Received: 29 October 2019 | Revised: 12 October 2020 | Accepted: 30 October 2020

DOI: 10.1111/desc.13060

Developmental Science WILEY

PAPER

Electroencephalography decoding of Chinese characters in primary school children and its prediction for word reading performance and development

2023

Kelvin F. H. Lui[1] | Jason C. M. Lo[1] | Urs Maurer[1,2] | Connie S.-H. Ho[3] | Catherine McBride[1,2]

[1]Department of Psychology, The Chinese University of Hong Kong, Hong Kong

[2]Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong

[3]Department of Psychology, The University of Hong Kong, Hong Kong

Correspondence
Catherine McBride, Room 333 Sino Building, Department of Psychology, The Chinese University of Hong Kong, Shatin, Hong Kong.
Email: cmcbride@psy.cuhk.edu.hk

Funding information
Lui Che Woo Institute of Innovative Medicine, Grant/Award Number: 8303402; Hong Kong Special Administrative Region Research Grants Council, Grant/Award Number: CUHK8/CRF/13G and C4054-17WF

Abstract

Research on what neural mechanisms facilitate word reading development in non-alphabetic scripts is relatively rare. The present study was among the first to adopt a multivariate pattern classification analysis to decode electroencephalographic signals recorded for primary school children (N = 236) while performing a Chinese character decision task. Chinese is an ideal script for studying the relationship between neural discriminability (i.e., decodability) of the orthography and behavioral word reading skills since the mapping from orthography to phonology is relatively arbitrary in Chinese. This was also among the first empirical attempts to examine the extent to which decoding performance can predict current and subsequent word reading skills using a longitudinal design. Results showed that neural activation patterns of real characters can be distinguished from activation patterns for pseudo-characters, non-characters, and random stroke combinations in both younger and older children. Topography of the transformed classifier weights revealed two distinct cognitive sub-processes underlying single character recognition, but temporal generalization analysis suggested common neural mechanisms between the distinct cognitive sub-processes. Suggestive evidence from correlational and hierarchical regression analyses showed that decoding performance, assessed on average 2 months before the year 2 behavioral testing, predicted both year 1 word reading performance and the development of word reading fluency over the year. Results demonstrate that decoding performance, one indicator of how the neural system is functionally organized in processing characters and character-like stimuli, can serve as a useful neural marker in predicting current word reading skills and the capacity to learn to read.

KEYWORDS
Chinese character, EEG, machine learning, MVPA, reading development, word reading

1 | INTRODUCTION

Word recognition is a fundamental skill to reading comprehension (Zhang et al., 2012) which is important to academic, career, and life success. The development of word reading expertise and the corresponding neural network is an important issue in developmental cognitive neuroscience. In the present study, we focused on early neural markers of Chinese character recognition. Predicting current and future behavior from brain features has been argued to be an initial and important step in neuroscience to understand how the brain gives rise to cognition and behavior (Rosenberg et al., 2018). We particularly sought to determine the extent to which neural indicators could predict behavioral performance in both timed and untimed word recognition longitudinally. There have been a few previous attempts (e.g., Leppänen et al., 2010; Maurer et al., 2006, 2009; Pugh et al., 2014; van Zuijen et al., 2013) to understand the utility of neural markers for explaining word reading longitudinally. In the present study, we adopted multivariate pattern classification analysis (MVPA), which is arguably a more sensitive method in detecting subtle but widespread effects to decode electroencephalography (EEG) signals collected from primary school children when they were performing a Chinese character decision task. Chinese is a complex script with about 5000 commonly used characters (Lee, 2000) and the mapping from orthography to phonology is relatively arbitrary in Chinese; hence, the neural system's ability to recognize the orthography of each Chinese character may be particularly important for Chinese reading development. The decoding performance, reflecting a difference between distributed but unique neural activation patterns in processing of different types of Chinese character and character-like stimuli, was our primary focus. We looked both at this performance in relation to children's current word reading skills and their timed and untimed word reading skills longitudinally, 1 year later.

Research Highlights

- Neural activation patterns of Chinese characters can be decoded and significantly distinguished from those of pseudo-characters, non-characters, and stroke combinations in children.
- The topography of transformed classifier weights showed two temporally and spatially distinctive sub-processes underlying character reading, while temporal generalization analysis suggested common mechanisms between the sub-processes.
- Decoding performance was not correlated with age, suggesting that neural discriminability of Chinese characters does not change obviously with neural maturation in this age range.
- Decoding performance tended to be associated with word reading development assessed using a longitudinal design and may be a useful neural marker to predict reading development.

1.1 | Differences in cognitive processes and the neural network between word and nonword reading

Single word reading is a complex process involving several levels of cognitive sub-processes, such as visual/orthographic recognition of words, activation of the corresponding phonological forms, and activation of the meanings of words. The classical dual-route theory of reading aloud (Coltheart et al., 1993) proposes that there are two separate cognitive routes responsible for reading of words and nonwords, respectively. The lexical route enables skilled readers to directly recognize known written words and to determine the corresponding spoken words without the phonological analysis process for the constituent graphemes. In contrast, the nonlexical route enables readers to read aloud a written stimulus, either a word or nonword, by converting the graphemes to associated phonemes through established grapheme and phoneme association rules of the language.

Research using various neuroimaging techniques has shed light on the corresponding neural network. Pugh et al. (2001) suggested a neural reading network with two left hemisphere posterior systems. The occipito-temporal (ventral) system may be responsible for the early visual word identification process, which is a fast and automatic process without heavy dependence on attentional resources. The temporo-parietal (dorsal) system, in contrast, may be responsible for an effortful phonological analysis that is slow and attention-demanding, for mapping the orthography of words onto corresponding phonological forms.

Subsequent studies have further suggested that the occipito-temporal system contains a so-called visual word form area (VWFA; Cohen et al., 2002; Sandak et al., 2004) which appears to be a critical region specialized for visual word form representation. In one study, for example, the VWFA showed stronger activation to words than to non-pronounceable consonant strings and showed stronger activation to alphabetic strings than to checkerboards (Cohen et al., 2002). A more recent meta-analysis on cross-linguistic effects in word reading (Bolger et al., 2005) suggested that the VWFA is critical to word recognition across writing systems, including English and Chinese. Taken together, findings indicate that the neural system is functionally organized for processing of words and word-like stimuli and shows different neural activation patterns in a highly distributed fashion across the brain in response to words and nonwords, regardless of the writing system.

1.2 | Temporal dynamics of word reading processes

Studies using high-temporal resolution techniques such as magnetoencephalography (MEG) and EEG have shed light on the time course of word reading. The visual word identification process responsible by the occipito-temporal system occurs approximately 150–200 ms after stimulus onset (Salmelin et al., 1996; Tarkiainen et al., 1999). The neural responses in this early stage show a preference for letter strings as compared to symbol strings and differ between fluent and dyslexic readers; these are correlated with participants' word reading speed. A subsequent event-related potential (ERP) study found that an ERP component N170, which indicates fast and automatic visual recognition process, was larger for orthographic than nonorthographic stimuli in the left hemispheric occipito-temporal sites (Bentin et al., 1999); hence, the N170 may be closely related to the visual encoding process of the orthographic form of the printed word. In addition, Bentin et al. (1999) found that an ERP component N350, distributed in the temporo-parietal areas, was elicited by phonological legal but not by phonologically illegal stimuli, and an ERP component N450, activated in areas including the fronto-central regions, distinguished not only phonologically legal and illegal words but also meaningful and meaningless words. A more recent ERP study on Chinese character and word reading (Lo et al., 2019) has also revealed a very similar time course of word reading as compared to that found in alphabetic scripts. The N170 and N400 components have been shown to be differentially activated by words and nonwords across several studies (e.g., Brem et al., 2005; Maurer et al., 2005; Sánchez-Vincitore et al., 2018 for N170; see Lau et al., 2008 for a review for N400).

These findings from electrophysiological studies are consistent with the two left hemisphere posterior systems model proposed by Pugh et al. (2001), which assumes that single word reading is comprised of an earlier (peak around 170–200 ms) orthography encoding process occurring in the occipito-temporal sites and later (peak around 400 ms) phonological and semantic analysis processes occurring in the temporo-parietal areas and fronto-central regions. Taken together, these findings suggest that the difference in neural activation patterns in response to words (or word-like stimuli) and nonwords is highly distributed both spatially and temporally.

1.3 | Development of the neural reading network and word reading expertise

The development of the occipito-temporal system is likely to be critically dependent on word reading acquisition (Shaywitz et al., 2002). Simos et al. (2001) found that children (8–15 years old) lacked a clear temporal distinction in engaging in the occipito-temporal and temporo-parietal systems, and the activations to word reading in the former system tend to be bilaterally symmetrical in children but appeared to become progressively specialized in the left hemisphere with increasing reading experience. Studies that have examined the ERP component N170 have viewed the early orthography encoding stage of word reading as a special case of perceptual expertise (e.g., Maurer & McCandliss, 2007). Two cross-sectional studies (Tong et al., 2016; Zhao et al., 2019) have showed that the N1 (how the N170 is called in children) response was different between Chinese characters, pseudo-characters, and non-characters and this N1 specialization was related to children's word reading skills. In a longitudinal study, Maurer et al. (2006) showed that larger N1 for words than symbol strings was not observed for kindergarten children who had not yet started learning to read but was found for the same children after they had mastered basic reading skills in second grade. This perceptual expertise framework has been used to account for the N170 responses that are increased by numerous classes of visual stimuli associated with perceptual expertise, such as faces (Rossion et al., 2003), birds (Tanaka & Curran, 2001), cars (Gauthier et al., 2003), and also words, even those that are not required to be read (Maurer et al., 2005).

In contrast to the perceptual expertise framework, several meta-analyses on word reading (e.g., Houdé et al., 2010; Turkeltaub et al., 2002) have suggested that adults and children (5.8–15 years old) engage very similar word reading networks, including the frontal, temporo-parietal, and occipito-temporal systems. Specifically, children have been found to engage in the visual word form area situated at the left occipito-temporal junction in word reading. According to these meta-analytic studies, the development of the occipito-temporal system is relatively independent of word reading acquisition. Moreover, Maurer et al. (2006) have found the N170 expertise effect only in contrasting words to symbol strings but have not found clear N170 specialization for words over pseudowords. Also, they did not include consonant strings, which would allow for more fine-grained contrasts of perceptual expertise in word reading (Zhao et al., 2014). Longitudinal research on this topic is scarce, especially for Chinese scripts. As suggested by a recent review (Vandermosten et al., 2016), more longitudinal studies should be conducted to examine whether the development of the occipito-temporal system depends on word reading acquisition.

In contrast, some studies have suggested that the development of the temporo-parietal system for phonological and semantic analysis occurred much earlier than the development of the occipito-temporal system and was less dependent on word reading acquisition (Friedrich & Friederici, 2006; Pugh et al., 2001; Rämä et al., 2013) while others have suggested that the system continues to be involved in word reading through adulthood (Turkeltaub et al., 2003). In addition to the dorsal–ventral model of reading acquisition, Turkeltaub et al. (2003) has further asserted that word reading development is associated with increased activity in the left inferior frontal and middle temporal areas and decreased activity in the right inferotemporal area. To conclude, research findings regarding the development of the neural reading network and its associations with the development of behavioral word reading skills have been mixed. It is not clear whether and at what age the neural mechanisms of the cognitive sub-processes underlying word reading in children can be dissociated temporally. It is also unclear whether and how much the development of each neural mechanism of the cognitive sub-processes is associated with the development of word recognition skills. Since the neural reading network and the difference in neural activation patterns in processing words and nonwords are highly distributed both spatially and temporally, the present study applied MVPA, which is arguably a more sensitive method in detecting subtle but widespread effects, to decode high-temporal resolution EEG data recorded for grade 1 to grade 5 Chinese children to examine these questions.

1.4 | A general introduction of multivariate pattern classification analysis

The application of multivariate pattern classification analysis or "brain decoding" methods to the analysis of neuroimaging data has become prevalent in the field of cognitive neuroscience (Grootswagers et al., 2017; Haynes, 2015; Pereira et al., 2009). The common practice is to test whether we can predict which experimental condition the participant is in based on their patterns of brain activation. The decoding process involves training a classifier (e.g., a support vector machine) to associate brain activation patterns with the experimental conditions using a subset of the data, and then using the trained classifier to predict the experimental conditions for new data that were not used for training. If the classification accuracy is significantly higher than chance, we can conclude that some information relevant to the experimental manipulation exists in the data.

Compared with the prevalence in fMRI, only a relatively small number of studies have applied decoding methods to the analysis of time series neuroimaging data such as MEG and EEG data. Most of them have focused on decoding the neural signals of visual perception such as face perception (Cauchoix et al., 2014; Sandberg et al., 2013), object categorization (Carlson et al., 2011, 2013; Nieuwenhuijzen et al., 2013), object recognition (Isik et al., 2014), position perception (Hogendoorn et al., 2015), communicative gestures perception (Redcay & Carlson, 2015), and auditory perception (King et al., 2014) including music perception (Schaefer et al., 2011) and speech perception (Ding & Simon, 2012). A few of them have focused on decoding the mental representation of semantic categories (Chan et al., 2011; Murphy et al., 2011; Simanova et al., 2010, 2014; Sudre et al., 2012), decision outcomes (Bode et al., 2012), and memory contents (Jafarpour et al., 2013; Wolff et al., 2015). Few studies have applied decoding methods to classify across words, pseudowords, and nonwords.

Multivariate pattern classification analysis can examine the neural activation pattern distributed across multiple time points and spatial locations simultaneously. Hence, MVPA is usually a more sensitive method in detecting subtle and widespread effects that were previously undetectable using univariate techniques. As a multivariate technique, MVPA pools together weak information across time points and electrodes. For example, Cauchoix et al. (2014) found no modulation of the two early face processing components (P100 and N170) in most of the occipitotemporal channels using traditional ERP analyses but revealed significant face category information at a very early time window (94 ms poststimulus onset) and across occipitolateral channels using MVPA. Typically, the neural activity will be normalized across features for each trial before training and testing of the classifier; therefore, MVPA targets neural activation patterns distributed across the brain but not the difference in neural activation levels in a few channels between conditions. Electrodes in isolation cannot provide as much information as their combined effects, and the electrode contributions assessed by MVPA are qualitatively different from those measured using classical univariate analyses (Cauchoix et al., 2012). In addition, by considering relationships between all features concurrently, MVPA needs not correct for multiple comparisons for the number of features to control for false positives, so it has a larger statistical power than traditional univariate techniques in analyzing the same data (Chan et al., 2011). Finally, MVPA can perform at the single trial level in individual subjects; this is better than examining the event-related potentials averaged across trials and participants when there is a certain amount of inter-individual variability across trials for the spatiotemporal activation patterns in one or more conditions (Salmelin et al., 1996).

1.5 | The present study

The present study was part of a large-scale longitudinal twin study conducted in Hong Kong examining both neural and genetic factors underlying early development of word reading skills. We applied multivariate pattern classification analysis to decode category information related to different types of character-like stimuli in single trial EEG data with a relatively large sample size. EEG activity was recorded when the children were performing a Chinese character decision task (Tong et al., 2016) in which they were required to judge whether the stimulus was a Chinese character. The children also completed a 1-min Chinese word reading task and an untimed Chinese word reading task twice across 1 year. The 1-min and untimed word reading tasks measured two components skills of word reading—word reading fluency and word reading accuracy, respectively (Cheng et al., 2017; Ho et al., 2017). Chinese is a complex script with about 5000 commonly used characters (Lee, 2000) which are usually formed from radicals consisting of multiple strokes. In addition, the mapping from orthography to phonology is relatively arbitrary in Chinese, so that orthographic awareness is important to the Chinese word reading performance (Ho et al., 2002; Lin et al., 2011). Due to these properties, the neural system's ability to differentiate Chinese characters from non-characters may be crucial for Chinese character and word reading development. As a result, Chinese is an ideal language for studying the relationship between neural discriminability (i.e., decodability) of the orthography and behavioral word reading skills.

Three types of analyses were performed. First, decoding analyses were performed to classify pairwise between EEG activity of the four types of stimuli (i.e., Chinese characters, pseudo-characters, non-characters, and stroke combinations) within each individual. Based on the previous studies which showed differential VWFA activations (Bolger et al., 2005) and N170 responses (Tong et al., 2016; Zhao et al., 2019) to Chinese characters and non-characters, we expected that the decoding performance would be significantly above chance in the orthographic recognition stage, reflecting differences in neural mechanisms in processing real characters and stimuli with different degrees of similarity to real characters. For the phonological analysis stage, the nonlexical route of the classical dual-route theory of reading aloud may not apply for Chinese non-characters since the mapping from orthography to phonology is relatively arbitrary in Chinese. However, there should still be differences in the neural responses between Chinese characters and non-characters, given that the phonological form of the former but not the latter is activated in the neural system. The decoding accuracy between real characters and stroke combinations, hence, was expected to be the highest since the two conditions differed dramatically in both orthographic and phonological forms. The topography of the classifier weights was plotted across time to examine how the contributing features were distributed temporally and spatially. The decoding performance and topography were compared between younger and older children to examine the age effects on the neural mechanisms of the cognitive sub-processes underlying word reading.

Second, decoding analyses were performed using temporal generalization by training the classifier on a particular time window and then testing it in different time windows. If the classifier trained on one time window can successfully predict the types of stimuli for data at other time windows, the neural activation patterns for the stimuli should be similar between the trained and tested time windows, suggesting the same cognitive sub-process across time. In contrast, if the temporal generalization performance is not significantly above chance, the neural activation patterns for the stimuli should change between time windows, suggesting different cognitive sub-processes across time. The temporal generalization analysis can provide insights about whether the cognitive sub-processes underlying single character recognition can be temporally disassociated in primary school children.

Third, correlations and regression analyses were performed to test whether decoding performance could serve as a neural marker in explaining the development of word reading fluency and accuracy. In the hierarchical regression analyses, the individual differences in word recognition skills in the previous year was statistically controlled to ensure that the additional variance explained by decoding performance reflects the extent to which the development of word read skills over the past year is related to the neural system's ability to differentiate real Chinese characters and different types of character-like stimuli. Since word reading fluency and accuracy are two different components skills of word reading, their associations with neural discriminability of Chinese characters are possibly different. Given that word reading fluency indicates automatization of reading processes to a larger degree, while word reading accuracy indicates more controlled aspects of reading, we expected that word reading accuracy would show relatively stronger associations with neural discriminability between different types of character-like stimuli that are more similar to each other (e.g., real characters vs. pseudo characters) than those that are more distinct (e.g., real characters vs. stroke combinations). Still, overall associations between neural discriminability and reading might be strongest for reading fluency given previous studies that focused on N1 print tuning and reading (Maurer et al., 2007; Tong et al., 2016).

2 | METHOD

2.1 | Participants

One hundred twenty-one pairs of Chinese twins from grades 1 to 5 participated in the current study in the first year. One child did not complete all the tasks and two children scored on fewer than 10 trials for at least one condition of the Chinese character decision task after EEG data preprocessing. The three pairs of twins involved were hence discarded from analyses. The remaining 236 children included 108 males and 128 females. They were 5.39–9.76 years old (M = 7.36, SD = 0.92) in the first year and were 7.08–10.17 years old (M = 8.28, SD = 0.82) when they completed the EEG testing. Seven twin pairs quit the study in the second year. The remaining 222 children included 103 males and 119 females and were 6.69–10.99 years old (M = 8.44, SD = 0.98) in the second year. The children were divided into two age groups for analyses based on their year 2 age: (a) 106 younger group children from 6.67 to 8 years old (M = 7.61, SD = 0.37, 51 males); (b) 116 older group children from 8.08 to 10.92 (M = 9.11, SD = 0.80, 52 males). All children were native Cantonese speakers, not previously diagnosed as having developmental dyslexia, and had normal or corrected to normal visual ability. Informed consent was obtained in written form from the parents. The study protocol was approved by the Survey and Behavioral Research Ethics Committee of the Chinese University of Hong Kong (Ref. CUHK8/CRF/13G/2300035) and the Joint Chinese University of Hong Kong-New Territories East Cluster Clinical Research Ethics Committee (Ref. 2017.479).

2.2 | Procedure

Participants completed the two behavioral sessions, both including a 1-min Chinese word reading task and an untimed Chinese word reading task, across the period of 1 year. The behavioral session completed in the first year is referred to as wave 1 and the behavioral session completed in the second year is referred to as wave 2 throughout this paper. The children completed the behavioral measures either in their home or their school. Participants completed an EEG session for a Chinese character decision task scheduled on average 2 months before the wave 2 behavioral testing. For the EEG session, the children were individually tested in a sound-attenuated laboratory in the Chinese University of Hong Kong. The Chinese character decision task was presented using an E-prime program and the EEG activity was collected using the HydroCel GSN EGI 128-channel system (EGI net station; Electrical Geodesics Inc.).

2.3 | One-minute Chinese word reading task

In this task, a list of 90 Chinese two-character words was presented to the children. They were asked to read aloud the words one by one as accurately and as quickly as possible. They were instructed to read the next word when they did not know how to read a word. The number of words they read correctly within 1 mi served as the indicator of their word reading fluency.

2.4 | Untimed Chinese word reading task

In this task, a list of 150 Chinese two-character words was presented to the children. They were asked to read aloud the words one by one as accurately as possible. The testing stopped when the children had made 15 consecutive errors or had finished reading all the words. The total number of words they read correctly served as the indicator of their word reading accuracy.

2.5 | Chinese character decision task

Participants performed a Chinese character decision task during the EEG recording. The task stimuli were presented via an E-prime program. In each trial, a fixation cross was presented on the screen for a random interval varying from 400 to 600 ms followed by the presentation of the target stimulus for 1500 ms. Participants were seated at a distance of 80 cm from the computer monitor and were required to judge whether the stimuli were real characters or not as fast and as accurately as possible by pressing button "1" or "4," respectively, on an response box. Responses had to be made within the 1500 ms stimulus presentation time window or otherwise considered as incorrect. A blank screen then appeared for 1000 ms; this served as the inter-trial interval.

The stimulus for each trial was randomly selected from four types of stimuli including real Chinese characters, pseudo-characters, non-characters, and random stroke combinations. Figure 1 shows examples of the four types of stimuli and an example trial. All real characters were left right compound characters with a semantic radical on the left and a phonetic radical on the right. All real characters were selected from a published wordlist in Hong Kong and were typically learned from grade 1 to grade 3 (Chinese Language Education Section, 2009). The pseudo-characters were created by combining a semantic and a phonetic radical following the correct orthographic rules in Chinese while the non-characters were created by reversing the positions of the semantic and phonetic radicals of a real character. The stroke combinations, adopted from Su et al. (2015), were composed of two non-existing radicals that were formed by randomized strokes. The total number of strokes of the four types of stimuli was matched to control for the visual complexity across conditions.

Participants performed a practice block of 10 trials followed by performing 6 test blocks with 40 trials each. The number of stimuli of each condition was equal within each block, which resulted in 60 trials for each condition.

2.6 | EEG recording and preprocessing

Electroencephalography was recorded using the HydroCel GSN EGI 128-channel system (EGI net station; Electrical Geodesics Inc.) at a sampling rate of 500 Hz and with the Cz electrode as the online reference. Electrode impedance levels were set at less than 50 kΩ. There was a 12 ms timing delay between the event trigger and the actual stimulus onset created by the amplifier's internal anti-aliasing filter which was corrected in the following preprocessing.

Preprocessing steps were done using EEGLAB v14.1.2 (Delorme & Makeig, 2004). Continuous EEG data were first filtered with a 0.3–30 Hz band-pass filter and then downsampled to 250 Hz. Only channels from the standard EEG montage 10–10 system were included in the further analyses. Bad channels were further removed with the PREP pipeline (Bigdely-Shamlo et al., 2015). Independent

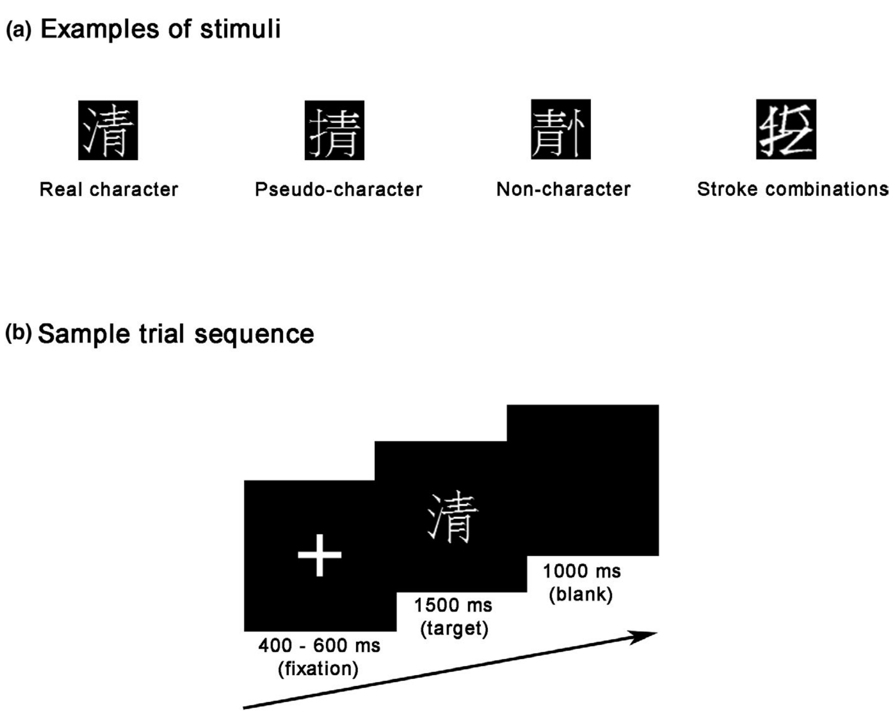

FIGURE 1 Examples of target stimuli for each experimental condition (a) and a sample trial sequence of the real Chinese character condition (b)

component analysis (ICA) was then performed on the continuous EEG data with an optimization algorithm – CUDAICA (Raimondo et al., 2012). Components related to eye movement artifacts were then removed using ADJUST (Mognon et al., 2011). Trials were then epoched from 150 ms pre-stimulus onset to 850 ms post-stimulus onset. Epochs with an absolute amplitude larger than 80 μV were removed from further analyses. On average, 28 epochs (7 per condition) were removed for each participant. Finally, potentials were referenced to the common average and baseline corrected using the pre-stimulus interval.

2.7 | Multivariate pattern classification analyses

After the preprocessing steps, trials with correct responses were subjected to multivariate pattern classification analyses using the linear support vector machine (SVM) classifier. The SVM classifier was trained to distinguish between any two types of stimuli (pairwise) among the four using a 10-fold cross-validation. The data were randomly partitioned into 10 portions with an equal number of trials. Each time the classifier was trained using 90% of trials and then tested on the remaining 10% of trials. This training and testing process was repeated 10 times with each portion of data being used in training for nine times and in testing for 1 time. To improve the signal-to-noise ratio and examine the time course of the classification accuracy, a sliding time window approach was adopted to classify across five time points (20 ms) simultaneously. The step size of the moving time window was two time points (8 ms) which resulted in three overlapping time points between adjacent time windows. Whether the classification accuracy of each time window was significantly above chance (50%) was tested by the Student's t-test with subjects as a random factor. To conform to the independent observation assumption in the general linear model, only one child's datum from each twin pair was randomly selected for the t-test. However, to minimize the noise caused by random selection, we repeated the selection process and performed the t-test for 1000 times and averaged the 1000 t-statistics as the final t-statistic. Multiple comparisons across time windows were corrected by the threshold-free cluster enhancement method (TFCE; Smith & Nichols, 2009). Two types of decoding analyses were performed including (a) the usual decoding analysis using data from the same time window in training and testing and (b) a temporal generalization decoding analysis using data from one time window in training and data from another time window in testing.

2.8 | Regression analyses

Two hierarchical regression analyses were performed to examine whether decoding performance could predict the development of word reading fluency and accuracy, respectively. The regression analyses treated the wave 2 word reading performance as the dependent variable, and then included age in Block 1, wave 1 word reading performance in Block 2, and then further all pairwise decoding performance (neural markers) in Block 3 to examine whether early character sensitivity at time 1 could predict unique variance in word reading performance after controlling for age and individual difference in word reading performance over 1 year. Both the change in explained variance from Block 2 to Block 3 and p-values for individual predictors were examined. As there were two word reading measures and six decoding predictors, the Bonferroni-adjusted p-values were also examined.

3 | RESULTS

3.1 | Behavioral results

3.1.1 | Chinese character decision task

Table 1 shows the mean accuracy and RT for trials with correct responses for each condition and age group combination. Two two-way mixed model ANOVAs with the experimental condition as the within-subject factor and age group as the between-subject factor were performed on accuracy and RT separately. To conform to the independent observation assumption in the general linear model, only one child's datum from each twin pair was randomly selected for the analyses. The main effect of condition was significant for both accuracy, $F(3,107) = 228.62$, $p < 0.001$, $\eta_p^2 = 0.865$, and RT, $F(3,107) = 80.09$, $p < 0.001$, $\eta_p^2 = 0.692$. Follow-up contrasts showed significantly higher accuracy and faster RT in the stroke combinations condition than all other conditions and found significantly lower accuracy and slower RT in the pseudo-characters

TABLE 1 Mean accuracy and RT for trials with correct responses for the Chinese character decision task for the two age groups

Stimuli category	Accuracy (younger)	Accuracy (older)	RT (younger)	RT (older)
Real characters	0.85 (0.11)	0.85 (0.13)	814.3 (113.3)	745.3 (114.1)
Pseudo-characters	0.48 (0.27)	0.56 (0.26)	887.0 (165.2)	797.2 (153.1)
Non-characters	0.84 (0.10)	0.88 (0.10)	846.1 (140.0)	749.5 (137.4)
Stroke combinations	0.93 (0.10)	0.95 (0.07)	764.1 (115.0)	681.8 (107.3)

Note: Younger group children were 6.67–8 years old and older group children were 8.08–10.92 years old in completing the year 2 testing. Numbers inside the parentheses showed the standard deviations.

TABLE 2 Amount of words correctly read in the two word reading tasks for the two age groups

	Wave 1		Wave 2		Improvement		
Task	Younger group	Older group	Younger group	Older group	Younger group	Older group	Combined
COM	42.7 (18.1)	60.2 (18.9)	58.5 (20.1)	72.0 (20.4)	15.8 (14.3)	11.8 (13.6)	13.7 (14.0)
CWR	50.3 (27.1)	85.5 (31.9)	81.6 (26.5)	105.5 (25.1)	31.3 (15.5)	20.7 (12.0)	25.8 (14.8)

Note: Numbers inside the parentheses showed the standard deviations.

Abbreviations: COM, Chinese one-minute word reading; CWR, Chinese untimed word reading.

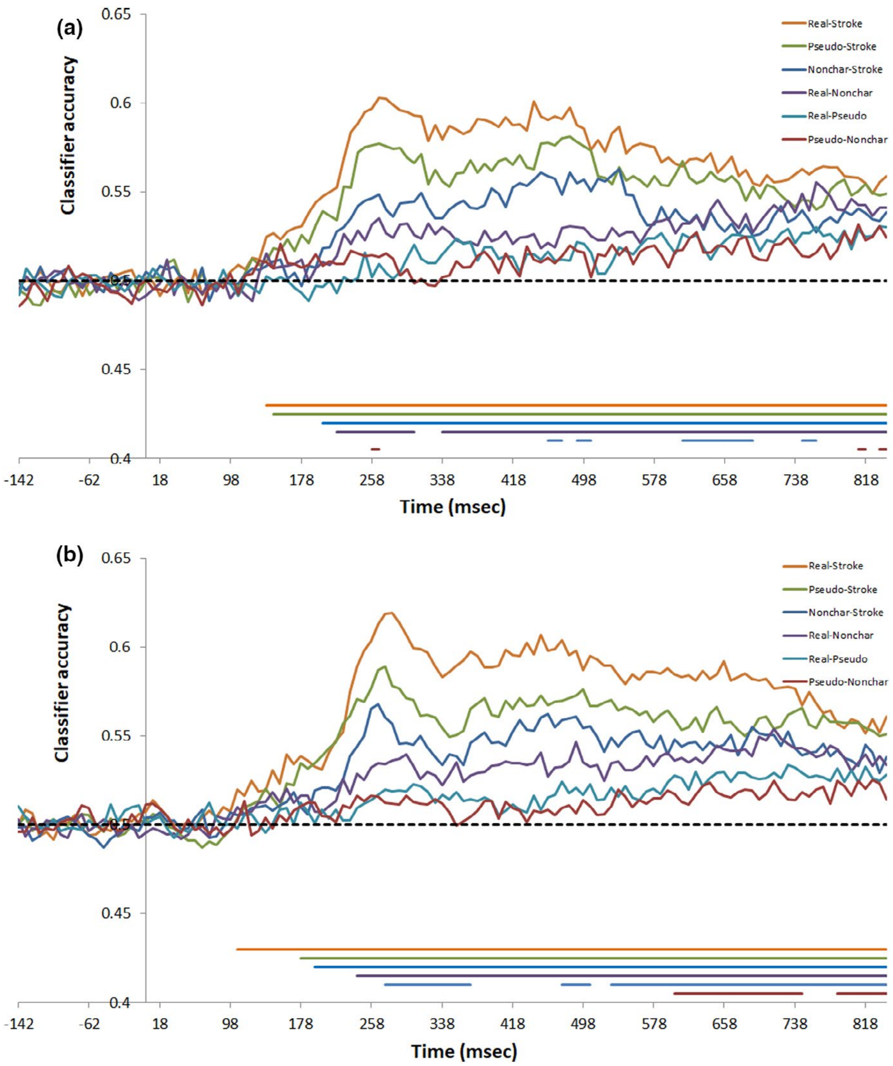

FIGURE 2 Within-individual classification accuracies of all pairwise combinations of the four stimuli categories (a: younger children; b: older children). Discs above the *x*-axis indicated the time points where classification accuracies were significantly above chance. Appendix D shows the classification accuracy plots faceted by age group

condition than all other conditions for both groups, all *p*s < 0.001. The main effect of age was marginally significant for accuracy, $F(1,109) = 3.31$, $p = 0.072$, $\eta_p^2 = 0.03$, and significant for RT, $F(1,109) = 13.53$, $p < 0.001$, $\eta_p^2 = 0.11$, with higher accuracy and faster RT in the older group for all experimental conditions. The Condition × Group interaction was not significant for either accuracy $F(3,107) = 0.89$, $p = 0.449$, $\eta_p^2 = 0.024$, or RT $F(3,107) = 0.93$, $p = 0.431$, $\eta_p^2 = 0.03$.

3.1.2 | Chinese word reading performance

Table 2 shows the number of words the children read correctly in both word reading tasks for both age groups. Two two-way mixed model ANOVAs with Time (wave 1 vs. wave 2) as the within-subject factor and age group as the between-subject factor were performed on the two word reading tasks separately. The main effect of Time was significant for both 1-min, $F(1,109) = 109.06$,

$p < 0.001$, $\eta_p^2 = 0.500$, and untimed word reading performance, $F(1,108) = 389.21$, $p < 0.001$, $\eta_p^2 = 0.783$ with significantly better performance in wave 2 than wave 1. The main effect of age was also significant for both 1-min, $F(1,109) = 20.25$, $p < 0.001$, $\eta_p^2 = 0.157$, and untimed word reading performance, $F(1,108) = 34.20$, $p < 0.001$, $\eta_p^2 = 0.240$, with a significantly better performance for older than younger children. The Time × Group interaction was not significant for the 1-min word reading task $F(1,109) = 2.19$, $p = 0.142$, $\eta_p^2 = 0.020$, but significant for the untimed word reading task $F(1,108) = 16.13$, $p < 0.001$, $\eta_p^2 = 0.130$, with a larger improvement of word reading performance from wave 1 to wave 2 for the younger than the older group.

3.2 | Within-individual decoding results

Figure 2 shows the decoding accuracies for all pairwise classifications for the two age groups, respectively. As shown, for both age groups, all the six pairwise classifications yielded significantly above chance (0.50) accuracies at certain time points across the trial (*ps* < 0.05). The same decoding accuracy plots were also organized by the pairwise classifications and presented in Appendix D. To examine the effects of stimuli and age on the classification performance, a two-way mixed model ANOVA with Stimuli (six pairwise classification analyses) as the within-subject factor and age group as the between-subject factor was performed. The main effect of Stimuli was significant, $F(5,105) = 49.16$, $p < 0.001$, $\eta_p^2 = 0.701$, with the highest classification accuracy found in classifying between real characters and stroke combinations and the lowest classification accuracy found in classifying between pseudo-characters and non-characters. In general, classification analyses involving the stroke combinations showed higher accuracies as compared to those not involving the stroke combinations. Neither the main effect of age nor the interaction effect were significant, $F(1,109) = 0.929$, $p = 0.337$, $\eta_p^2 = 0.008$, and $F(5,105) = 0.473$, $p = 0.796$, $\eta_p^2 = 0.022$, respectively. As shown in Table 3, follow-up *t*-tests revealed no significant age group differences for any of the pairwise classification analyses.

TABLE 3 *T*-tests results in comparing the decoding performance between the two age groups

	Accuracy (young)	Accuracy (old)	*t*-score	*p*-value	Cohen's *d*
Real-Stroke	0.5604 (0.0297)	0.5670 (0.0367)	1.04	0.301	0.20
Pseudo-Stroke	0.5517 (0.0355)	0.5509 (0.0343)	−0.12	0.904	0.02
Nonchar-Stroke	0.5347 (0.0263)	0.5383 (0.0238)	0.749	0.455	0.14
Real-Nonchar	0.5230 (0.0287)	0.5277 (0.0239)	0.929	0.355	0.18
Real-Pseudo	0.5144 (0.0222)	0.5154 (0.0175)	0.229	0.819	0.05
Pseudo-Nonchar	0.5092 (0.0209)	0.5124 (0.0221)	0.800	0.426	0.15

Note: Decoding performance of each pairwise classification was computed by averaging the classification accuracies across all post-stimulus time windows. Numbers inside the parentheses showed the standard deviations. Real represents the real character condition, Pseudo represents the pseudo-character condition, Nonchar represents the Non-character conditions, Stroke represents the random stroke combinations condition.

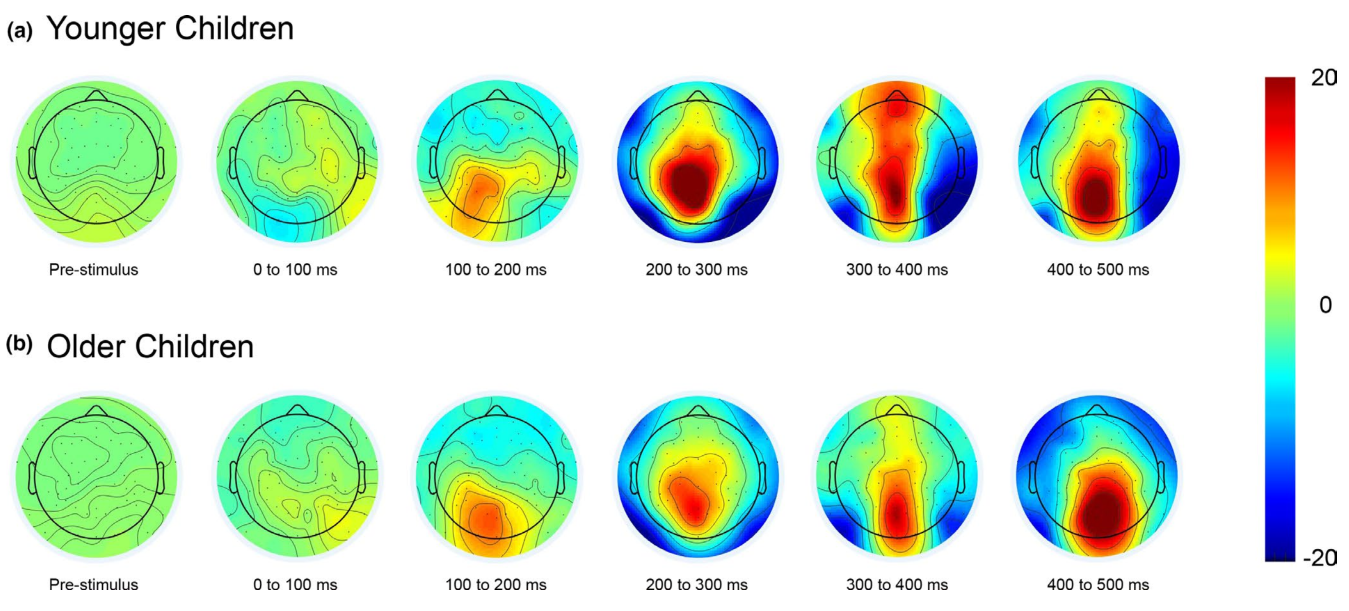

FIGURE 3 Topography of the transformed classifier weights in classifying between real characters and pseudo-characters across time windows for both the younger children (a) and older children (b). For each topography, transformed classifier weights were averaged across time points within the time window

3.3 | Topography and temporal generalization decoding results

The topography of the transformed classifier weights in classifying between real characters and pseudo-characters was examined to see if the underlying neural processes of Chinese character recognition were qualitatively different between the younger and older children. As shown in Figure 3, both groups showed left-lateralized activation patterns around the occipito-temporal channels starting at the 100–200 ms time window. The activation patterns then appeared to shift to the parietal and fronto-central channels at the 200–300 ms time window and finally became bi-lateralized at 300–500 ms time windows. As the two age groups did not differ significantly in most of the decoding analyses and showed qualitatively the same pattern of topography, we combined the two age groups in the subsequent analyses.

The temporal generalization decoding analysis was performed to classify between real characters and stroke combinations. The unified and larger cluster in Figure 4b suggested that decoding performance for the sub-processes reflected certain common neural mechanisms. As a result, for the correlational and regression analyses, the decoding performance for each individual was calculated by averaging the classification accuracies across all post-stimulus time windows. More details of the topography and temporal generalization analyses were included in Appendices A and B, respectively.

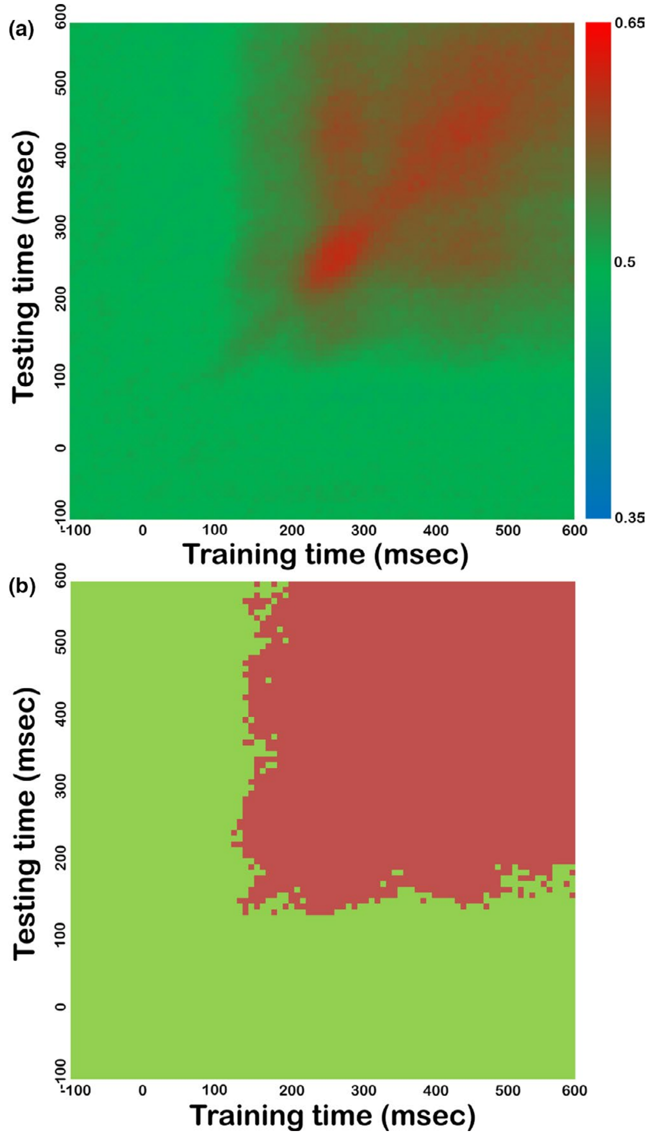

FIGURE 4 (a) Temporal generalization classification accuracies between real characters and stroke combinations. (b) Statistical significance of the temporal generalization classification accuracies. Red color indicates significantly above chance accuracy, after controlling for multiple comparisons using the Bonferroni correction

3.4 | Correlations among decoding performance, age, and reading skills

In order to consider the longitudinal associations among the neural and behavioral markers of literacy, we first examined the correlations among wave 2 age, decoding performance, and both wave 1 and wave 2 one-minute and untimed word reading performance. As shown in Table 4, the word reading skills were all significantly and positively correlated with each other, *rs* > 0.69, *ps* < 0.001, and the decoding performances were in general significantly and positively correlated among themselves (9 correlations out of 15 were significant). More importantly, both wave 2 one-minute and untimed word reading performance were significantly and positively correlated with some decoding performances, suggesting that better neural indicators of decoding performance tend to be associated with better behavioral word reading skills. More details of the correlation analyses were included in Appendix C.

3.5 | Decoding performance predicts improvement in word reading performance

Two hierarchical multiple regression analyses were performed separately to explain variability in 1-min word reading and the untimed word reading performance at wave 2 after controlling for age and wave 1 word reading performance. As there were many significant correlations among the decoding performances, variance inflation factors (VIFs) of the six decoding performances were computed to examine the multicollinearity issue. Multicollinearity among predictors are typically considered to be acceptable, if the VIFs are smaller than 5. As the VIFs of the six predictors ranged from 1.23 to 1.82, we conclude that multicollinearity is not a problem in our analysis. Table 5 shows the results. For word reading fluency, decoding performance additionally explained 4.8% variance, $\Delta R^2 = 0.048$, $F(6,102) = 2.316$, $p = 0.039$. Specifically, decoding performance in classifying between pseudo-characters and stroke combinations significantly predicted word reading fluency, $\beta = 0.219$, $t = 2.794$, $p = 0.006$. For word reading accuracy, decoding performance additionally explained 1.8% variance; the statistical test was marginally significant, $\Delta R^2 = 0.018$,

TABLE 4 Correlations among age, decoding performance, and reading skills

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
V1. Age (Wave 2)	–	0.08	0.21*	0.11	0.08	0.09	0.08	0.54**	0.66**	0.39**	0.56**
V2. Real-Pseudo		–	0.38**	0.22*	0.27**	0.09	0.05	0.18	0.14	0.19*	0.13
V3. Real-Nonchar		0.37**	–	0.30**	0.32**	0.08	0.04	0.28**	0.21*	0.24*	0.27**
V4. Real-Stroke		0.23*	0.29**	–	0.13	0.60**	0.35**	0.26**	0.17	0.27**	0.19*
V5. Pseudo-Nonchar		0.27**	0.31**	0.12	–	0.19*	0.03	0.01	0.03	0.13	0.07
V6. Pseudo-Stroke		0.09	0.07	0.59**	0.18	–	0.40**	0.24*	0.13	0.36**	0.20*
V7. Nonchar-Stroke		0.04	0.02	0.35**	0.02	0.41**	–	0.13	0.13	0.09	0.13
V8. Wave 1_COM		0.17	0.21*	0.24*	−0.04	0.22*	0.10	–	0.83**	0.78**	0.81**
V9. Wave 1_CWR		0.12	0.10	0.13	−0.03	0.10	0.11	0.75**	–	0.69**	0.91**
V10. Wave 2_COM		0.17	0.17	0.26**	0.11	0.35**	0.06	0.73**	0.63**	–	0.74**
V11. Wave 2_CWR		0.09	0.18	0.17	0.04	0.19*	0.10	0.73**	0.86**	0.69**	–

Note: Values presented above the diagonal are Pearson's correlation coefficients without controlling for age and values presented below the diagonal are partial correlation coefficients controlling for age. Real represents the real character condition, Pseudo represents the pseudo-character condition, Nonchar represents the Non-character conditions, Stroke represents the random stroke combinations condition.

Abbreviations: COM, Chinese one-minute word reading; CWR, Chinese untimed word reading.

*$p < .05$,

**$p < .01$.

$F(6,102) = 2.316$, $p = 0.080$. As two regression analyses were performed and each included six decoding predictors, we also examined the Bonferroni-adjusted p-values to control for the overall type I error rate. After adjusting the p-values for the two regression analyses, decoding performance marginally significantly predicted word reading fluency, $\Delta R^2 = 0.048$, $F(6,102) = 2.316$, $p = 0.078$, but was not significantly associated with word reading accuracy $\Delta R^2 = 0.018$, $F(6,102) = 2.316$, $p = 0.160$. After adjusting the p-values for the six predictors, the decoding performance in classifying between pseudo-characters and stroke combinations was still significantly associated with word reading fluency, $\beta = 0.219$, $t = 2.794$, $p = 0.036$. Taken together, there is suggestive evidence that a larger difference in the neural system in reacting to different types of character-like stimuli was associated with better learning and improvement on word reading fluency, but not word reading accuracy, across 1 year in children.

4 | DISCUSSION

4.1 | Decodability of EEG signals of Chinese characters in children

The present study adopted multivariate pattern classification analysis to decode EEG signals recorded for children in grades 1–5 while performing a Chinese character decision task. Within-individual decoding analysis showed that the neural activation patterns in children's early processing of real characters could be significantly distinguished from their processing of pseudo-character, non-character, and stroke combinations for both younger children (6.67–8 years old) and older children (8.08–10.92 years old). The decoding performance between real characters and stroke combinations was highest in general and had the earliest onset time for significantly above chance performance. For older children, the onset time for significant performance in classifying between real characters and stroke combinations occurred as early as 100 ms after stimulus presentation which was shortly after visual processing had begun. In addition, although the orthographic forms between real characters and pseudo-characters are highly similar, the neural activation patterns in response to the two could still be distinguished in the decoding analysis, suggesting that the EEG signals of single Chinese characters and other character-like stimuli were decodable in children. Thus, MVPA is a powerful technique in decoding the neural activation patterns of Chinese characters at the single trial level. In group comparisons, older and younger children did not show significant differences in their performances across all the pairwise classification analyses. In the correlation analyses, age was only weakly correlated with the decoding performance in classifying between real characters and non-characters. It seems that the neural discriminability of Chinese characters and character-like stimuli does not increase obviously as age increases. This is probably because the characters used had been selected to be relatively easy to both younger and older children (curriculum specified they were to have been learned by grade 3). As shown in the behavioral results, the difference in performance between the younger and older children was small (<0.1 accuracy for all conditions). Taken together, this may suggest that neural discriminability is more related to word reading skills than to neural maturation in this age range.

4.2 | Distinct and common neural mechanisms between the cognitive sub-processes

The topography of the transformed classifier weights in classifying between real characters and pseudo-characters showed clear

TABLE 5 Results of multiple regression on wave 2 reading skills

	One-minute word reading							Untimed word reading						
Variables	ΔR^2	*B*	*SE*	*B*	*t/F*	*p-value*	Adjusted *p*-value	ΔR^2	*B*	*SE*	*B*	*t/F*	*p-value*	Adjusted *p*-value
Block 1	0.153				19.7	<0.001		0.314				49.6	<0.001	
(Intercept)		−6.0	16.2		−0.37	0.712			−41.7	19.5		−02.1	0.035	
Age		0.71	0.16	0.39	4.4	<0.001			1.4	0.19	0.56	7.0	<0.001	
Block 2	0.450				122.4	<0.001		0.512				315.2	<0.001	
(Intercept)		29.2	11.6		2.5	0.014			56.1	11.3		5.0	<0.001	
Age		−0.07	0.13	−0.04	−0.50	0.620			−0.15	0.13	−0.06	−1.2	0.247	
W1 COM/CWR		0.83	0.08	0.80	11.1	<0.001			0.78	0.04	0.95	17.8	<0.001	
Block 3	0.048				2.32	0.039	0.078	0.018				1.95	0.080	0.160
(Intercept)		−47.5	48.6		−0.98	0.331			12.8	43.5		0.29	0.770	
Age		−0.06	0.13	−0.03	−0.47	0.642			−0.19	0.13	−0.08	−1.5	0.131	
W1 COM/CWR		0.79	0.08	0.76	10.2	<0.001			0.78	0.04	0.94	17.7	<0.001	
Real-Pseudo		32.5	69.9	0.03	0.47	0.643	1		−84.2	62.8	−0.06	−1.3	0.183	1
Real-Nonchar		−5.3	56.5	−0.01	−0.10	0.925	1		106.9	49.7	0.10	2.2	0.034	0.204
Real-Stroke		−24.7	49.7	−0.04	−0.50	0.620	1		−20.9	44.7	−0.03	−0.47	0.642	1
Pseudo-Nonchar		83.9	64.1	0.09	1.3	0.194	1		30.1	56.6	0.02	0.53	0.596	1
Pseudo-Stroke		134.2	48.1	0.22	2.8	0.006	0.036		87.7	42.3	0.11	2.1	0.041	0.246
Nonchar-Stroke		−72.2	55.2	−0.09	−1.3	0.194	1		−31.1	49.6	−0.03	−0.63	0.532	1

Note: Decoding performance of each pairwise classification was computed by averaging the classification accuracies across all post-stimulus time windows. Real represents the real character condition, Pseudo represents the pseudo-character condition, Nonchar represents the Non-character conditions, Stroke represents the random stroke combinations condition.

Abbreviations: COM, Chinese one-minute word reading; CWR, Chinese untimed word reading.

left-lateralized neural activation patterns at the occipito-temporal channels at the 100–300 ms time windows and bi-lateralized activation patterns at the parietal and fronto-central channels at the 300–500 ms time windows for both age groups, suggesting two distinct cognitive sub-processes underlying single character reading. Consistently, the classification accuracy (Figures 2 and 4a) also showed two peaks with peak latency at about 250 and 450 ms, respectively. These two apparent sub-processes underlying single Chinese character reading were highly consistent with the interpretation of an orthographic encoding process occurring earlier (peak around 170–200 ms) in the occipito-temporal sites and the phonological and semantic analysis processes occurring later (peak around 400 ms) in the temporo-parietal and fronto-central regions, as suggested in some previous studies (e.g., Lau et al., 2008; Maurer et al., 2005; Pugh et al., 2001). However, the temporal generalization analysis did not obviously disassociate the two cognitive sub-processes underlying single character recognition as generalization from one sub-process to the another was also statistically significant, resulting in a large cluster of significant decoding performance (Figure 4b). Results suggest that there should be some shared neural mechanisms between the cognitive sub-processes. It is likely that the phonological and semantic analysis processes are required in order to activate the orthographic forms of characters before one can map the orthographic form onto the corresponding phonological and semantic forms.

4.3 | Potential mechanisms underlying the decoding performance

The decoding performance was found to be significantly above chance, indicating that there is category-level information related to each type of stimulus in the neural activation patterns which allows the MVPA algorithm to differentiate across different experimental conditions. In other words, different neural activation patterns resulted in processing of different types of stimuli, and better decoding performance implies larger differences in the neural activation patterns between conditions. However, what are the exact neural mechanisms or characteristics captured by the decoding performance? One possibility is that the decoding performance in the current study reflects the extent to which the brain is functionally organized in processing stimuli with different degrees of similarity to real Chinese characters. According to the dual-route theory (Coltheart et al., 1993), the dorsal–ventral reading network (Pugh et al., 2001), and findings from previous neuroimaging studies (e.g., Cohen et al., 2002; Lau et al., 2008; Maurer et al., 2005), the human neural system shows different neural mechanisms in various cognitive sub-processes, including the orthographic encoding process and the phonological and semantic analysis processes, in processing words, pseudo-words, and nonwords. In this sense, better decoding performance may reflect larger differences in these neural processes.

Alternatively, decoding performance may reflect some more general mechanisms or brain characteristics. This is also likely given that we found shared mechanisms between the cognitive sub-processes as well as many significant correlations among the decoding performances of different pairwise classification analyses. One example can be the issue of network modularity which is defined as the extent to which each brain subnetwork is segregated from other brain modules (Gallen & D'Esposito, 2019). Higher network modularity indicates higher functional specialization of different brain modules which may lead to better functional organization in processing of stimuli with different degrees of similarity to real Chinese characters.

Another possibility is that better decoding performance reflects better neural representation of orthographic forms of Chinese characters and character-like stimuli. If this is true, we should be able to discriminate the neural representation of individual characters from each other even when they are of the same type of stimuli such as real characters. Unfortunately, we do not have enough trials for each individual character to perform MVPA for such within-category classifications. Future studies should include more trials for each character and test this possibility. Given the high complexity of the orthographic forms of Chinese characters, it is intriguing to pursue the issue of whether neural representations of individual characters can be decoded in young children and how the neural discriminability of individual characters affects their word reading development.

4.4 | Decoding performance predicts word reading development

An important characteristic of the design of the current study is that children's word reading development was assessed using a longitudinal design in which the word reading tasks were administered to each child twice across 1 year. Many studies examining the relationships between neural markers and word reading skills performed only simple correlation at a single time point (e.g., Rämä et al., 2013; Shaywitz et al., 2002; Tong et al., 2016) and thus are not able to conclude a causal relationship between the two because a potential confound in such analyses is the initial differences in word reading performance. In the current study, the association between the decoding performance and word recognition development was examined by regressing the wave 2 word reading performance on decoding performance while statistically controlling for age and wave 1 word reading performance. The neural decoding performance was found to be significantly associated with wave 2 word reading performance when wave 1 word reading performance was not statistically controlled. This is not surprising since better decoding performance may indicate better functional organization of the brain for character reading and, hence, behaviorally more accurately and efficiently in recognizing and reading the words. This is also consistent with many of the previous studies that found associations between neural processes underlying word reading and reading skills measured at a single time point (e.g., Maurer et al., 2005; Tong et al., 2016; Turkeltaub et al., 2003).

More importantly, neural decoding performance was found to be significantly associated with wave 2 word reading fluency even when statistically controlling for age and wave 1 word reading fluency. Specifically, this neural decoding performance explained an additional 4.8% of word reading fluency. Although the association did not survive a Bonferroni correction and was marginally significant (Bonferroni-adjusted *p*-value = 0.078), this is still promising given that age and wave 1 word reading fluency together had already explained about 60.3% variance of the wave 2 word reading fluency. In contrast, the additional variance of word reading accuracy explained by decoding performance was only 1.8%. The statistical test was marginally significant before Bonferroni correction and did not survive the correction (Bonferroni-adjusted *p*-value = 0.160). This was consistent with our prediction that the association between neural discriminability and word reading should be stronger for reading fluency than accuracy given previous N1 print tuning studies (Maurer et al., 2007; Tong et al., 2016). Nevertheless, we found suggestive evidence that are consistent with results from a longitudinal study (Maurer et al., 2006) which found a N170 expertise effect only in children in second grade but not the same children in kindergarten. It should be noted, however, that the present study made use of a more powerful statistical technique (i.e., MVPA) in a different script (i.e., Chinese). Perhaps partly as a consequence, the current study has showed that not only decoding performance between real characters and stroke combinations but also decoding performance between real characters and pseudo-characters, as well as that between real characters and non-characters, can predict word reading skills longitudinally. One possible reason for the suggestive association between decoding performance and word reading development over the year is that a highly functionally specialized neural network may be easier to further refine for specialization in adapting to learning of new characters. This is consistent with the idea of brain modularity, which suggests that highly modular networks will lead to larger neural plasticity due to the relatively few between-module connections (Gallen & D'Esposito, 2019). One limitation to note for the current study relates to the timing of the EEG testing. Ideally it should have been conducted at about the same time as the wave 1 behavioral testing; however, for various practical reasons, it was conducted on average 2 months before wave 2 behavioral testing. As the association between decoding performance and word reading development did not survive the Bonferroni correction, further research is needed in order to attempt to replicate the findings. Nevertheless, this association suggests that decoding performance is potentially a useful neural marker both for predicting current and subsequent word reading skills.

5 | CONCLUSION

The present study was among the first to apply MVPA to decode category information specific to Chinese characters in single trial EEG data, particularly with a view to longitudinal prediction of behavioral word reading. Neural activation patterns of Chinese characters can be decoded and significantly distinguished from those of the pseudo-characters, non-characters, and stroke combinations in both younger and older children. Topography of the transformed classifier weights showed two temporally and spatially distinctive cognitive sub-processes underlying single character reading while the temporal generalization method found significant generalization across time, suggesting common mechanisms between the sub-processes. The decoding performance in general was significantly correlated across categories but not correlated with age. Given that the Chinese characters included were easy and the behavioral performance in the character decision task was comparable between age groups, the lack of significant correlations with age may suggest that basic neural discriminability of Chinese characters does not change obviously with neural maturation in this age range. The association between decoding performance and word reading development suggests that decoding performance is potentially a useful neural marker to indicate both current word recognition skills and also subsequent capacity to learn to read words.

ACKNOWLEDGEMENTS

This research was funded by the Collaborative Research Fund from the Hong Kong Special Administrative Region Research Grants Council (CUHK8/CRF/13G, and C4054-17WF) and the BRAIN grant from the Lui Che Woo Institute of Innovative Medicine (8303402) awarded to C. McBride. We thank all the research assistants and participants for their contribution to this work.

CONFLICT OF INTEREST

The authors hereby declare no conflict of interest.

AUTHOR CONTRIBUTION

K. Lui developed the study concept and designed the study. All authors collected the data, and K. Lui and J. Lo analyzed the data. All authors interpreted the data. K. Lui drafted the manuscript. All authors revised the manuscript and approved the final version of the manuscript for submission.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the Open Science Framework at https://osf.io/5wu9s/.

ORCID

Kelvin F. H. Lui https://orcid.org/0000-0003-4818-3136
Jason C. M. Lo https://orcid.org/0000-0002-8428-1086

REFERENCES

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience*, *11*(3), 235–260. https://doi.org/10.1162/089892999563373

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K. M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for

large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*, 1–20. https://doi.org/10.3389/fninf.2015.00016
Bode, S., Sewell, D. K., Lilburn, S., Forte, J. D., Smith, P. L., & Stahl, J. (2012). Predicting perceptual decision biases from early brain activity. *Journal of Neuroscience*, *32*(36), 12488–12498. https://doi.org/10.1523/JNEUROSCI.1708-12.2012
Bolger, D. J., Perfetti, C. A., & Schneider, W. (2005). Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Human Brain Mapping*, *25*(1), 92–104. https://doi.org/10.1002/hbm.20124
Brem, S., Lang-Dullenkopf, A., Maurer, U., Halder, P., Bucher, K., & Brandeis, D. (2005). Neurophysiological signs of rapidly emerging visual expertise for symbol strings. *NeuroReport*, *16*(1), 45–48. https://doi.org/10.1097/00001756-200501190-00011
Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turret, J. (2011). High temporal resolution decoding of object position and category. *Journal of Vision*, *11*(10), 1–17. https://doi.org/10.1167/11.10.9
Carlson, T., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, *13*(10), 1–19. https://doi.org/10.1167/13.10.1
Cauchoix, M., Arslan, A. B., Fize, D., & Serre, T. (2012). The neural dynamics of visual processing in monkey extrastriate cortex: A comparison between univariate and multivariate techniques. In *Machine learning and interpretation in neuroimaging* (pp. 164–171). Springer.
Cauchoix, M., Barragan-Jason, G., Serre, T., & Barbeau, E. J. (2014). The neural dynamics of face detection in the wild revealed by MVPA. *Journal of Neuroscience*, *34*(3), 846–854. https://doi.org/10.1523/JNEUROSCI.3030-13.2014
Chan, A. M., Halgren, E., Marinkovic, K., & Cash, S. S. (2011). Decoding word and category-specific spatiotemporal representations from MEG and EEG. *NeuroImage*, *54*(4), 3028–3039. https://doi.org/10.1016/j.neuroimage.2010.10.073
Cheng, Y., Zhang, J., Li, H., Wu, X., Liu, H., Dong, Q., Li, L., Nguyen, T. P., Zheng, M., Zhao, Y., & Sun, P. (2017). Growth of compounding awareness predicts reading comprehension in young Chinese students: A longitudinal study from grade 1 to grade 2. *Reading Research Quarterly*, *52*(1), 91–104. https://doi.org/10.1002/rrq.155
Chinese Language Education Section. (2009). *Lexical items with English explanations for fundamental Chinese learning in Hong Kong schools*. Government of Hong Kong Special Administrative Region, Education Bureau, Curriculum Development Institute.
Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, *125*(5), 1054–1069. https://doi.org/10.1093/brain/awf094
Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589–608. https://doi.org/10.1037/0033-295X.100.4.589
Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009
Ding, N., & Simon, J. Z. (2012). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *Journal of Neurophysiology*, *107*(1), 78–89. https://doi.org/10.1152/jn.00297.2011
Friedrich, M., & Friederici, A. D. (2006). Early N400 development and later language acquisition. *Psychophysiology*, *43*(1), 1–12. https://doi.org/10.1111/j.1469-8986.2006.00381.x
Gallen, C. L., & D'Esposito, M. (2019). Brain modularity: A biomarker of intervention-related plasticity. *Trends in Cognitive Sciences*, *23*(4), 293–304. https://doi.org/10.1016/j.tics.2019.01.014
Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*(4), 428–432. https://doi.org/10.1038/nn1029
Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. *Journal of Cognitive Neuroscience*, *29*(4), 677–697. https://doi.org/10.1162/jocn_a_01068
Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J. D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, *87*, 96–110. https://doi.org/10.1016/j.neuroimage.2013.10.067
Haynes, J. D. (2015). A primer on pattern-based approaches to fMRI: Principles, pitfalls, and perspectives. *Neuron*, *87*(2), 257–270. https://doi.org/10.1016/j.neuron.2015.05.025
Ho, C. S. H., Chan, D. W. O., Tsang, S. M., & Lee, S. H. (2002). The cognitive profile and multiple-deficit hypothesis in Chinese developmental dyslexia. *Developmental Psychology*, *38*(4), 543. https://doi.org/10.1037/0012-1649.38.4.543
Ho, C. S. H., Zheng, M., McBride, C., Hsu, L. S. J., Waye, M. M. Y., & Kwok, J. C. Y. (2017). Examining an extended simple view of reading in Chinese: The role of naming efficiency for reading comprehension. *Contemporary Educational Psychology*, *51*(September), 293–302. https://doi.org/10.1016/j.cedpsych.2017.08.009
Hogendoorn, H., Verstraten, F. A., & Cavanagh, P. (2015). Strikingly rapid neural basis of motion-induced position shifts revealed by high temporal-resolution EEG pattern classification. *Vision Research*, *113*, 1–10. https://doi.org/10.1016/j.visres.2015.05.005
Houdé, O., Rossi, S., Lubin, A., & Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: An fMRI meta-analysis of 52 studies including 842 children. *Developmental Science*, *13*(6), 876–885. https://doi.org/10.1111/j.1467-7687.2009.00938.x
Isik, L., Meyers, E. M., Leibo, J. Z., & Poggio, T. A. (2014). The dynamics of invariant object recognition in the human visual system. *Journal of Neurophysiology*, *111*, 91–102. https://doi.org/10.1152/jn.00394.2013
Jafarpour, A., Horner, A., Fuentemilla, L., Penny, W. D., & Duzel, E. (2013). Decoding oscillatory representations and mechanisms in memory. *Neuropsychologia*, *51*(4), 772–780. https://doi.org/10.1016/j.neuropsychologia.2012.04.002
King, J. R., Gramfort, A., Schurger, A., Naccache, L., & Dehaene, S. (2014). Two distinct dynamic modes subtend the detection of unexpected sounds. *PLoS One*, *9*(1), e85791. https://doi.org/10.1371/journal.pone.0085791
Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920. https://doi.org/10.1038/nrn2532
Lee, H. M. (2000). *List of frequently used character [in Chinese]*. Hong Kong Institute of Education.
Leppänen, P. H. T., Hämäläinen, J. A., Salminen, H. K., Eklund, K. M., Guttorm, T. K., Lohvansuu, K., Puolakanaho, A., & Lyytinen, H. (2010). Newborn brain event-related potentials revealing atypical processing of sound frequency and the subsequent association with later literacy skills in children with familial dyslexia. *Cortex*, *46*(10), 1362–1376. https://doi.org/10.1016/j.cortex.2010.06.003
Lin, S. E., Chen, H. C., Zhao, J., Li, S., He, S., & Weng, X. C. (2011). Left-lateralized N170 response to unpronounceable pseudo but not false Chinese characters—The key role of orthography. *Neuroscience*, *190*, 200–206. https://doi.org/10.1016/j.neuroscience.2011.05.071
Lo, J. C. M., McBride, C., Ho, C. S., & Maurer, U. (2019). Event-related potentials during Chinese single-character and two-character word reading in children. *Brain and Cognition*, *136*, 103589. https://doi.org/10.1016/j.bandc.2019.103589
Maurer, U., Brandeis, D., & McCandliss, B. D. (2005). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behavioral and Brain Functions*, *1*(1), 1–12. https://doi.org/10.1186/1744-9081-1-13

Maurer, U., Brem, S., Bucher, K., Kranz, F., Benz, R., Steinhausen, H. C., & Brandeis, D. (2007). Impaired tuning of a fast occipito-temporal response for print in dyslexic children learning to read. *Brain*, *130*(12), 3200–3210. https://doi.org/10.1093/brain/awm193

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

Maurer, U., Bucher, K., Brem, S., Benz, R., Kranz, F., Schulz, E., van der Mark, S., Steinhausen, H.-C., & Brandeis, D. (2009). Neurophysiology in preschool improves behavioral prediction of reading ability throughout primary school. *Biological Psychiatry*, *66*(4), 341–348. https://doi.org/10.1016/j.biopsych.2009.02.031

Maurer, U., & McCandliss, B. D. (2007). The development of visual expertise for words: The contribution of electrophysiology. In E. L. Grigorenko & A. Naples (Eds.), *Single-word reading: Cognitive, behavioral and biological perspectives* (pp. 43–64). Erlbaum.

Michel, C. M., & Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *NeuroImage*, *61*(2), 371–385. https://doi.org/10.1016/j.neuroimage.2011.12.039Get

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*(2), 229–240. https://doi.org/10.1111/j.1469-8986.2010.01061.x

Murphy, B., Poesio, M., Bovolo, F., Bruzzone, L., Dalponte, M., & Lakany, H. (2011). EEG decoding of semantic category reveals distributed representations for single concepts. *Brain and Language*, *117*(1), 12–22. https://doi.org/10.1016/j.bandl.2010.09.013

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *NeuroImage*, *45*(1), S199–S209. https://doi.org/10.1016/j.neuroimage.2008.11.007

Pugh, K. R., Frost, S. J., Rothman, D. L., Hoeft, F., Del Tufo, S. N., Mason, G. F., Molfese, P. J., Mencl, W. E., Grigorenko, E. L., Landi, N., Preston, J. L., Jacobsen, L., Seidenberg, M. S., & Fulbright, R. K. (2014). Glutamate and choline levels predict individual differences in reading ability in emergent readers. *Journal of Neuroscience*, *34*(11), 4082–4089. https://doi.org/10.1523/JNEUROSCI.3907-13.2014

Pugh, K. R., Mencl, WEinar, Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., Shaywitz, S. E., & Shaywitz, B. A. (2001). Neurobiological studies of reading and reading disability. *Journal of Communication Disorders*, *34*(6), 479–492. https://doi.org/10.1016/S0021-9924(01)00060-0

Raimondo, F., Kamienkowski, J. E., Sigman, M., & Slezak, D. F. (2012). CUDAICA: GPU optimization of infomax-ICA EEG analysis. *Computational Intelligence and Neuroscience*, *2012*, 1–8. https://doi.org/10.1155/2012/206972

Rämä, P., Sirri, L., & Serres, J. (2013). Development of lexical–semantic language system: N400 priming effect for spoken words in 18-and 24-month old children. *Brain and Language*, *125*(1), 1–10. https://doi.org/10.1016/j.bandl.2013.01.009

Redcay, E., & Carlson, T. A. (2015). Rapid neural discrimination of communicative gestures. *Social Cognitive and Affective Neuroscience*, *10*(4), 545–551. https://doi.org/10.1093/scan/nsu089

Rosenberg, M. D., Casey, B. J., & Holmes, A. J. (2018). Prediction complements explanation in understanding the developing brain. *Nature Communications*, *9*(1), 1–13. https://doi.org/10.1038/s41467-018-02887-9

Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *NeuroImage*, *20*(3), 1609–1624.

Salmelin, R., Kiesilä, P., Uutela, K., Service, E., & Salonen, O. (1996). Impaired visual word processing in dyslexia revealed with magnetoencephalography. *Annals of Neurology*, *40*(2), 157–162. https://doi.org/10.1002/ana.410400206

Sánchez-Vincitore, L. V., Avery, T., & Froud, K. (2018). Word-related N170 responses to implicit and explicit reading tasks in neoliterate adults. *International Journal of Behavioral Development*, *42*(3), 321–332. https://doi.org/10.1177/0165025417714063

Sandak, R., Mencl, W. E., Frost, S. J., & Pugh, K. R. (2004). The neurobiological basis of skilled and impaired reading: Recent findings and new directions. *Scientific Studies of Reading*, *8*(3), 273–292. https://doi.org/10.1207/s1532799xssr0803_6

Sandberg, K., Bahrami, B., Kanai, R., Barnes, G. R., Overgaard, M., & Rees, G. (2013). Early visual responses predict conscious face perception within and between subjects during binocular rivalry. *Journal of Cognitive Neuroscience*, *25*(6), 969–985. https://doi.org/10.1162/jocn_a_00353

Schaefer, R. S., Farquhar, J., Blokland, Y., Sadakata, M., & Desain, P. (2011). Name that tune: Decoding music from the listening brain. *NeuroImage*, *56*(2), 843–849. https://doi.org/10.1016/j.neuroimage.2010.05.084

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., Constable, R. T., Marchione, K. E., Fletcher, J. M., Lyon, G. R., & Gore, J. C. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry*, *52*(2), 101–110. https://doi.org/10.1016/S0006-3223(02)01365-3

Simanova, I., Van Gerven, M., Oostenveld, R., & Hagoort, P. (2010). Identifying object categories from event-related EEG: Toward decoding of conceptual representations. *PLoS ONE*, *5*(12), e14465. https://doi.org/10.1371/journal.pone.0014465

Simanova, I., Van Gerven, M. A., Oostenveld, R., & Hagoort, P. (2014). Predicting the semantic category of internally generated words from neuromagnetic recordings. *Journal of Cognitive Neuroscience*, *27*(1), 35–45. https://doi.org/10.1162/jocn_a_00690

Simos, P. G., Breier, J. I., Fletcher, J. M., Foorman, B. R., Mouzaki, A., & Papanicolaou, A. C. (2001). Age-related changes in regional brain activation during phonological decoding and printed word recognition. *Developmental Neuropsychology*, *19*(2), 191–210. https://doi.org/10.1207/S15326942DN1902_4

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Su, M., Wang, J., Maurer, U., Zhang, Y., Li, J., McBride, C., Tardif, T., Liu, Y., & Shu, H. (2015). Gene–environment interaction on neural mechanisms of orthographic processing in Chinese children. *Journal of Neurolinguistics*, *33*, 172–186. https://doi.org/10.1016/j.jneuroling.2014.09.007

Sudre, G., Pomerleau, D., Palatucci, M., Wehbe, L., Fyshe, A., Salmelin, R., & Mitchell, T. (2012). Tracking neural coding of perceptual and semantic features of concrete nouns. *NeuroImage*, *62*(1), 451–463. https://doi.org/10.1016/j.neuroimage.2012.04.048

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science*, *12*(1), 43–47. https://doi.org/10.1111/1467-9280.00308

Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain*, *122*(11), 2119–2132. https://doi.org/10.1093/brain/122.11.2119

Tong, X., Lo, J. C. M., McBride, C., Ho, C.-H., Waye, M. M. Y., Chung, K. K. H., Wong, S. W. L., & Chow, B.-Y. (2016). Coarse and fine N1 tuning for print in younger and older Chinese children: Orthography, phonology, or semantics driven? *Neuropsychologia*, *91*, 109–119. https://doi.org/10.1016/j.neuropsychologia.2016.08.006

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *NeuroImage*, *16*(3), 765–780. https://doi.org/10.1006/nimg.2002.1131

Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A., & Eden, G. F. (2003). Development of neural mechanisms for reading. *Nature Neuroscience*, *6*(7), 767. https://doi.org/10.1038/nn1065 van de Nieuwenhuijzen, M. E., Backus, A. R., Bahramisharif, A., Doeller, C. F., Jensen, O., & van Gerven, M. A. (2013). MEG-based decoding of the spatiotemporal dynamics of visual category perception. *NeuroImage*, *83*, 1063–1073. https://doi.org/10.1016/j.neuroimage.2013.07.075

van Zuijen, T. L., Plakas, A., Maassen, B. A., Maurits, N. M., & van der Leij, A. (2013). Infant ERPs separate children at risk of dyslexia who become good readers from those who become poor readers. *Developmental Science*, *16*(4), 554–563. https://doi.org/10.1111/desc.12049

Vandermosten, M., Hoeft, F., & Norton, E. S. (2016). Integrating MRI brain imaging studies of pre-reading children with current theories of developmental dyslexia: A review and quantitative meta-analysis. *Current Opinion in Behavioral Sciences*, *10*, 155–161. https://doi.org/10.1016/j.cobeha.2016.06.007

Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. *Frontiers in Systems Neuroscience*, *9*, 1–12. https://doi.org/10.3389/fnsys.2015.00123

Zhang, J., McBride-Chang, C., Tong, X., Wong, A. M. Y., Shu, H., & Fong, C. Y. C. (2012). Reading with meaning: The contributions of meaning-related variables at the word and subword levels to early Chinese reading comprehension. *Reading and Writing*, *25*(9), 2183–2203. https://doi.org/10.1007/s11145-011-9353-4

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. https://doi.org/10.1162/jocn_a_00660

Zhao, J., Maurer, U., He, S., & Weng, X. (2019). Development of neural specialization for print: Evidence for predictive coding in visual word recognition. *PLOS Biology*, *17*(10), e3000474. https://doi.org/10.1371/journal.pbio.3000474

How to cite this article: Lui KF, Lo JC, Maurer U, Ho CH, McBride C. Electroencephalography decoding of Chinese characters in primary school children and its prediction for word reading performance and development. *Dev Sci*. 2020;24:e13060. https://doi.org/10.1111/desc.13060

APPENDIX A

Topography of transformed classifier weights

As a significant age group difference was found in classifying between real characters and pseudo-characters, the topography of the transformed classifier weights in classifying between these two types of stimuli was examined to see if the underlying neural processes of Chinese character recognition were qualitatively different between the younger and older children. The classifier weights were reconstructed by multiplying them with the covariance of the data matrix (Haufe et al., 2014). Larger transformed classifier weights indicated that the channel contained more information specific to the experimental conditions; hence, the neural activity picked up by the channel was used to a larger extent in classifying between the experimental conditions. Topographic differences in the EEG indicate different underlying neural sources (Michel & Murray, 2012), suggesting different cognitive sub-processes. As shown in Figure 3, both groups showed left-lateralized activation patterns around the occipito-temporal channels starting at the 100–200 ms time window. The activation patterns then appeared to shift to the parietal and fronto-central channels at the 200–300 ms time window and finally became bi-lateralized at 300–500 ms time windows. The apparent sub-processes underlying Chinese character recognition found in the topography were highly consistent with those suggested by previous studies, including the fact that the orthographic encoding process occurred earlier (peak around 170–200 ms) in the occipito-temporal sites and the phonological and semantic analysis processes occurred later (peak around 400 ms) in the temporo-parietal and fronto-central regions (Bentin et al., 1999; Pugh et al., 2001; Salmelin et al., 1996; Tarkiainen et al., 1999). As the two age groups did not differ significantly in most of the decoding analyses and showed qualitatively the same pattern of topography, we combined the two age groups in the subsequent analyses.

APPENDIX B

Temporal generalization decoding results

The temporal generalization decoding analysis was performed to classify between real characters and stroke combinations, for slightly shorter epochs (−100 to 600 ms) than that of the above same time window decoding analysis (−150 to 850 ms) to examine whether the sub-processes underlying Chinese character recognition can be dissociated. These two types of stimuli were selected to maximize the differences between stimulus categories in terms of both orthography and phonology. A slightly shorter epoch length was adopted because the decoding performance dropped significantly after 500 ms (as shown in Figure 2) and the decoding performance of temporal generalization was expected to be lower than the standard same time window analysis. As the two age groups did not differ significantly in most of the decoding analyses and showed qualitatively the same pattern of topography, we combined the two age groups in the subsequent analyses. Figure 4 shows the temporal generalization classification accuracy averaged across participants and the significance in comparison to chance level. The diagonal basically represents decoding results of the standard same time window decoding analysis, so it is reasonable that the decoding accuracy drops gradually away from the diagonal.

In Figure 4a, two clusters of highly significant classification performance were observed as highlighted in the two black squares, suggesting at least two separate sub-processes underlying Chinese character reading. The first cluster in roughly the 200–300 ms time window seems to represent the orthography encoding process while the second cluster in roughly the 350–500 ms time window seems to represent the phonology and semantic analysis processes. However, the statistical significance pattern, as shown in Figure 4b, revealed a much larger cluster ranging from 150 to 600 ms instead of two smaller clusters as suggested in Figure 4a. The unified and larger cluster suggested that decoding performance for the sub-processes reflected certain common neural mechanisms. As a result, for the correlational and regression analyses, the decoding performance for each individual was calculated by averaging the classification accuracies across all post-stimulus time windows.

APPENDIX C

Correlations among decoding performance, age, and reading skills

In order to consider the longitudinal associations among the neural and behavioral markers of literacy, we first examined the correlations among wave 2 age, decoding performance of all pairwise classification analyses, and both wave 1 and wave 2 one-minute and untimed word reading performance. In Table 4, values presented above the diagonal are zero order correlation coefficients and values presented below the diagonal are partial correlation coefficients controlling for age. As shown, age was significantly and positively correlated with all wave 1 and wave 2 word reading performances, *rs* > 0.39, *ps* < 0.001, but only significantly correlated with the decoding performance in classifying between real characters and non-characters among all the classification analyses, $r(109) = 0.206$, $p = .03$. The word reading skills were all significantly and positively correlated with each other, *rs* > 0.69, *ps* < 0.001, and the decoding performances were in general significantly and positively correlated among themselves (9 correlations out of 15 were significant). Finally, both wave 2 one-minute and untimed word reading performance were significantly correlated with some decoding performances. One-minute word reading was positively correlated with decoding performance between real characters and pseudo-characters, $r(109) = 0.190$, $p = .045$, decoding performance between real characters and non-characters, $r(109) = 0.239$, $p = .011$, decoding performance

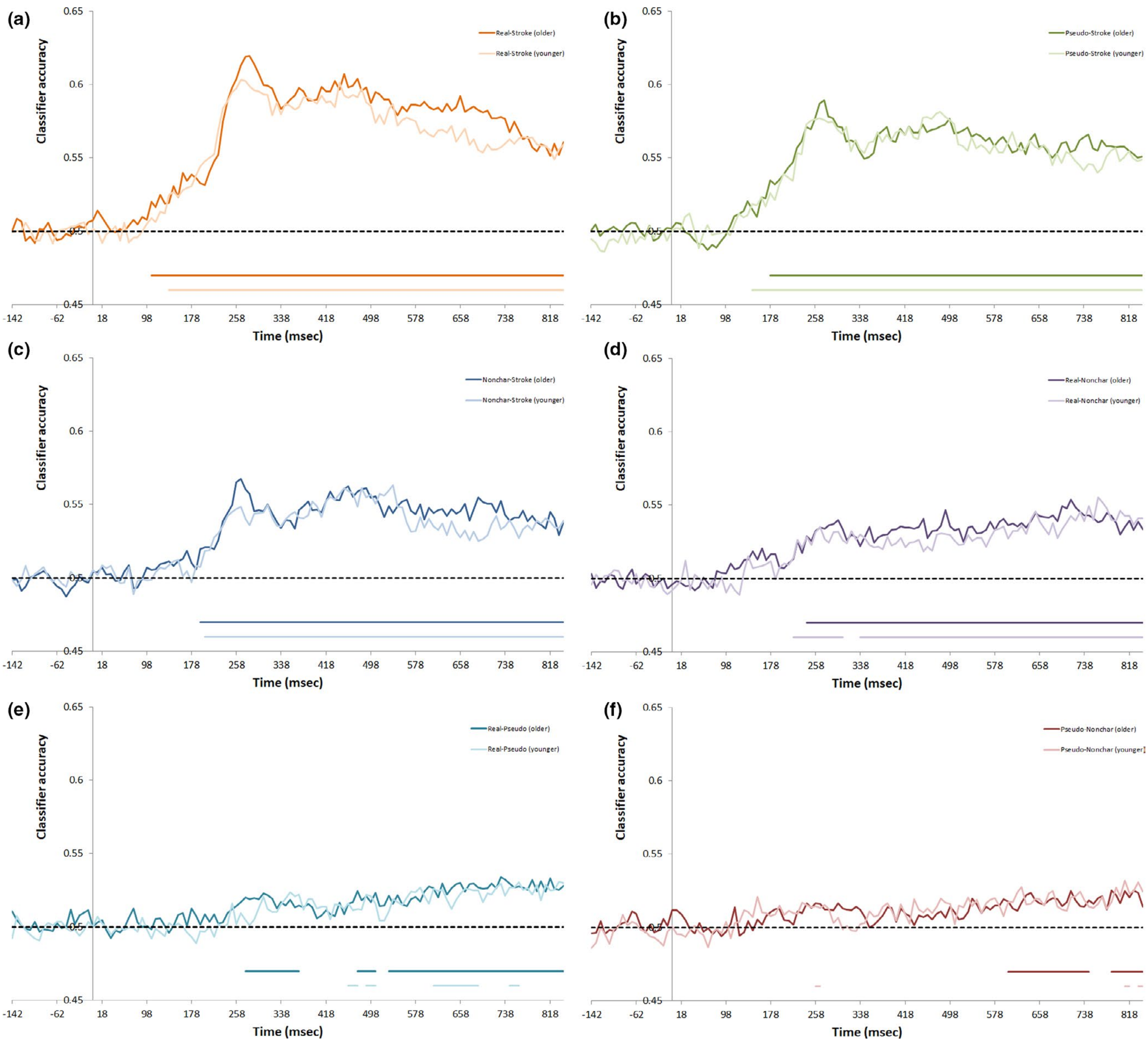

FIGURE D1 Within-individual classification accuracies of all pairwise combinations of the four stimuli categories (a: Real-Stroke; b: Pseudo-Stroke; c: Nonchar-Stroke; d: Real-Nonchar; e: Real-Pseudo; f: Pseudo-Nonchar). Discs above the *x*-axis indicated the time points where classification accuracies were significantly above chance between real characters and stroke combinations, $r(109) = 0.274$, $p = .004$, and decoding performance between pseudo-characters and stroke combinations, $r(109) = 0.356$, $p < .001$. Untimed word reading was positively correlated with decoding performance between real characters and non-characters, $r(109) = 0.266$, $p = .005$, decoding performance between real characters and stroke combinations, $r(109) = 0.189$, $p = .047$, and decoding performance between pseudo-characters and stroke combinations, $r(109) = 0.202$, $p = .034$. The pairwise decoding analyses that were highly correlated with word reading performance appeared to be those with higher decoding accuracies such as decoding between real characters and stroke combinations and decoding between pseudo-characters and stroke combinations. The partial correlation results controlling for age were qualitatively the same as the zero order correlation results. All the significant correlations were positive, suggesting that better neural indicators of decoding performance tend to be associated with better behavioral word reading skills.

APPENDIX D

Classification accuracy plots faceted by age group