test subjects (blue) showed no significant differences in the mean or distribution of individual PCs (KS test p>0.3 for all PCs) or the multidimensional space across all PCs (ANOVA F(1,78)=0.03, p=0.87). (b) Scatter plots present the coefficients for each PC dervied 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 subjects. Black line shows the cumulative proportion of subjects.

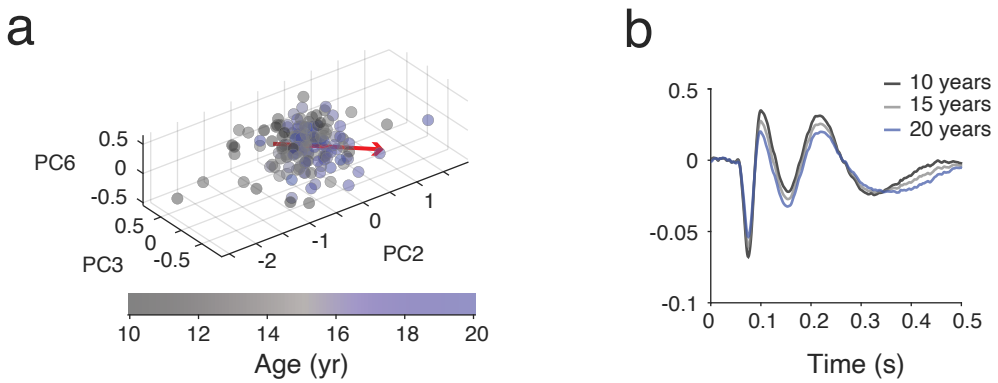


Figure 3: PC coefficients change systematically with age. (a) The response from each may be described by their PC2, PC3, and PC6 coefficients. These three dimensions of the PCA model had the strongest association with subject age. Plot points are colored from black to blue indicating increasing subject 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.