Reading fluency development is predicted by cortical latency.

Fang Wang*[1], Quynh Trang H. Nguyen*[1], Blair Kaneshiro[1],

Anthony M. Norcia[2,3], and Bruce D. McCandliss†[1]

[1] *Graduate School of Education, Stanford University, Stanford, CA, USA*

[2] *Department of Psychology, Stanford University, Stanford, CA, USA*

[3] *Wu Tsai Neurosciences Institute, Stanford, CA, USA*

February 25, 2025

*These authors contributed equally to this work.

†Corresponding author: Bruce D. McCandliss, brucemc@stanford.edu. Graduate School of Education, Stanford, CA 94305 USA

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethics Approval Statement

The study was approved by the Institutional Review Board of Stanford University. As participants were children, a parent or legal guardian of each participant reviewed a written description of the study and gave written informed consent before the session; each participant also assented to taking part in the research.

Acknowledgments

We thank the students and their families for participating. We also thank Elizabeth Y. Toomarian, Angie M. Wang, and Lindsey Hasak for their help with data collection.

Abstract

The development of fluent literacy skills from childhood to adolescence is strongly constrained by the temporal dynamics of word recognition. Capturing the neural basis of these subtle timing changes in word recognition has remained challenging with EEG measures that lack reliability at the individual subject level. Here, we leverage phase information from Steady-State Visual Evoked Potentials (SSVEPs) to derive precise and reliable temporal dynamics of neural signatures underlying visual word form recognition at the individual level and examine their relationship to reading fluency and comprehension. Typically developing readers (N = 68), aged 8–15 years, viewed a stream of four-character stimulus strings presented at 3 Hz. Significant SSVEP signals emerged for nearly all participants. Signals at 3, 6, and 9 Hz harmonics exhibited a phase pattern consistent with a delay model, indicating a mean latency of approximately 170 milliseconds. Individual variations in latencies demonstrated (a) high internal consistency ($R = .94$); (b) stability across variations in letter string forms (familiar words, nonwords with familiar letters, nonwords with unfamiliar pseudo-characters); (c) a linear relationship with age; and most remarkably, (d) a predictive relationship with individual variation in reading fluency and reading comprehension. These results establish SSVEP visual word form latency as a promising approach for investigating the neural basis of reading development, paving the way for future translational applications in education and offering potential solutions to broader societal challenges in promoting population-level reading fluency.

Keywords EEG SSVEP N170 Reliable Components Analysis (RCA) Visual Word Recognition Latency Typical Developing Children

1 Introduction

Reading fluency, which extends beyond mere accuracy measures, plays a pivotal role in both reading comprehension and overall literacy development (Yildiz et al., 2014). While various factors influence reading fluency across individuals (Pennington, 2006), word recognition speed stands out as a unique and significant predictor of reading ability (Landerl & Wimmer, 2008), independent of phonological processing and general performance skills (Macmillan & Creelman, 2004). Therefore, understanding the (neuro)temporal dynamics underlying visual word form recognition is crucial for a robust understanding of childhood reading fluency development.

Electroencephalography (EEG) holds promise for capturing the timing of neural dynamics of visual word recognition processes with millisecond-level precision. Attempts to study the peak timing of Event-Related Potential (ERP) components, such as N170, have been shown to be insensitive to developmental differences in reading development (Amora et al., 2022). In contrast, Steady-State Visual Evoked Potential (SSVEP) paradigms—where a sequence of stimuli are presented at a predefined periodic rate (e.g., 3 Hz)—provide robust and high signal-to-noise ratio (SNR) responses in nearly every subject with as few as 3 minutes of stimulation, and have been linked to reading development (van de Walle de Ghelcke et al., 2021; Wang et al., 2022).

The phase of SSVEP signals has proven to be highly informative in studies of the precise timing of neural processes. In 1970, Da Silva and colleagues calculated what has become to be called "apparent latency" of early cortical processes for vision by examining the slope of the *phase-vs-stimulus temporal frequency* function (Da Silva et al., 1970) after having presented the same stimulus condition at multiple temporal frequencies. When the phase-by-frequency plot is linear over a substantial range, the apparent latency can be calculated

Abbreviations: Reliable Components Analysis (RCA); steady-state visual evoked potentials (SSVEP).

using the following formula:

$$d = 1/360 \times \delta\varphi/\delta f$$

Here, d is the apparent latency in seconds, $\delta\varphi$ is the change in phase over the frequency range in degrees, and δf is the change in frequency in Hz. $\delta\varphi/\delta f$ is the slope of the phase-versus-frequency plot.

Yeatman & Norcia (Yeatman & Norcia, 2016) extended this phase-by-frequency latency method to study the temporal dynamics of word-selective SSVEP responses. They alternated full-screen images of scrambled and intact text across a range of frequencies from 1-12 Hz. Calculating the slope across these stimulation frequencies revealed early (141–160 msec) and late (>250 msec) left lateralized text-selective responses.

Presenting a single stimulus condition at multiple frequencies, as in Yeatman & Norcia (Yeatman & Norcia, 2016), is both time consuming in general and arduous in particular for children. Fortunately, a more efficient method has recently been validated, taking advantage of SSVEP paradigms that elicit multiple harmonics of the stimulus frequency, thereby providing the necessary frequency range to model latencies using the same procedures. Norcia et al. (Norcia et al., 2020) noted that the phase of the SSVEPs to contrast-modulated luminance probes increased linearly over the first several harmonics of the stimulation frequency. This linear function was then used to derive the latency for the SSVEPs using the *phase-by-harmonic* slope that emerged when results were averaged across the group of participants. The assumption underlying this approach is that the harmonics of the group-level evoked response come from a single cortical source with a fixed delay. The current study was designed to validate and extend this approach by fitting phase-by-harmonic slopes to multiple harmonics at the single-subject level.

Prior SSVEP work has used the "oddball" paradigm—where a sequence of *base* stimuli (e.g., pseudofont strings) are periodically replaced by contrastive *oddballs* (e.g., words) at a slower predefined rate. SSVEP responses emerging at the slower oddball frequencies are thought to reflect categorical discriminations related to higher-level visual processes such as

visual word recognition, while the responses at the faster base frequency (and its harmonics) are often interpreted as only reflecting lower-level sensory processes that are insensitive to visual word form processing (Lochy et al., 2016; Volfart et al., 2021). However, recent SSVEP studies using a 3 Hz base stimulation rate across adults and children have revealed multiple neural sources underlie the base response, some of which are consistent with higher-level visual word processing in regions such as the visual word form area (VWFA) (Wang et al., 2021, 2022). Such results highlight the importance of analyzing base responses in greater detail, as this could enhance our understanding of the neural basis of both general visual and word-related processes and their relationship with reading ability, potentially informing future research in this area.

Examination of the temporal dynamics of the base response is also merited by recent converging evidence suggesting that many aspects of developmental challenges in reading are linked to subtle temporal processing deficits (Stein, 2023). For example, McLean et al. found that word recognition speed is a unique and significant predictor of reading ability, independent of phonological processing or domain-general cognitive skills (McLean et al., 2011). Accordingly, the current study examines the temporal dynamics of the base responses using the phase-by-harmonic function validated in (Norcia et al., 2020) to derive cortical latencies of visual word form processes. Critically, this approach goes beyond modeling the central tendencies of group-level SSVEP responses, but also derives individual subject-level responses, which enable a rigorous examination of the connection between these cortical latencies and measures of reading skills in children at multiple level of analysis.

2 Methods

2.1 Participants

Data from 68 typical developing children (3rd-8th grade, between 8 and 15 years old, median age 11 years, 36 females) were analyzed. Data from three additional children were collected

but not analyzed here due to EEG data-quality issues (i.e., excessive movement during recording, incomplete EEG data, N = 2) and having cataract (N = 1). All participants had normal or corrected-to-normal visual acuity and had no diagnosed reading disabilities. After the study, each participant received a small toy for their time.

2.2 Behavioral Assessments

Each participant completed a 30-minute behavioral assessment either before or after the EEG session. All children were tested for handedness (Edinburgh Handedness Inventory, Oldfield (1971)), silent reading efficiency (i.e., speed and accuracy) and comprehension (Test of Silent Reading Efficiency and Comprehension, TOSREC, Wagner et al. (2010)), word reading efficiency (Test of Word Reading Efficiency, Second Edition, TOWRE-2, Torgesen et al. (2012)), and non-verbal intelligence and fluid reasoning (the Matrices subtest of the Kaufman Brief Intelligence Test, Second Edition Revised, KBIT-2, Kaufman (1990)).

2.3 Stimuli

Three types of stimuli—words, nonwords, and pseudofonts—were used, each comprising four elements (letters or pseudoletters). There were 30 exemplars for each type of stimulus, for 90 exemplars in total.

All words were common monosyllabic singular nouns that started and ended with consonants. They were chosen to have relatively high frequency (averaged 97.7 per million) with limited orthographic neighbors (average 2.3, range from 0 to 4) based on the Children's Printed Word Database (Masterson et al., 2010).

Nonwords were generated on an item-by-item basis by semi-randomly rearranging the letters of a corresponding Word while ensuring that they also started and ended with a consonant. Each nonword was meant to be unpronounceable and not follow English phonological rules. Nonwords did not differ significantly from the words on average unconstrained bigram frequency ($t(55) = -0.59, p = 0.55$).

Pseudofonts were the same words stimuli presented in Brussels Artificial Character Set font (BACS-2 Serif font, Vidal & Chetail (2017)), a strictly controlled set of artificial characters that were designed by manipulating the visual strokes of a letter while removing their associated characters.

Using these three types of stimuli, three experimental conditions were investigated. Conditions A and C involved Word embedded in a base stream of pseudofonts (Word & pseudofont) and nonwords (Word & nonword), respectively. For condition B, nonword were embedded in a stream of pseudofont (nonword & pseudofont). The three conditions were completed in a randomly assigned order and were presented at a 3 Hz base stream Figure 1.

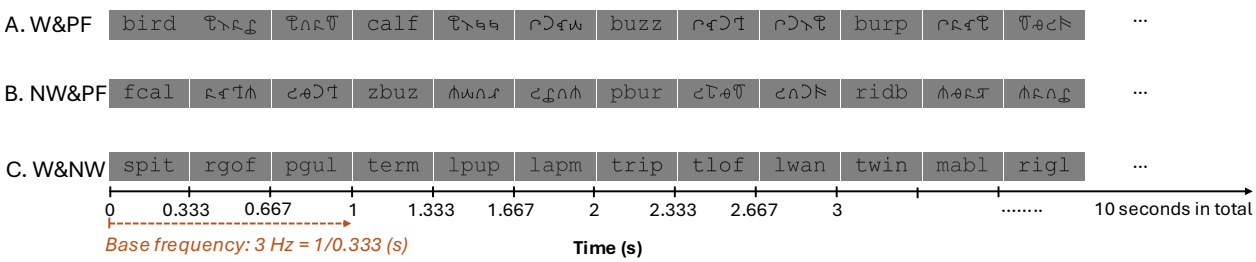

Figure 1: **Experimental Design.** Example sequences of stimuli presented in the experiment. 3 total presentations per second result in a 3 Hz base stream in all three conditions. A: Word (W) and pseudofont (PF) condition; B: Nonword (NW) and pseudofont (PF) condition; C: Word (W) and nonword (NW) condition. All stimuli were presented centered on the screen.

2.4 Experimental Procedure

Prior to the EEG experiment, the procedure was explained to the participant, after which the participant performed a brief practice session.

Participants were seated 1 meter away from the computer monitor in a darkened room during their EEG session. Each condition started with a blank screen shown for a random interval of 1500–2500 ms. During this time, participants were asked to fixate on a white cross in the middle of the screen. Then, each second (aka an epoch) includes the presentation

of three stimuli. Twelve successive epochs made up one trial. Ten trials were randomly presented for each condition.

In order to maintain participants' attention throughout the experiment, a font size change detection task was used. During the recording, the participant continuously fixated on the center of the screen, and pressed a button whenever they detected that the font size of the stimulus got bigger. The font size change detection task was on a "staircase" mode, during which the size of font change became smaller as the accuracy increased, or became bigger when the accuracy decreased. Two size change targets occurred in a non-periodic fashion within a trial.

Overall, the entire EEG experiment took around 30 minutes including setup, practice, and breaks between trials.

2.5 EEG Recording and Preprocessing

EEG data were collected using 128-sensor HydroCell arrays (MagstimEGI), Electrical Geodesics NetAmp300, and NetStation 5.4.2 software. Data were acquired using Cz as reference at a sampling rate of 500 Hz, and impedances were kept below 50 kΩ. Stimuli were presented using an in-house stimulus presentation software.

Before pre-processing, the data were shifted backwards in time to account for a 60 ms delay introduced by NetAmp300/NetStation software. Recordings were filtered in NetStation using a 0.3-50 Hz bandpass filter that was applied twice to minimize power from frequencies outside the filter range. Data were then imported into the in-house signal processing software for further pre-processing.

During pre-processing, EEG data were re-sampled to 420 Hz to ensure that the frame rate of 60 Hz included an integer number of time samples. Sensors were excluded if more than 15% of samples from the sensor exceeded a 60 μV amplitude threshold. Data from these sensors were replaced by the average value from six of their nearest neighboring sensors.

The continuous EEG data were then re-referenced to average reference (Lehmann &

Skrandies, 1980) and segmented into 1-second epochs. Epochs with more than 10% of data samples exceeding the noise threshold of 30 μV or any part of the sample exceeding the blink threshold of 60 μV were excluded from the analysis on a sensor-by-sensor basis. Artifact rejection was performed to remove epochs containing artifacts such as blinks or eye movements. If an epoch exceeded the peak/blink threshold in more than 7 sensors, the entire epoch would be removed in all sensors.

Recursive Least Squares (RLS) filters were then used to filter continuous EEG signals in the time domain (Tang & Norcia, 1995). The filters were tuned to each of the analysis frequencies (i.e., 3Hz, 6Hz, 9Hz) and converted to complex amplitude values by means of Fourier transform. Complex-valued RLS outputs were decomposed into real and imaginary coefficients for input to the spatial filtering computations of Reliable Components Analysis (RCA).

The first and last 1-second epochs of each 12-epoch trial—each comprising three stimulus events—were excluded from further analysis. This was done to reduce the influence of transient responses associated with initial exposure to the stimuli and due to more blinking often occurring at the beginning and end of a trial. Thus, analysis was performed on 10 1-second epochs per trial.

2.6 Analysis of EEG Data

2.6.1 Reliable Components Analysis (RCA)

Reliable Components Analysis (RCA (Dmochowski et al., 2012, 2015)) is a matrix decomposition technique that derives a set of physiologically interpretable *reliable components* (RCs) by maximizing trial-to-trial covariance relative to within-trial covariance. Since response phases of SSVEP are constant over repeated stimulations, RCA uses this trial-to-trial reliability to decompose the entire 128-sensor array into RCs, the activations of which reflect phase-locked activities.

Given a sensor-by-feature (i.e., real and imaginary Fourier coefficients) data matrix, RCA

computes linear weightings of sensors such that the resulting projected data would exhibit the maximal Pearson Product Moment Correlation Coefficients (Pearson, 1896) across trials. Each RC is a linear combination of sensors that can be visualized as scalp topographies using a forward-model projection of the eigenvectors (spatial filter vectors) (Parra et al., 2005).

Relative to other spatial filtering approaches such as Principal Components Analysis (PCA), RCA achieves high output SNR with a lower trial count (Dmochowski et al., 2015). This is especially useful for investigations involving children since it reduces the amount of data required to observe a robust signal, thereby shortening the EEG data acquisition time.

2.6.2 RCA Calculation

To test brain responses to visual word form stimuli, we computed RCA at the base frequency and its harmonics. Specifically, we input the real and imaginary frequency coefficients of the first three harmonics of the base frequency (i.e, 3 Hz, 6 Hz, and 9 Hz) over the 128-sensor array. To enable a direct quantitative comparison of the three conditions in a shared component space, we computed the RCA weights over the three conditions together.

2.6.3 Analysis of Component-Space Data

To choose which RCA components to examine, we first assessed statistical significance of each component's eigenvalue coefficient via permutation test (Supplement Figure S1). The permutation test involved generating a null distribution of eigenvalue coefficients against which to compare the observed (intact) RCA coefficients by performing the RCA on 500 versions of surrogate "phase scrambled" sensor-space EEG data. This method disrupts the across-trials covariance of the data and temporal structure of the signal while preserving autocorrelation characteristics and magnitude spectra (Prichard & Theiler, 1994). Here, the phases were randomized in the frequency domain by means of rotation matrices, as described by (Wang et al., 2022). Following this, we further identified the statistical significance at each harmonic for a given significant component via Hotelling's two-sample t^2 tests (Victor & Mast, 1991)

on distributions of real and imaginary Fourier coefficients on a per-harmonic, per-component basis. Specifically, for each RC, a 1-second epoch of component-space data contained 6 data points (3 harmonics $\times$ 2 real and imaginary Fourier coefficients). Component-space data were first averaged across 1-second epochs (10 epochs per trial, for 10 trials in total in present study) on a per-participant basis. Statistical analyses were then performed across distributions of participants. Multiple comparisons were corrected using False Discovery Rate (FDR (Benjamini & Yekutieli, 2001)). Twenty-seven comparisons (3 harmonics $\times$ 3 components $\times$ 3 conditions) were corrected for base analyses where RCA was computed over the data from three conditions together.

To report SSVEP amplitudes across multiple significant harmonics, we used the square root of the summed squared amplitudes ("root sum square", RSS) across multiple harmonics to combine harmonic response amplitude (Tlumak et al., 2011). We refer to this value as the RSS amplitude, in units of μV.

2.6.4 Group- and Individual-Level Latencies of RCs

Following the statistical analyses of RCs, we derived group-level latency calculations by fitting a line through the phases at harmonics with significant responses using a linear regression method, where the slope and intercept were obtained with MATLAB's `polyfit` function (Norcia et al., 2020). The slope of the line represents the response latency.

In order to examine the relationship between brain processing speed and reading abilities, latency was also calculated at the individual subject level. The high SNR provided by SSVEP facilitates this individual-level analysis. We projected each subject's sensor-space data through the spatial filter vectors derived from group-level RC analyses, resulting in individual projected data. As with the group-level calculations, we fit a linear regression line when responses were statistically significant in at least two harmonics, with the precondition that including the third non-significant harmonic would not substantially change the slope (i.e., latency).

2.6.5 Validation of Individual-Level Latency

To confirm the reliability of our individual latency estimation method and ensure that the estimated latency represents signal rather than noise, we conducted several comparisons. First, we compared latencies at both group and individual levels to check for consistency in their range. Second, we compared individual latencies across the three conditions using paired sample t-tests. Finally, we assessed the split-half reliability to rule out potential confounds from different stimulus sets/conditions in latency calculations. Specifically, we concatenated all 30 trials (epochs/seconds) from the three conditions and randomly split them into two halves, each containing 15 trials—five randomly selected from each condition. We then applied the individual-level latency calculation methods described in the previous section to obtain latencies for each subject in both halves and computed the correlation between them.

2.6.6 Visualization of Component-Space Data

We visualized the RCA data in three ways. First, we present topographic maps for the spatial filters of each component. Second, we present bar plots of amplitudes (μV) across harmonics, with significant responses (according to adjusted p_{FDR} values of Hotelling's t^2 tests of the complex data) indicated with asterisks. Finally, we present phase values (radians) plotted as a function of harmonic, when responses are significant for at least two harmonics, this is accompanied by a line of best fit and slope (latency estimate).

2.6.7 Analysis of brain-behavior relationships

Brain-behavior analyses were then performed to assess individual variations in the relationship between latency and participants' reading scores and biological age. The reading scores analyzed were the participants' raw scores of TOWRE and TOSREC, representing reading efficiency and reading comprehension, respectively. Because latency did not differ significantly between conditions, the averaged latency across conditions was used for this brain-behavior

correlation. Permutation test results showed RC1 explained the most reliability (Supplement Figure S1), so only latencies of RC1 were included.

Linear correlations were calculated between individual latencies and the TOWRE, TOSREC scores, and age. When significant correlation was found, outliers were identified and removed using a three standard deviation ($3SD$) criterion, where data points falling more than three standard deviations above or below the mean were excluded from further analyses. Correlation analyses were then performed again to assess whether the significant relationship still held after removal of influential data points. As reading skills are highly correlated with age, multiple linear regression analyses, with age as a co-variate, were further performed to access whether latency is genuinely correlated with reading skills, independent of age. Similarly, multiple linear regression analyses with reading scores as covariates were conducted to access whether latency remains correlated with age after controlling for reading scores.

Given the well-established association between TOWRE (single word level reading fluency) and TOSREC (sentence level reading fluency and comprehension), and parallel findings linking latency to reading, we conducted mediation analyses to assess whether TOWRE mediates the relation between latency and TOSREC. A nonparametric bootstrap method with 500 simulations was used to estimate confidence intervals and p-values for the indirect effects. Standardized regression coefficients were calculated and reported to assess effect sizes.

2.7 Behavioral Analysis

Behavioral responses for the font size change detection task served as a measure of participants' attention during EEG recording. We conducted one-way ANOVAs separately for reaction time and accuracy to determine whether participants were highly engaged during the whole experiment.

3 Results

3.1 Participants were highly engaged during the whole experiment

Participants completed a font size change detection task as a means of sustaining attention during the EEG session. The mean and standard deviation (SD) of d-prime and reaction time across three conditions are summarized in Table 1. Separate one-way ANOVAs indicate significant condition differences in d-prime ($F(2, 194) = 6.02, p < 0.01$). Post-hoc t-tests revealed that size-change detection was better in words&nonwords stream compared to the words&pseudofont stream ($t = 2.26, p < 0.05$), and better in the words&pseudofont stream compared to the nonwords&pseudofonts stream ($t = 1.79, p < 0.05$). There was no significant difference across three conditions in reaction time ($F(2, 194) = 2.14, p = 0.12$).

3.2 Base stimulation elicits reliable spatial topographies and latencies across multiple harmonics

We performed RCA on responses at 3 Hz, 6 Hz, and 9 Hz in order to investigate neural activity related to processing of category non-specific visual word features with the results being summarized in Figure 2 for the first, most reliable component (RC1, for RCs 2 and 3, see Figure S2 in the Supplement). Figure 2A displays topographic visualizations of the spatial filter out of RCA pooled over stimulus conditions. RC1 was maximal over right occipito-temporal area. The bar plots in Figure 2B present amplitudes of component space data at harmonics in bar plots, with statistically significant responses in all three harmonics (all $p_{FDR} < 0.05$, corrected for 27 comparisons) for all three conditions. RSS amplitude comparisons across conditions showed that there is no significant difference between conditions ($F(2, 203) = 0.05$, $p = 0.95$). The line plots of Figure 2B display the best-fit line of phase values across three harmonics, accompanied by group-level response latencies, represented and calculated by the slope of the plot. Latencies are comparable across the three conditions (165±2.2 ms, 162±1.2 ms, and 164±2.2 ms, respectively).

	words–pseudofonts	nonwords–pseudofonts	words–nonwords
d-prime	1.70(±0.38)	1.58(±0.42)	1.82(±0.36)
RT(s)	0.62(±0.11)	0.61(±0.13)	0.58(±0.10)

Table 1: **d-prime and reaction time (RT) of the font size change task performance during EEG sessions for each of three stimulus streams.** d-prime was computed based on the z-transformed probabilities of hits and false alarms. Values are mean(±SD).

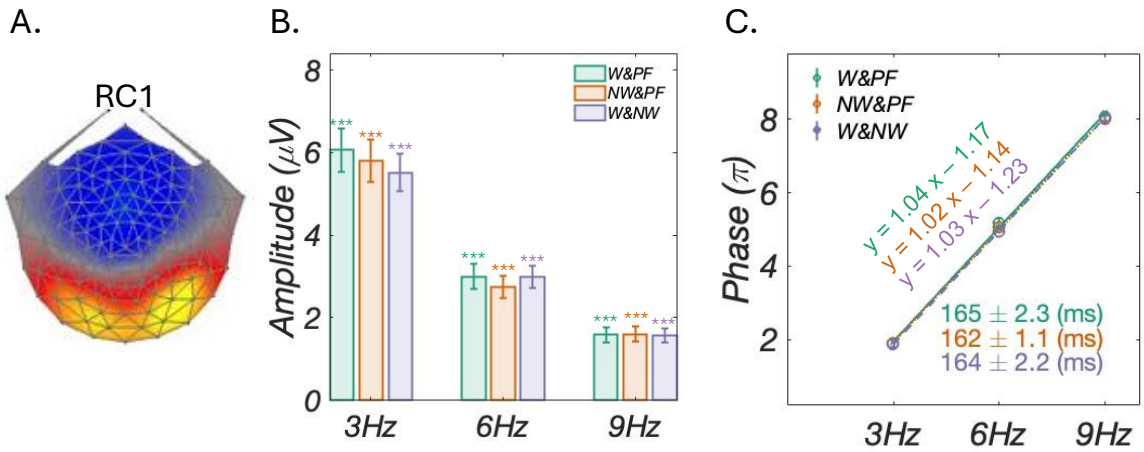

Figure 2: **RCA Analyses Results.** A: Topographic visualizations of the spatial filter for the first, most reliable components (RC1); B: Amplitude (μV) for each harmonic in bar charts, all response amplitudes are statistically significant (*: $p_{FDR} < 0.05$; **: $p_{FDR} < 0.01$; ***: $p_{FDR} < 0.001$); C: Group-level phase (π) and latencies (ms) of the three conditions. There are no significant differences across the three conditions (165±2.3 ms, 162±1.1 ms, and 164±2.2 ms, respectively).

3.3 Individual-level latencies are highly reliable

At the individual level, latencies of 64 of 68 (94%) participants could be estimated for RC1 (i.e., at least two harmonics had significant signals with the third non-significant harmonic not affecting the slope of line of best fit), allowing for further analyses. The group-level latencies (Supplement Figure S3.A) are broadly similar to the individual-level latencies (Supplement Figure S3.B), both are ranged from 150-250 ms, highly consistent with the well-known 170 ms findings in ERP field. There is no main effect of condition on the individual latencies (Supplement Figure S3.B), and individual-level latencies are significantly correlated between pairs of conditions (all $p < 0.001$ in Supplement Figure S3.C). These findings support the most promising split-half reliability test ($p < 0.001$ in Figure 3D), which demonstrated high

reliability of individual latency calculations across different stimulus onsets and various visual word forms.

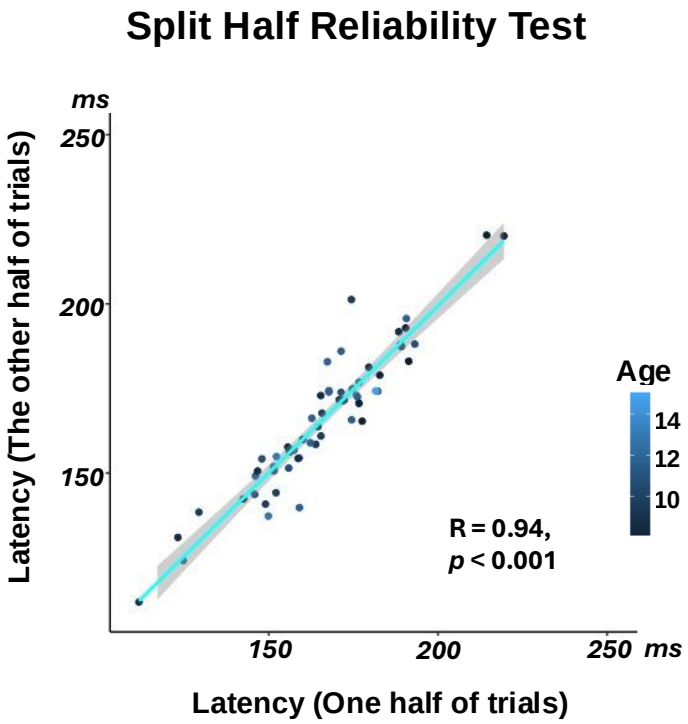

Figure 3: **Validation of individual-level latencies.** The split-half reliability test was computed to rule out potential confounds from different stimulus sets/conditions in latency calculations. Specifically, we concatenated all 30 trials (epochs/seconds) from the three conditions and randomly split them into two halves, each containing 15 trials—five randomly selected from each condition. The split-half reliability test of all 30 trials across the three conditions demonstrates high reliability, indicating the latencies represent robust signal rather than noise, and more importantly, they are stable across variations in letter string forms.

3.4 Individual latencies were negatively correlated with reading scores and age

We tested brain-behavior relationships using the estimated RC1 latencies of individual subjects in conjunction with their reading scores and age. We focus on response latencies only for base processing, as responses at base harmonics were stronger than those at oddball harmonics and most individual-level responses were statistically significant at at least two

harmonics (see Methods).

We observed negative correlations for all three comparisons (Figure 4A). Linear correlations between latency and each reading score showed that participants with better word (TOWRE) and sentence (TOSREC) reading fluency and comprehension had shorter latencies. Multiple linear regression correlations showed that the effects of TOWRE ($p < 0.001$) and TOSREC ($p < 0.01$) on latency are still statistically significant when controlling for age. In addition, the effect of age on latency is statistically significant when controlling for reading efficiency and reading comprehension (all $p < 0.05$).

3.5 Relationship between latency and sentence reading fluency is mediated by single word reading fluency

A causal mediation analysis was conducted to examine the effect of latency on sentence reading fluency (TOSREC), with single-word reading fluency (TOWRE) as the mediator. The effect of latency on the TOSREC was significantly mediated via TOWRE. The standardized regression coefficient between latency and TOSREC and the regression coefficient between TOWRE and TOSREC were significant, both $p < .05$ (Figure 4B). The indirect effect was (-0.37) $\times$ (0.54) = -0.2. We tested the significance of this indirect effect using bootstrapping procedures (500 samples). The bootstrapped indirect effect was -0.07, 95% CI [-0.15, 0.02] and was significant, $p < .05$.

4 Discussion

4.1 Latency calculations using SSVEP and RCA

Examining the latencies of spatial components can offer valuable insights into the temporal dynamics of cognitive operations, including processing speed. Unlike commonly used ERP paradigms which typically provide only group-level latencies and require long record-

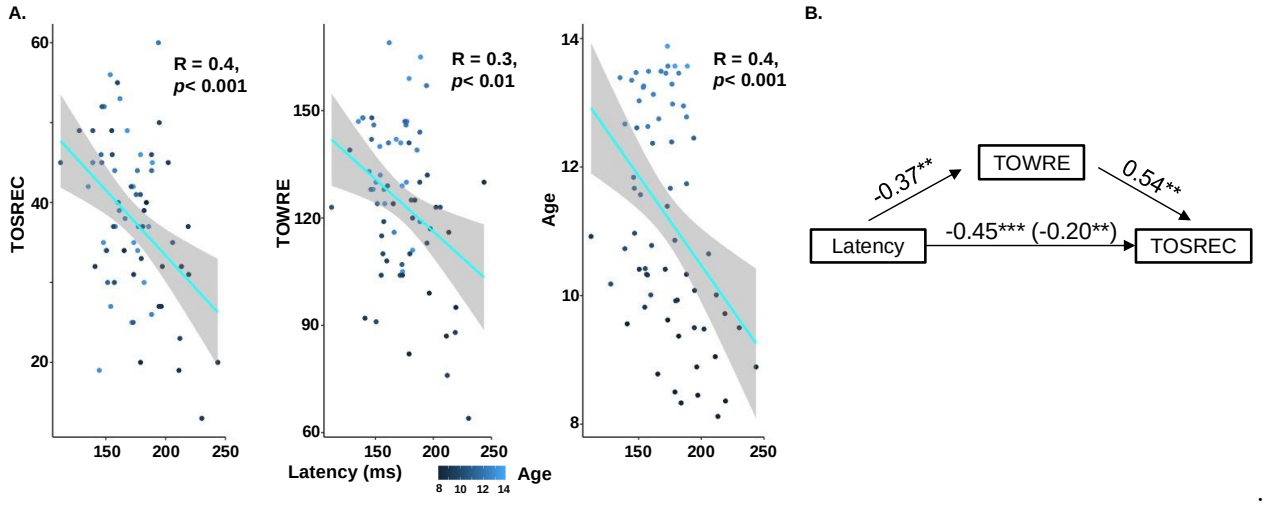

Figure 4: **Statistically significant correlations between latency, reading skills, and age, as well as a significant mediation effect.** A: Linear regressions between individual-level latency and sentence reading fluency and comprehension (TOSREC, A-left), single word reading fluency (TOWRE, A-middle), and age (A-right). Children with shorter latency had better reading fluency. Multiple linear regression correlations showed that the effects of latency on reading fluency are still statistically significant when controlling for age, but the effect of sight word efficiency is not. Meantime, older children had shorter latency. This effect of age on latency remains statistically significant after controlling for reading fluency. B: The direct relationship between latency and TOSREC was significantly mediated by an indirect pathway. Specifically, latency influenced TOWRE, which in turn affected TOSREC (see residual values in parentheses).

ing times, this study employed a SSVEP paradigm combined with an RCA approach. With its high signal-to-noise ratio, our SSVEP paradigm requires only a few minutes of stimulation and can provide not only group-level but, importantly, single-subject latencies using an estimation approach that considers phase angles across multiple SSVEP harmonics (Norcia et al., 2020). Additionally, unlike conventional ERP analyses, which usually involve one or a few preselected sensors in a cluster, RCA computes data-driven weighted linear combinations across the entire montage of sensors, thereby enhancing the feasibility of obtaining the individual-level latencies and the discovery of multiple underlying brain sources. Furthermore, by providing latencies, this study extends beyond the majority of previous SSVEP research, which has primarily focused on topographic analysis of response amplitudes.

4.2 Group-level latencies and brain sources underlying visual word form processing

Typically, SSVEP researchers use electrode responses at base frequencies as a baseline check, confirming that no significant differences exist between conditions before proceeding to more extensive analyses at the oddball frequency. In contrast with most previous SSVEP studies, including our own prior work that focused on the oddball response (Wang et al., 2021, 2022; Wang, Kaneshiro, et al., 2024; Wang, Toomarian, et al., 2024), here we closely examined the base frequency response. We find that signals at the base harmonics are more robust than those at the oddball frequencies, making them suitable for latency analysis, particularly when extended to the single-subject level. Second, using the data-driven RCA approach, we find multiple brain sources underlie base frequencies, making it promising to analyze these signals in greater detail. This is particularly significant given that the previous literature implicitly assumes the base response, being independent of stimulus category, is likely generated in primary visual areas early in the object-processing hierarchy (Lochy et al., 2016).

RCA of base streams produced three components with differing topographies and time courses (see Figure 2 and Figure S2). Here, we focus on the most robust component, RC1. The finding of brain responses to stimulus streams over occipital (RC1) areas aligns with many previous studies, including those with fMRI (Brem et al., 2009), ERP (Maurer et al., 2006), and SSVEP (van de Walle de Ghelcke et al., 2021) methodologies, which revealed preferential responses to word-like stimuli. The occipito-temporal (OT) sulcus, at a site known as the *visual word form area* that develops with reading expertise, is particularly attuned to written strings (Dehaene & Cohen, 2011). The OT region plays a pivotal role in efficient reading by receiving feedforward signals from early visual cortex and feedback projected from high-level language areas, such as angular gyrus (Lerma-Usabiaga et al., 2018).

Phase-lag quantification of RC1 provides latencies of approximately 160–170 ms. This 160–170 ms latency and the OT activations are consistent with the N170 component, a

neuropsychological marker of visual specialization for print, which has been consistently found in previous ERP studies comparing words or word-like stimuli with symbols (Brem et al., 2006; Maurer et al., 2006; Wang & Maurer, 2017).

Therefore, the presence of response components over occipito-temporal regions suggests that the base response is generated in cortical areas that ultimately encode object category. This indicates that object category information could be derived from the base stream, particularly under longer stimulus durations and slower presentation rates (Wang et al., 2022). How these non-selective responses in higher order areas are converted into category specific responses is an important question for future research.

4.3 Correlation of individual latencies of visual word form processing with reading and age

The possibility that higher-level orthographic processing could be derived from the base stream opens up opportunities to explore meaningful relationships between these distinctive neural responses and behavioral measures, including reading- and age-related effects.

The higher SNR of SSVEP signals, combined with RCA's efficiency in extracting phase-locked activities, allows us to obtain highly reliable individual latencies from the majority (94%) of participants, particularly for the most robust component, RC1. This approach enables investigation of individual differences and examination of how latency varies with children's biological age and reading abilities, providing a new direction to understand cognitive development and literacy skills.

Negative correlations were found between latencies of RC1 and reading skills, as well as age. Children with shorter latency had better single word (TOWRE) and sentence (TOSREC) reading fluency and comprehension, even after controlling for age. Additionally, older children have shorter latencies, even after reading skills were controlled. These results are consistent with several previous findings in the ERP literature of relevance. For example, cross-sectional comparison studies have shown a latency shift from childhood to

adulthood due to increased reading skills (Amora et al., 2022; Eberhard-Moscicka et al., 2016; Maurer et al., 2011). Moreover, group comparison studies between typically developing children and those with reading difficulties also provided supporting evidence. For instance, P1 latencies were found to be delayed among dyslexic as compared to normal readers (Breznitz & Meyler, 2003). In addition, abundant research have reported prolonged lower-level visual processing, which may affect further higher-level processing, in readers with dyslexia, particularly in attention demanding situation (Hari & Renvall, 2001) rather than in passive tasks (Regtvoort et al., 2006), presumably reflecting lower quality of automatised performance in readers with reading difficulties.

However, some studies have also found contradictory relationships between latency and reading. In school-aged children, some ERP studies have shown similar mean latencies for both typically developing children and those with reading difficulties or poor readers (Hasko et al., 2013; Kast et al., 2010; Maurer et al., 2011; Zhao et al., 2014). One study even found longer latencies in controls compared with children with reading difficulty (van Setten et al., 2019). This inconsistency also extends to the relationship between latency and age. For example, two studies divided a school-aged sample into young and old subgroups (Maurer et al., 2011; Tong et al., 2016): Maurer et al. (2011) found a longer latency for the younger children compared to the older ones, whereas Tong et al. (2016) reported the opposite pattern. Therefore, further studies are needed to explore the relationship between latency and behavioral measures. However, the consistent patterns across conditions observed in the present study, along with robust split-half reliability, suggest that this approach in the current study may hold promise for future research, particularly when short recording times are required.

4.4 Single word reading fluency mediates latency and sentence reading fleuncy and comprehension

We further examined the relationship between latency and sentence reading fluency and comprehension, and found that single word reading fluency mediates this relationship. Our findings provide evidence that latency predicts single word reading efficiency, which, in turn, predicts sentence reading fluency and comprehension. These results align with previous ERP findings showing a latency effect for print, where shorter latencies were observed for familiar single-word recognition in the native language compared to unfamiliar visual scripts from other language systems (Wang & Maurer, 2017, 2020). Increasing processing speed for familiar visual word form processing may reflect increased automatization, driven by reading expertise developed through education and daily exposure (McCandliss et al., 2003), which in turn predicts comprehension and development of text reading. This is an important finding, as it suggests that single-word reading fluency can be explicitly measured and potentially trained through tasks that engage automatic processing, even before formal reading acquisition and instruction begin. Future training studies could lay the groundwork for remediation efforts aimed at improving single-word reading efficiency, ultimately enhancing text reading fluency, comprehension, and overall literacy development.

4.5 Challenges and limitations

The individual-level latency estimate approach requires that the signal must meet a specific criteria—the phase-by-harmonic best-fit line must form a linear pattern of phase distribution. This constraint explains why we obtained individual-level latencies only for RC1, the most robust component, and not for other components, including those related to oddball responses.

Additionally, to examine individual differences in cortical latency, a study must include a sufficient number of participants with at least two significant harmonics (provided that

including a third non-significant harmonic does not substantially alter the slope) at the individual level for meaningful across-subject comparisons.

The stimulus streams in the current study contained mixed categories of stimuli (e.g., words, pseudofonts) rather than separate streams for each category. However, our split-half reliability test—where all trials across conditions were concatenated and randomly split into two halves—confirmed that this did not pose a problem. Nevertheless, future studies using pure stimulus streams for each category would provide more precise evidence on the neurotemporal dynamics of hierarchical processing and its relationship to reading skills.

5 Conclusion

The current study examined latencies of SSVEP base stream processing, a response measure that to date has not received much attention in SSVEP research on object and text processing. Latencies for individual participants were remarkably stable across the three stimulus streams and also highly reliable, as confirmed by split-half reliability testing. Moreover, individual-participant latencies were found to correlate significantly with reading skill and age: Children who are better readers or more developmentally advanced tended to have faster pre-categorical image-level processing speed—i.e., shorter latencies. These findings pave the way for future exploration of the developmental patterns of neural temporal dynamics in children with dyslexia and other reading or learning differences. Early assessment of these patterns could help identify children at risk for developing dyslexia before formal diagnosis, potentially serving as an early diagnostic tool.

References

Amora, K. K., Tretow, A., Verwimp, C., Tijms, J., Leppänen, P. H., & Csépe, V. (2022). Typical and atypical development of visual expertise for print as indexed by the visual word n1 (n170w): a systematic review. *Frontiers in Neuroscience*, *16*, 898800.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of statistics*, 1165–1188.

Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., & Brandeis, D. (2006). Evidence for developmental changes in the visual word processing network beyond adolescence. *Neuroimage*, *29*(3), 822–837. doi: https://doi.org/10.1016/j.neuroimage.2005.09.023

Brem, S., Halder, P., Bucher, K., Summers, P., Martin, E., & Brandeis, D. (2009). Tuning of the visual word processing system: Distinct developmental ERP and fMRI effects. *Human Brain Mapping*, *30*(6), 1833—1844. doi: https://doi.org/10.1002/hbm.20751

Breznitz, Z., & Meyler, A. (2003). Speed of lower-level auditory and visual processing as a basic factor in dyslexia: Electrophysiological evidence. *Brain and Language*, *85*(2), 166–184.

Da Silva, F. L., Van Rotterdam, A., Van Leeuwen, W. S., & Tielen, A. (1970). Dynamic characteristics of visual evoked potentials in the dog. ii. beta frequency selectivity in evoked potentials and background activity. *Electroencephalography and clinical neurophysiology*, *29*(3), 260–268.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in cognitive sciences*, *15*(6), 254–262. doi: https://doi.org/10.1016/j.tics.2011.04.003

Dmochowski, J. P., Greaves, A. S., & Norcia, A. M. (2015, April). Maximally reliable spatial filtering of steady state visual evoked potentials. *NeuroImage*, *109*, 63–72. doi: 10.1016/j.neuroimage.2014.12.078

Dmochowski, J. P., Sajda, P., Dias, J., & Parra, L. C. (2012). Correlated components of ongoing EEG point to emotionally laden attention–a possible marker of engagement? *Frontiers in human neuroscience*, *6*, 112. doi: https://doi.org/10.3389/fnhum.2012.00112

Eberhard-Moscicka, A. K., Jost, L. B., Fehlbaum, L. V., Pfenninger, S. E., & Maurer, U. (2016). Temporal dynamics of early visual word processing–early versus late N1 sensitivity in children and adults. *Neuropsychologia*, *91*, 509–518.

Hari, R., & Renvall, H. (2001). Impaired processing of rapid stimulus sequences in dyslexia. *Trends in cognitive sciences*, *5*(12), 525–532.

Hasko, S., Groth, K., Bruder, J., Bartling, J., & Schulte-Körne, G. (2013). The time course of reading processes in children with and without dyslexia: an ERP study. *Frontiers in human neuroscience*, *7*, 570.

Kast, M., Elmer, S., Jancke, L., & Meyer, M. (2010). ERP differences of pre-lexical processing between dyslexic and non-dyslexic children. *International Journal of Psychophysiology*, *77*(1), 59–69. doi: https://doi.org/10.1016/j.ijpsycho.2010.04.003

Kaufman, A. S. (1990). *Kaufman brief intelligence test: Kbit*. AGS, American Guidance Service Circle Pines, MN.

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: an 8-year follow-up. *Journal of educational psychology*, *100*(1), 150.

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and clinical neurophysiology*, *48*(6), 609–621. doi: https://doi.org/10.1016/0013-4694(80)90419-8

Lerma-Usabiaga, G., Carreiras, M., & Paz-Alonso, P. M. (2018). Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *Proceedings of the National Academy of Sciences*, *115*(42), E9981–E9990. doi: https://doi.org/10.1073/pnas.1803003115

Lochy, A., Van Reybroeck, M., & Rossion, B. (2016). Left cortical specialization for visual letter strings predicts rudimentary knowledge of letter-sound association in preschoolers. *Proceedings of the National Academy of Sciences*, *113*(30), 8544–8549. doi: https://doi.org/10.1073/pnas.1520366113

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

Masterson, J., Stuart, M., Dixon, M., & Lovejoy, S. (2010). Children's printed word database: Continuities and changes over time in children's early reading vocabulary. *British Journal of Psychology*, *101*(2), 221–242.

Maurer, U., Brem, S., Kranz, F., Bucher, K., Benz, R., Halder, P., ... Brandeis, D. (2006). Coarse neural tuning for print peaks when children learn to read. *Neuroimage*, *33*(2), 749–758. doi: https://doi.org/10.1016/j.neuroimage.2006.06.025

Maurer, U., Schulz, E., Brem, S., van der Mark, S., Bucher, K., Martin, E., & Brandeis, D. (2011). The development of print tuning in children with dyslexia: Evidence from longitudinal ERP data supported by fMRI. *Neuroimage*, *57*(3), 714–722.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in cognitive sciences*, *7*(7), 293–299. doi: https://doi.org/10.1016/S1364-6613(03)00134-7

McLean, G. M., Stuart, G. W., Coltheart, V., & Castles, A. (2011). Visual temporal processing in dyslexia and the magnocellular deficit theory: the need for speed? *Journal of Experimental Psychology: Human Perception and Performance*, *37*(6), 1957.

Norcia, A. M., Yakovleva, A., Hung, B., & Goldberg, J. L. (2020). Dynamics of contrast decrement and increment responses in human visual cortex. *Translational Vision Science & Technology*, *9*(10), 6–6. doi: https://doi.org/10.1167/tvst.9.10.6

Oldfield, R. C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Parra, L. C., Spence, C. D., Gerson, A. D., & Sajda, P. (2005). Recipes for the linear analysis of EEG. *NeuroImage*, *28*(2), 326–341. doi: https://doi.org/10.1016/j.neuroimage.2005.05.032

Pearson, K. (1896). Vii. mathematical contributions to the theory of evolution.—iii. regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*(187), 253–318. doi: https://doi.org/10.1098/rsta.1896.0007

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, *101*(2), 385–413.

Prichard, D., & Theiler, J. (1994, Aug). Generating surrogate data for time series with several simultaneously measured variables. *Phys. Rev. Lett.*, *73*, 951–954. doi: 10.1103/PhysRevLett.73.951

Regtvoort, A. G., van Leeuwen, T. H., Stoel, R. D., & van der Leij, A. (2006). Efficiency of visual information processing in children at-risk for dyslexia: habituation of single-trial erps. *Brain and Language*, *98*(3), 319–331.

Stein, J. (2023). Theories about developmental dyslexia. *Brain Sciences*, *13*(2), 208.

Tang, Y., & Norcia, A. M. (1995). An adaptive filter for steady-state evoked responses. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *96*(3), 268-277. doi: https://doi.org/10.1016/0168-5597(94)00309-3

Tlumak, A. I., Durrant, J. D., Delgado, R. E., & Boston, J. R. (2011). Steady-state analysis of auditory evoked potentials over a wide range of stimulus repetition rates: Profile in adults. *International Journal of Audiology*, *50*(7), 448–458.

Tong, X., Lo, J. C. M., McBride, C., Ho, C. S.-h., Waye, M. M. Y., Chung, K. K. H., ... Chow, B. W.-Y. (2016). Coarse and fine N1 tuning for print in younger and older Chinese children: Orthography, phonology, or semantics driven? *Neuropsychologia*, *91*, 109–119.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). TOWRE-2 examiner's manual. *Austin, TX: Pro-Ed*.

van de Walle de Ghelcke, A., Rossion, B., Schiltz, C., & Lochy, A. (2021). Developmental changes in neural letter-selectivity: A 1-year follow-up of beginning readers. *Developmental science*, *24*(1), e12999.

van Setten, E. R., Maurits, N. M., & Maassen, B. A. (2019). N1 lateralization and dyslexia: An event-related potential study in children with a familial risk of dyslexia. *Dyslexia*, *25*(1), 84–102.

Victor, J. D., & Mast, J. (1991). A new statistic for steady-state evoked potentials. *Electroencephalography and Clinical Neurophysiology*, *78*(5), 378 - 388. doi: https://doi.org/10.1016/0013-4694(91)90099-P

Vidal, C., & Chetail, F. (2017). BACS: The Brussels Artificial Character Sets for studies in cognitive psychology and neuroscience. *Behavior Research Methods*, *49*(6), 2093–2112. doi: https://doi.org/10.3758/s13428-016-0844-8

Volfart, A., Rice, G. E., Ralph, M. A. L., & Rossion, B. (2021). Implicit, automatic semantic word categorisation in the left occipito-temporal cortex as revealed by fast periodic visual stimulation. *NeuroImage*, *238*, 118228.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). *Test of silent reading efficiency and comprehension*. Pro-Ed.

Wang, F., Kaneshiro, B., Strauber, C. B., Hasak, L., Nguyen, Q. T. H., Yakovleva, A., ... McCandliss, B. D. (2021). Distinct neural sources underlying visual word form processing as revealed by steady state visual evoked potentials (SSVEP). *Scientific reports*, *11*(1), 1–15.

Wang, F., Kaneshiro, B., Toomarian, E. Y., Gosavi, R. S., Hasak, L. R., Moron, S., ... McCandliss, B. D. (2024). Progress in elementary school reading linked to growth of cortical responses to familiar letter combinations within visual word forms. *Developmental Science*, *27*(2), e13435.

Wang, F., & Maurer, U. (2017). Top-down modulation of early print-tuned neural activity in reading. *Neuropsychologia*, *102*, 29–38.

Wang, F., & Maurer, U. (2020). Interaction of top-down category-level expectation and bottom-up sensory input in early stages of visual-orthographic processing. *Neuropsychologia*, *137*, 107299.

Wang, F., Nguyen, Q. T. H., Kaneshiro, B., Hasak, L., Wang, A. M., Toomarian, E. Y., ... McCandliss, B. D. (2022). Lexical and sublexical cortical tuning for print revealed by steady-state visual evoked potentials (ssveps) in early readers. *Developmental Science*, e13352.

Wang, F., Toomarian, E. Y., Gosavi, R. S., Kaneshiro, B., Norcia, A. M., & McCandliss, B. (2024). Brain plasticity for visual words: Elementary school teachers can drive changes in weeks that rival those formed over years. *bioRxiv*, 2024–09.

Yeatman, J. D., & Norcia, A. M. (2016). Temporal tuning of word- and face-selective cortex. *Journal of cognitive neuroscience*, *28*(11), 1820–1827. doi: https://doi.org/10.1162/jocn_a_01002

Yildiz, M., Yildirim, K., Ates, S., Rasinski, T., Fitzgerald, S., & Zimmerman, B. (2014). The relationship between reading fluency and reading comprehension in fifth-grade turkish students. *International Journal of School & Educational Psychology*, *2*(1), 35–44.

Zhao, J., Kipp, K., Gaspar, C., Maurer, U., Weng, X., Mecklinger, A., & Li, S. (2014). Fine neural tuning for orthographic properties of words emerges early in children reading alphabetic script. *Journal of Cognitive Neuroscience*, *26*(11), 2431–2442.