

The Development of Cortical Sensitivity to Visual Word Forms

Michal Ben-Shachar^{1,2}, Robert F. Dougherty¹, Gayle K. Deutsch¹,
and Brian A. Wandell¹

Abstract

■ The ability to extract visual word forms quickly and efficiently is essential for using reading as a tool for learning. We describe the first longitudinal fMRI study to chart individual changes in cortical sensitivity to written words as reading develops. We conducted four annual measurements of brain function and reading skills in a heterogeneous group of children, initially 7–12 years old. The results show age-related increase in children's cortical sensitivity to word visibility in posterior left occipito-temporal sulcus (LOTS), nearby the anatomical location of the visual word form area. Moreover, the rate of

increase in LOTS word sensitivity specifically correlates with the rate of improvement in sight word efficiency, a measure of speeded overt word reading. Other cortical regions, including V1, posterior parietal cortex, and the right homologue of LOTS, did not demonstrate such developmental changes. These results provide developmental support for the hypothesis that LOTS is part of the cortical circuitry that extracts visual word forms quickly and efficiently and highlight the importance of developing cortical sensitivity to word visibility in reading acquisition. ■

INTRODUCTION

In early schooling, children are exposed to hundreds of thousands of written words each year (Nagy & Anderson, 1984). Such extensive exposure to a particular class of stimuli may shape the functional properties of the visual pathways via experience-dependent cortical plasticity. This assumption underpins the hypothesis that readers have a “visual word form area” (VWFA)—a specific cortical region in left ventral occipito-temporal cortex specialized for visual word processing (Dehaene, Cohen, Sigman, & Vinckier, 2005; McCandliss, Cohen, & Dehaene, 2003; Cohen et al., 2000). Cortical specialization for written words is unlikely to be part of our genetic makeup, because reading is a relatively recent invention in human history. But the ability of experience-dependent plasticity to shape certain parts of cortex for efficiently identifying important visual categories in our environment may well be part of our genetic endowment (Dehaene & Cohen, 2007). In this article, we examine the developmental predictions of the visual word form hypothesis by measuring changes in cortical sensitivity to visual words as children learn to read.

The Visual Word Form Hypothesis: Background and Recent Developments

The search for a cortical area specialized for visual word recognition dates back to the early observations by Dejerine (1892) and to later neuropsychological studies (Warrington

& Shallice, 1980). Within the cognitive neuroscience literature, the visual word form hypothesis has been revived by Cohen and Dehaene (Dehaene et al., 2005; McCandliss et al., 2003; Cohen et al., 2000). Many studies have followed up on this idea, showing specific responsivity to visual words and their orthographic properties in the adult ventral occipito-temporal (VOT) cortex (Glezer, Jiang, & Riesenhuber, 2009; Baker et al., 2007; Vinckier et al., 2007; Binder, Medler, Westbury, Liebenthal, & Buchanan, 2006). In recent years, it has been argued that, rather than a discrete location, there is a visual word form system of print-sensitive regions extending along the VOT surface (Dehaene et al., 2005). This model has been supported by data showing a hierarchical organization along the VOT, with increasing specificity to larger orthographical units as we move ventrally toward the temporal pole (Vinckier et al., 2007; Brem et al., 2006).

The visual word form hypothesis stirred a debate in the neuroimaging literature pertaining to its modality and stimulus specificity (Vigneau, Jobard, Mazoyer, & Tzourio-Mazoyer, 2005; Price & Devlin, 2003, 2004). In previous work, we developed a parametric incidental reading paradigm that maps adults' cortical response functions to words and other stimuli. We showed that BOLD signals in the posterior occipito-temporal sulcus (OTS) increase with stimulus visibility, and this is true not only for words but also for line drawings of objects and false font strings (Ben-Shachar, Dougherty, Deutsch, & Wandell, 2007b). We refer to this region by its anatomical location (OTS), because the functional name VWFA suggests that it responds

¹Stanford University, ²Bar Ilan University, Ramat-Gan, Israel

only to words (contra to our findings) and that it is the only such region. Importantly, our results also demonstrated word specialization in this region, which was significantly more sensitive to words than to other stimuli. These results make a clear developmental prediction that the sensitivity to word visibility in OTS develops over the school years. The exact developmental profile and the cognitive changes that accompany these cortical changes are the topic of this study.

The Development of Word Sensitivity and Specificity in VOT Cortex

The principle that cortex learns to process visual word forms predicts that sensitivity and specificity to visual word forms develop in VOT cortex during early education. Several studies examined the development of word form responses in school age children. For example, a series of articles by Brem and colleagues (Brem et al., 2006, 2009, 2010) use both ERPs and fMRI to measure the development of responses to words and symbol strings in posterior VOT. Across age groups, they found a negative relationship between reading performance and the amplitude of fMRI responses to words in VOT (Brem et al., 2009). However, no significant age-dependent differences were found in the spatial distribution of responses to words and symbol strings (Brem et al., 2006, 2009). The amplitude of the N1 component of the ERP differed between ages, and this component was localized to the general position of the left occipito-temporal sulcus (LOTS) (Brem et al., 2006, 2009; Maurer et al., 2006, 2007).

Several fMRI comparisons between age groups (commonly adults and school age children) fail to find a difference in the OTS (Brown et al., 2005; Turkeltaub, Gareau, Flowers, Zeffiro, & Eden, 2003). Yet other fMRI studies find age effects as well as reading skill effects in groups of children versus adults in the vicinity of the OTS (Church, Coalson, Lugar, Petersen, & Schlaggar, 2008; Kronbichler et al., 2006). Differences between orthographic systems are unlikely to be the reason for this variability in fMRI results, because there are both negative and positive results in both German (compare Brem et al., 2006 with Kronbichler et al., 2006) and English (compare Church et al., 2008 with Turkeltaub et al., 2003). Differences in task and stimulus contrast, as well as in the composition of the research population (percentage of good and poor readers), may contribute to or conversely mask developmental differences around the OTS. Finally, the VOT is comprised of multiple cell populations with differing functional preferences. Group comparisons may miss individual differences in the spatial organization of these specialized assemblies.

A Longitudinal Perspective on Individual Cortical Development

Individual variability can be properly addressed by assessing cortical responses individually and tracking changes in re-

sponse profiles longitudinally. This experimental approach is methodologically advantageous (Casey, Tottenham, Liston, & Durston, 2005) but difficult to implement for two primary reasons: (a) *Changes in signal to noise ratio over time*: Measurement noise complicates comparisons between fMRI activation levels over developmentally relevant periods (months or years). (b) *Attrition*: A true longitudinal perspective requires assessment of change over three or more time points (Willett, Singer, & Martin, 1998), and attrition rates are particularly high in child populations, for example, because of family mobility.

For these reasons and others, long-term longitudinal studies of brain changes underlying reading development in the normal population are still quite rare, and the longitudinal perspective is used mostly for testing intervention effects (Brem et al., 2010; Shaywitz et al., 2004; Temple et al., 2003). Although longitudinal studies are particularly powerful for measuring *individual* change, no previous study has traced changes in individual children's cortical responses to visual words. Instead, previous studies compare group levels of activation at different time points, typically before and after intervention. This approach has the advantage of casting a wide net across the whole brain and enhancing the chances for detecting an existing pre/post difference. However, it does not provide a precise measure of the change within individual and, therefore, makes it hard to relate cortical and behavioral changes.

In this study, we use longitudinal fMRI and cognitive measurements to track the developmental profile of cortical sensitivity to word visibility and examine which components of reading change at the same pace as the sensitivity to word visibility in the OTS. We followed a diverse group of children over four annual measurements. To overcome long-term changes in signal to noise ratio, we developed a robust measure of cortical sensitivity to word visibility, which is defined as a threshold value on a stimulus-response curve. Similar to the psychophysical threshold, the sensitivity measure is "stimulus referred"—it is quantified in the domain of the stimulus (visibility), and it relies on the shape of the whole stimulus response function, not on the absolute or relative fMRI amplitude. To reduce the attrition problem, we recruited a large initial sample and maintained a stable team who kept constant communication with our participating families. The results provide a unique window into the cortical changes at the level of the individual participant in relation to cognitive changes within the same individual.

METHODS

We followed children with a wide range of reading skills and ages in an accelerated longitudinal design (Willett et al., 1998). Four annual waves of longitudinal measurements included anatomical and functional MRI scans as well as a battery of age-standardized cognitive tests. In

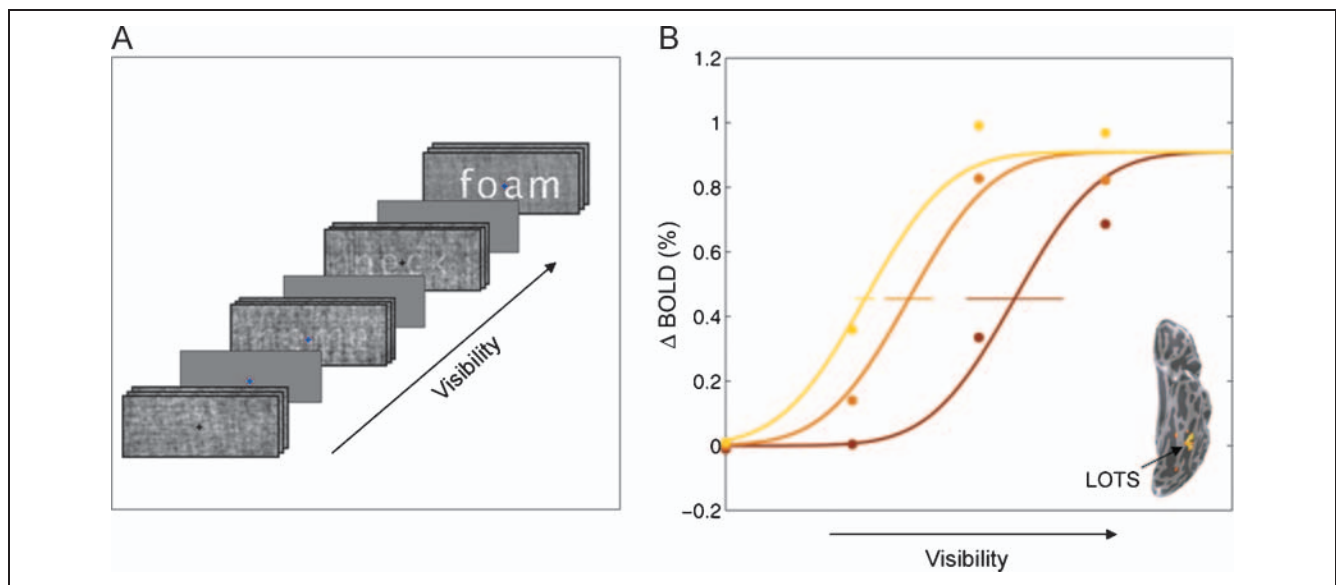


Figure 1. Experimental paradigm and age-grouped neurometric curves in LOTS. (A) Word visibility was controlled parametrically by varying the amount of phase scrambling applied to common four-letter English nouns. Stimuli were presented in blocks (six stimuli, 12 sec) in a pseudorandom order, interleaved with fixation blocks (12 sec, uniform gray rectangles). A black or dark blue fixation cross was refreshed every 2 sec throughout all blocks. Subjects indicated the color of the fixation cross using a response box. (B) Neurometric curves measured in the LOTS grouped by age: 7–8 years (brown), 9–11 years (orange), and 12–15 years (yellow). Mean change in contrast (Δ BOLD) is plotted as a function of visibility (inverse noise level). The Δ BOLD contrast is the difference between the fMRI responses for shape + noise and for noise alone. Circles represent the data, and separate curves were fitted to the data from each age group. The fitted curves were constrained to have the same upper asymptote and slope, differing only in horizontal position. Horizontal error bars represent the standard error, computed by bootstrapping.

fMRI, subjects viewed words presented at four different visibility levels (Figure 1A; see Ben-Shachar et al., 2007b, for a detailed description of the functional paradigm as implemented in adults). We measured children's fMRI signals as a function of word visibility, that is, their word visibility neurometric curve, and quantified changes in these curves within the same individual over time.

Participants

The initial cohort included 53 children (29 girls, 24 boys), 7–12 years old. Demographics describing the participant

population across the four sample points are shown in Table 1. Participants with a wide range of reading skills were recruited from the San Francisco Bay Area. Written informed consent or assent was obtained from both parents and children. All participants were physically healthy and had no history of neurological disease, head injury, psychiatric disorder, language disability (verified by the Clinical Evaluation of Language Fundamentals-3 (CELF-3) Screening Test), attention deficit/hyperactivity disorder (verified by Conners' Parent Rating Scale-Revised Short Form), or depression (verified by the Children's Depression Inventory Short Form). All participants were native English speakers and had normal or corrected-to-normal vision. Participants were paid per session; they also received

Table 1. Demographics of Participant Population

	Year 1	Year 2	Year 3	Year 4
Total sample size	53 (M: 24, F: 29)	44 (M: 23, F: 21)	37 (M: 17, F: 19)	29 (M: 16, F: 13)
Usable fMRI data sets	33 (62%) M: 16, F: 17	35 (79%) M: 17, F: 18	28 (76%) M: 12, F: 16	24 (83%) M: 11, F: 13
Age range	[7.1; 12.15]	[8; 12.9]	[9.1; 13.8]	[10; 14.8]
FSIQ: mean (<i>SD</i>) [range]	110.8 (13.2) [85; 145]			
Basic reading (WJ-III)	106 (13.2) [77; 134]	105.5 (13.5) [80; 135]	105.5 (12.2) [84; 132]	107.3 (11.9) [84; 129]

M = male, F = female, FSIQ = full-scale intelligence quotient (Wechsler Intelligence Scale for Children-IV, Wechsler, 2003), WJ-III = Woodcock Johnson III—Tests of Achievement (Woodcock, 1987).

a brain picture and small merchandise items for their participation each year (a water bottle and a t-shirt with a brain logo). The Stanford Panel on Human Subjects in Medical and Non-medical Research approved all procedures.

Subject attrition was about 20% each year (Table 1, top row). The principal reasons for attrition were (a) nonremovable dental braces, frequently installed around ages 9–10, causing artifacts in MRIs, and (b) loss of interest, particularly among the older children. Twenty-eight participants of the initial 53 remained with us for the entire four-measurement study (the twenty-ninth participant in Measurement 4 could not participate in Years 2 and 3 because of braces but returned for the fourth measurement when those were removed).

Procedure

Each year, starting in summer 2004 and ending in summer 2007, participants took part in three or four separate experimental sessions in the following order: cognitive assessment, anatomical MRI, and functional MRI (one or two sessions). The anatomical MRI session gave participants the opportunity to adjust to the scanner environment. This MRI scan also served as practice for reducing head motion; this artifact can be detected in short anatomical scans as image blurring.

Cognitive Assessment

Participants were administered a comprehensive assessment of reading and reading relevant cognitive skills that took approximately 4 hr to complete. Tests were administered by an experienced neuropsychologist (G.K.D). In following years, a shortened 2-hr assessment was administered. The assessment included the following tests: Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1999), subtests from the Woodcock-Johnson III Test of Achievement (WJ-III) (Woodcock, 1987), subtests from the Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999), and the Gray Oral Reading Test-4 (Wiederholt & Bryant, 2001). Intelligence was assessed by the Wechsler Intelligence Scale for Children-IV (first year only) (Wechsler, 2003), and general language was assessed by the CELF-3 Screening Test (Semel, Wiig, & Secord, 1995). Full-scale intelligence quotient of at least 85 and a passing score on the CELF-3 Screening Test were required for inclusion. See Supplementary Table 1 for a description of the relevant subtests referred to in this manuscript.

fMRI Stimulus Paradigm

Word stimuli were presented at four visibility levels (Figure 1A). In the least visible condition, it is impossible to see

a word because the contour information is fully scrambled by noise. In the two intermediate conditions, visibility is gradually enhanced. In the most visible condition, the word is fully visible.

Words ($n = 144$) were selected from a list of 192 common four-letter English nouns. The letters were centered on a fixation cross; there were two letters on each side of the cross, and no overlap between the letters and the cross (see Figure 1A). The letter font was white, “MS Sans Serif” (18 pt). The words and fixation were presented within a gray rectangular frame that spanned 7° visual angle; the words spanned 2.2 – 4.42° (mean = 3.5°).

The four visibility levels were created by randomizing the phase of the word image and creating a weighted sum of the original and phase scrambled image (see Ben-Shachar et al., 2007b, for more detail). Stimuli were presented through a backlit projection screen, visible to the subject by a mirror mounted on the top of the head coil. Responses were collected using an MRI compatible response box. Stimulus presentation and response collection were controlled by E-Prime (Psychology Software Tools, Pittsburgh, PA).

Stimuli were presented in a block design. Word presentation blocks were interleaved with fixation-only blocks. Each experimental run (six runs in total) began with a word presentation block (excluded from analysis, see below) and ended with a fixation block. Runs included seven experimental and seven fixation blocks. A fixation mark (plus sign, Arial font size of 16, bold), either dark blue or black, was present in all trials (including fixation blocks). Stimuli were presented centrally for 1700 msec, followed by a blank screen (300 msec), resulting in an ISI of 2 sec (six stimuli per block, 12-sec blocks).

In each trial, participants identified the fixation mark color with a button press; thus, reading was implicit. Children performed the fixation task with high accuracy (87–92% correct on different conditions and measurement points, no significant difference between conditions or measurement points). The main reasoning behind the choice of task was to avoid age-related and skill-related differences in task performance or in the motivation to process the stimuli. Our experience over more than 150 functional scans with children shows that (a) it is very important to include a behavioral task inside the scanner to maximize alertness and minimize motion, but (b) many children, particularly younger and poorer readers, are worried that they will be asked to read while in the scanner. In fMRI, we critically depend on the participants’ compliance; if they are unhappy with what they are asked to do, they may just close their eyes or attend elsewhere. The fixation task is a way to achieve these conflicting goals within the same experiment. Previous studies with adults and children showed that implicit reading is effective in activating the reading network (Turkeltaub et al., 2003; Moore & Price, 1999). As mentioned earlier, extensive assessment of reading skills was achieved in separate sessions outside the scanner.

MRI Data Acquisition

fMRI measurements were obtained using a 3T General Electric scanner with a custom head coil. Head movements were minimized by padding and surgical tape across the forehead. Functional MR data were acquired with a spiral-out pulse sequence (Glover, 1999). Twenty-six oblique slices (3-mm thick, no gap) were prescribed approximately along the ac–pc plane, covering the occipital and temporal lobes as well as ventral portions of the parietal and frontal lobes (TR = 2000 msec, TE = 30 msec, flip angle = 76°, voxel size = $2.5 \times 2.5 \times 3$ mm).

In-plane anatomical images were acquired before the functional scans using a fast spoiled gradient recalled (SPGR) sequence. These T1-weighted slices were taken at the same location as the functional slices and were used to align the functional data with high-resolution anatomical data acquired in a separate session. A week before the functional scan session, children attended an anatomical scan session in which we collected diffusion tensor imaging (DTI) data (not reported here) and several high-resolution ($1 \times 1 \times 1.2$ mm) T1-weighted whole-brain anatomical images on a GE 1.5-T Signa LX scanner (3-D SPGR pulse sequence). We averaged the high-resolution SPGR anatomical data sets for improved gray–white contrast and registered them year-to-year for each subject (Dougherty et al., 2005). Supplementary Figure 8 demonstrates the quality of the registration and the slight anatomical changes in the posterior part of the brain by overlaying segmentations of high-resolution anatomy from Year 1 on top of the anatomy from Year 4. Even in young children, posterior anatomical changes were small and below the resolution of our functional measurements.

Data Analysis

We analyzed the change in individual participants' sensitivity to visual word forms within predefined ROIs. First, ROIs were individually defined guided by individual activation maps (for all word conditions) and anatomical guidelines. For each individual and each ROI, we extracted the neuro-metric BOLD response curves as a function of word visibility and examined the year-to-year change in the curves. Finally, we quantified curve change over time and compared this change with behavioral change. We explain each of these steps below. fMRI data were analyzed using mrVista (white.stanford.edu/software/) and Matlab R2006b (Mathworks, Natick, MA).

Preprocessing

Data were analyzed voxel by voxel in individual subjects with no spatial smoothing. The first six volumes of each run were excluded from analysis to allow for signal stabilization. Baseline drifts were removed from the time series by high pass temporal filtering. Subjects were excluded from analysis because of excessive (>1 voxel) within-scan motion in more than two of the six scans. The remaining

data sets were corrected for between-scan motion by registering the mean images of each scan. See Table 1 for the number of data sets that passed our motion criteria in each year. Motion compensation methods are described in detail elsewhere (Ben-Shachar, Dougherty, Deutsch, & Wandell, 2007a).

Contrast Maps

Spatial contrast maps were computed for each measurement (year) by fitting a general linear model to each voxel's time course. The general linear model predictors were constructed by convolving the block timing with a standard hemodynamic response model (Boynton, Engel, Glover, & Heeger, 1996). Predictors were added for each run to model between run variations. Contrast maps were computed as voxel-wise t tests between the predictor weights for the relevant experimental conditions.

Visualization

The high-resolution volume anatomy of 12 subjects was segmented into gray and white matter using custom software and then hand-edited to minimize segmentation errors (Teo, Sapiro, & Wandell, 1997). The surface at the white–gray border was smoothed and rendered as a three-dimensional surface using VTK software (www.vtk.org/). Data from all gray layers were mapped to this surface, and the maximum value of those was assigned to each triangle on the surface. For visualization purposes only, spatial smoothing along the cortical surface was applied using an iterative neighborhood average roughly equivalent to a 4-mm FWHM Gaussian. These mesh representations were used in preparation of Figures 1B and 3 and Supplementary Figures 6 and 7.

ROI Analysis

ROIs were defined in individual participants' brains according to the following combination of anatomical and functional guidelines. Bilateral ROIs were constrained to fall within the following anatomical borders: OTS—in the depth of the OTS, at the junction between OTS and the inferior temporal sulcus; V1 (foveal)—at the posterior third of the calcarine sulcus; and posterior parietal cortex (PPC)—in the depth of the intraparietal sulcus, just lateral to the parieto-occipital sulcus and just medial to the bend of the angular gyrus. We selected voxels within these anatomical borders by functional criteria. For each measurement (year), we marked voxels that passed the threshold ($p < 10^{-3}$, uncorrected) for a functional localizer contrasting all word conditions (excluding full noise) versus fixation (Ben-Shachar et al., 2007b). We took the intersection of the functional localizer maps across the years. This ROI was transformed back to the functional slices from each year. Time course data were collected from voxels included in the intersection ROI within the anatomical borders detailed above, in the original in-plane, noninterpolated

data. In this way, we analyzed changes in the response properties of the same set of voxels along time. See Supplementary Figure 8 for an assessment of registration quality and anatomical change around the relevant brain regions.

Neurometric Functions

The word visibility neurometric function (Figure 1B) plots the difference between the BOLD contrast in each condition and the full-noise response contrast as a function of visibility (for a detailed explanation of the computation of BOLD contrast and word visibility neurometric functions, see Ben-Shachar et al., 2007b). The smooth curves are cumulative Gaussians fit to the mean data for each age group. The curves are fit using five parameters: To avoid overfitting, a common asymptote BOLD response and slope were found for all curves, and a Gaussian mean (horizontal shift) was determined for each of the three curves. Variability of the mean was estimated by bootstrapping (Efron & Tibshirani, 1993). We resampled with replacement (1000 samples) from the subject data within each age

group. We then found the five parameters and computed the standard deviations of the Gaussian means. Mean age grouped curves without fitting and with vertical error bars are shown in Supplementary Figures 1, 3, and 4. These curves were scaled to the mean response for the full-noise condition across all age groups.

To assess the change in individual's response curves along time, we first summarized the response curve measured each year by a single number that represents sensitivity. We defined neurometric sensitivity as one minus the visibility level that gives rise to half the maximum BOLD response (Figure 2A). We used this nonparametric sensitivity measure for individual data to avoid parametric assumptions about the shape of the function and to overcome fluctuations in overall signal strength across the years. We then plotted, for 28 individuals with three or more usable data sets (time points), the sensitivity value as a function of age. Again, given the small number of data points (3 or 4) per individual, we fitted these points with a simple line, and the slope of this line represents the change in sensitivity along the years for each individual (Figure 2B). We applied the same computation to the raw cognitive test scores plotted as a function of age, deriving the slope of a fitted

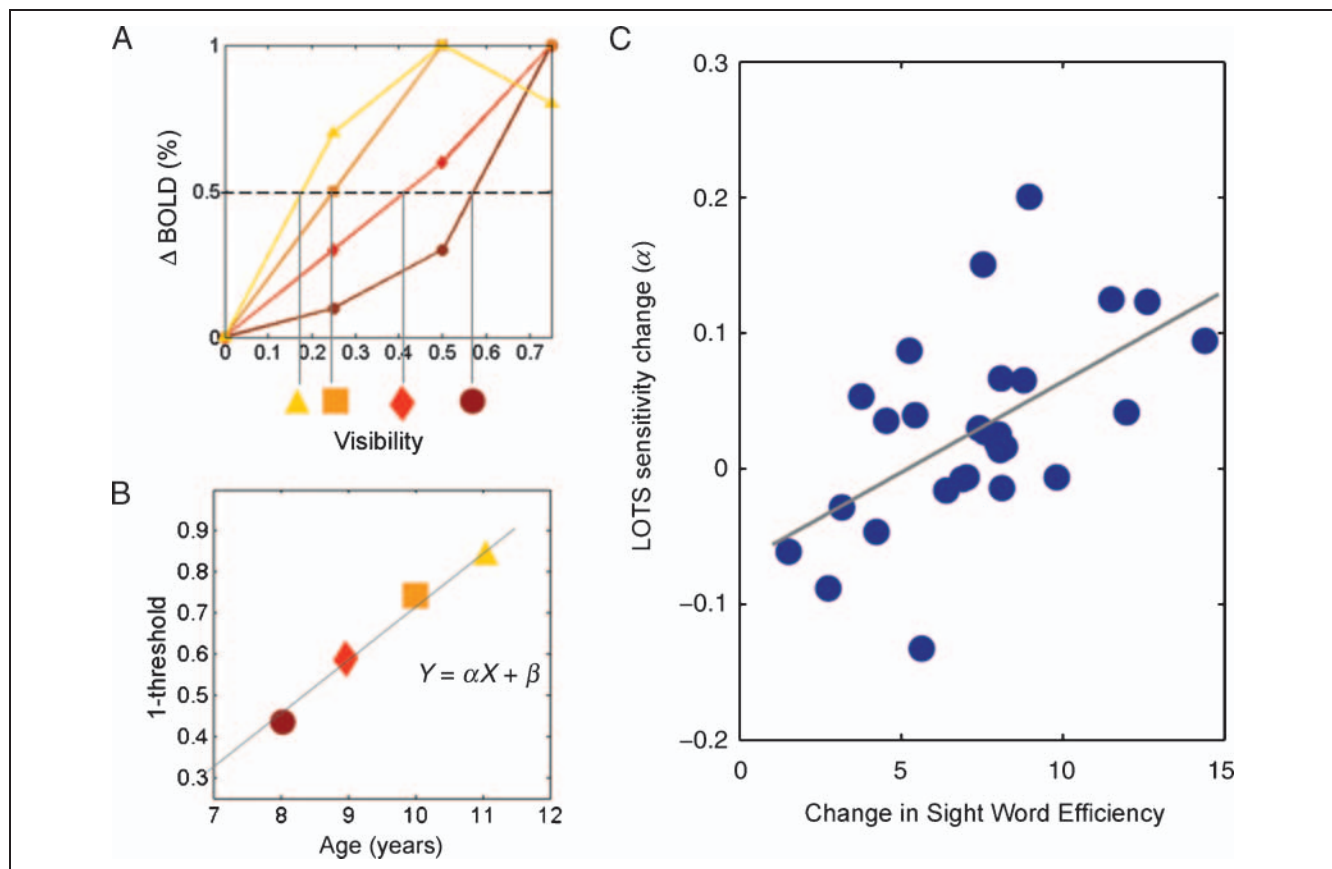


Figure 2. Longitudinal change in LOTS predicted by the change in SWE. (A) Threshold visibility is the visibility level that gives rise to half the maximum fMRI response, and sensitivity is $(1 - \text{threshold})$. In this example, the threshold decreases (sensitivity increases) across four measurements (brown to yellow shapes). (B) Change in sensitivity is computed as the slope (α) of a line fit to the sensitivity measures over time. Similarly, behavioral change is computed as the slope of a line through the raw behavioral scores over time. Subjects with fewer than three usable data sets are excluded from this analysis. (C) The correlation between longitudinal change in LOTS sensitivity and SWE is highly significant ($r = 0.564$, $SE = \pm 0.11$; SE is computed by bootstrapping, $n = 28$).

line as a measure of the change in a cognitive measure. We then compared the neurometric changes with the cognitive changes (Figure 2C and Supplementary Figure 2) and assessed the correspondence between localized changes in cortical sensitivity and changes in specific cognitive measures.

RESULTS

Development of Sensitivity to Word Forms: Cross-sectional Analyses

We first examine children's neurometric curves grouped by age. As shown in Figure 1B, there are systematic differences in the word visibility neurometric curve in LOTS as children develop. To quantify the difference in neurometric function across age groups, we fit the age-grouped data using a cumulative Gaussian function. This function is defined by three parameters: sensitivity, slope, and upper asymptote. We fix the slope and asymptote across the age groups, and we allow only the sensitivity parameter (horizontal shift) to differ between groups. Sensitivity should be interpreted similar to a threshold value, as it is the visibility level that gives rise to half the asymptote of the fMRI signal. Therefore, a lower sensitivity value means better performance, an ability to reach higher signal for a less visible word.

In the age-grouped data (Figure 1B), the sensitivity parameter changes substantially from the younger age group (0.57) to the older age group (0.275). The mean sensitivity value for words in the older age group agrees with the value we recently reported in adults (0.277) (Ben-Shachar et al., 2007b). This quantitative agreement provides independent validation of our measurements in children; it further suggests that, for English readers, measurements of LOTS sensitivity to words are stable after the age of 12 years. Other aspects of the cortical responses to words, such as stimulus specificity, may continue to develop after this age, as suggested by ERP findings in German (Brem et al., 2006).

We also measured word visibility neurometric functions in the right occipito-temporal sulcus (ROTS). Responses in the ROTS increase with word visibility, matching our results in adults (Ben-Shachar et al., 2007b). In contrast to the LOTS, however, the ROTS neurometric functions do not differ significantly between age groups (Supplementary Figure 1). The lack of a cross-sectional age effect in ROTS suggests that this region does not undergo the same experience-dependent changes as the LOTS.

Development of Sensitivity to Word Forms: Longitudinal Analyses

Longitudinal measurements enable us to make inferences about individual change, beyond the group comparison (Singer & Willett, 1996). To assess change robustly, one

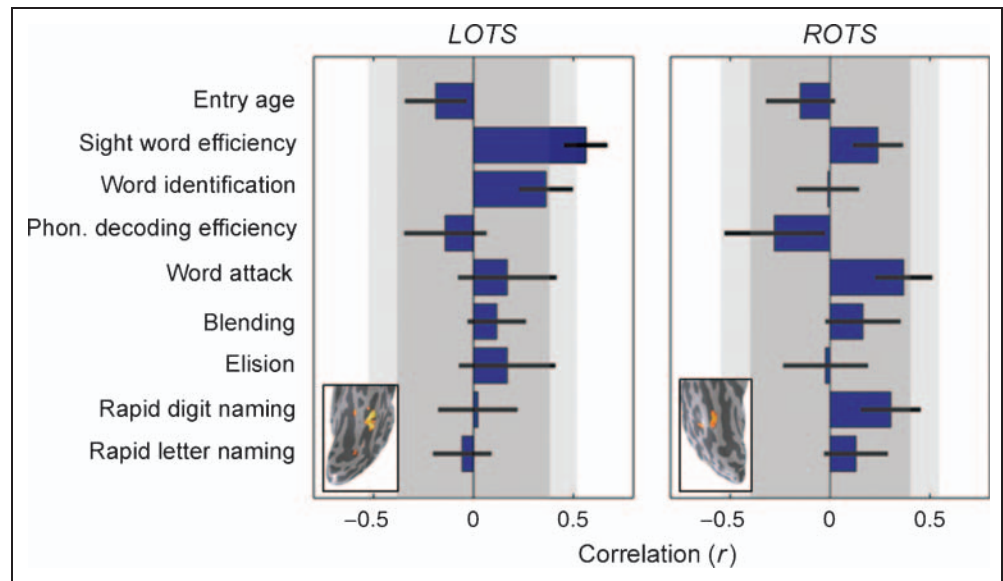
must acquire three measurements or more from the same individual. In our sample, 28 children provided data from three or more time points after attrition and exclusion of data because of excessive motion artifacts. To quantify individual change in these 28 participants, we first compute, for each temporal sample, each child's sensitivity to words. Individual sensitivity is defined as the stimulus visibility level that gives rise to half the maximum of the fMRI response in that child (Figure 2A). Notice that we avoid fitting a higher-order parametric function to individual subject data, because variability in single-subject data precludes a high-quality fit of any single parametric function; the measure we use for individual sensitivity is, therefore, a simple, nonparametric, $[1 - x \text{ at half maximum}]$ measure, with linear interpolation between the measured values. We then find a line through the sensitivity values plotted against age; the slope of this line is the change in sensitivity over time (Figure 2B). We use the same method to characterize the change in cognitive test scores. For longitudinal measurements, it is best to assess the change in raw scores rather than the changes in age-standardized scores, because the latter reflect changes in the child's status compared with his cohort rather than individual change (Willett et al., 1998).

Measured child by child, the change in LOTS cortical sensitivity significantly correlates with the change in sight word efficiency (SWE) raw scores (Torgesen et al., 1999) (Figure 2C). SWE is a measure of the number of frequent words read aloud in 45 sec. The correlation is highly specific: change in LOTS sensitivity does not correlate with raw scores in phonemic decoding efficiency (the number of pseudowords read in 45 sec) (Torgesen et al., 1999) or rapid letter naming (naming times were transformed to number of letters named per minute, for consistency with SWE and phonemic decoding efficiency) (Wagner et al., 1999) (Supplementary Figure 2).

Change in SWE is the only measure that significantly correlates with change in LOTS sensitivity (Figure 3A; see Bruno, Zumberge, Manis, Lu, & Goldman, 2008, for a related finding in adults). The second strongest correlation is with change in word identification (WJ-III) (Woodcock, 1987), an untimed measure that also estimates single-word reading skill. No significant correlations were detected with measures of phonological processing or rapid automatic naming. This result is important because it provides a functional interpretation of our findings, targeting the development of speeded sight word reading as the one most directly related to the development of cortical sensitivity to visual word forms in the LOTS.

Individual children's ROTS word form sensitivity showed no increase over time ($p > .3$); the same analysis of individual children in LOTS showed an increasing sensitivity ($p < .05$, two tailed). Change in ROTS did not correlate significantly with any cognitive skills (Figure 3B). The lateralization of the correlation to the LOTS suggests that the developmental effect is not explained by general brain development or instrumental factors.

Figure 3. Correlations between behavioral change and cortical sensitivity change. Horizontal bars represent correlations between the change in sensitivity in LOTS (left) or ROTS (right) and the change in several cognitive measures. Error bars are computed by bootstrapping. Dark gray areas denote uncorrected two-tailed significance range of $p > .05$; light gray areas denote Bonferroni corrected two-tailed significance range of $p > .05$. Only the correlation between LOTS change and SWE change is significant with Bonferroni (or FDR) correction for multiple comparisons. Word identification is close to significance, but none of the other correlations are significantly different from zero. Insets show the location of the cortical regions in a single data set (S8, f, at age 12). Mean (\pm SD) MNI coordinates of ROI across individuals: LOTS ($n = 28$), $[-49, -65, -9]$ ($\pm 5, 8, 5$) and ROTS ($n = 25$), $[46, -67, -12]$ ($\pm 6, 7, 4$).



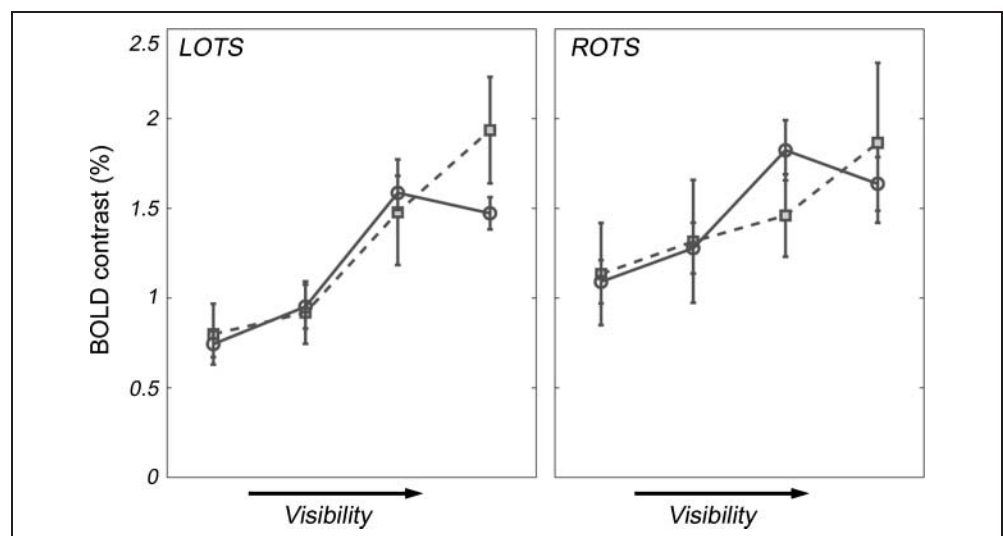
Differences between Good and Poor Readers' Word Visibility Curves

Although our study did not target dyslexia specifically, a few of the participants may be classified as poor readers (Basic Reading [WJ-III] age standardized score under 90, a criterion used by many, e.g., Shaywitz et al., 2007). To examine the difference between the word visibility curves in poor and good readers, we grouped the curves by reading scores (Figure 4). In both the LOTS and ROTS, good readers have a sigmoidal word visibility curve, with a slight signal drop-off at the most visible condition. Poor readers, on the other hand, have a monotonically ascending response

with word visibility. A significant difference is measured only for the most visible condition in the LOTS, wherein poor readers' responses exceed those of good readers ($t(31) = 2.35, p < .05$, participants who yielded more than one value in the group were counted only once).

In a subtraction design that only compared fully visible words to scrambled words, one would conclude that poor readers activate LOTS more than good readers. Given the full curves, it is clear that LOTS responses in good and poor readers are similarly modulated by visibility, except poor readers' responses do not saturate. Rather, poor readers appear to assign more and more cortical resources to word processing, even when the word

Figure 4. OTS word visibility curves grouped by reading skill. Dashed lines with square symbols depict mean curves for poor readers (basic reading standardized score < 90 , mean age = 10.85, $n = 7$ participants contributing 14 curves); full lines with circular symbols depict mean curves for good readers (basic reading standardized score ≥ 100 , mean age = 10.76, $n = 26$ participants contributing 58 curves). Error bars are calculated as SEM, wherein n is the number of unique participants (which is more conservative than the number of observations).



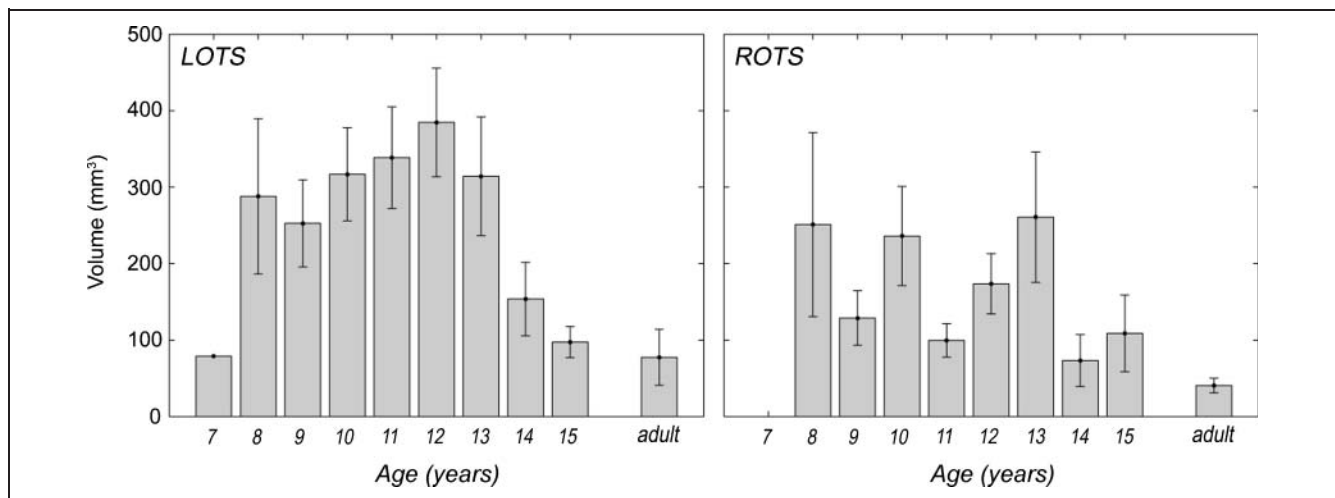


Figure 5. OTS activation volume varies by age. Bars show mean cluster volume for the word visibility contrast (two most visible word conditions vs. two least visible word conditions) for each age group ($n = 1, 6, 11, 21, 28, 25, 18, 6, 4, 9$, respectively) in LOTS and ROTS. Adult cluster volume calculated on data from Ben-Shachar et al. (2007). Error bars represent *SEM*.

form is completely visible and easy to extract. Good readers, on the other hand, have a more efficient curve, with response levels leveling out or even dropping when the words are fully visible.

Age-related Change in OTS Volume

Developmental changes in the volume of fMRI activations have been recorded in the domain of face processing (Golarai, Liberman, Yoon, & Grill-Spector, 2010; Golarai et al., 2007; Scherf, Behrmann, Humphreys, & Luna, 2007), but to our knowledge, there are no data on parallel changes in the VWFA. To assess age-related differences in the volume of the OTS, we calculated visibility contrast maps in each individual for each measurement (see Supplementary Figures 6–7C). We counted the activated voxels within the LOTS and ROTS, provided that they defined a coherent cluster on the high-resolution volume anatomy ($p < 10^{-3}$, uncorrected). Voxel counts were translated into volume estimates and averaged per age group (Figure 5). Adult estimates were derived in the same way from published data (Ben-Shachar et al., 2007b) using the same paradigm in nine adults (five men and four women) ages 23–53.

The volume analysis shows that, although children reach adult sensitivity in LOTS around 12 years, further changes in their responses are evident until adolescence, reaching adult cluster volume in LOTS around the age of 15. However, these changes in volume are not monotonic with age but take the form of an inverted U curve: volume increases until about age 12 and then decreases from age 13 to 15 years.

We followed up on this surprising finding with a second longitudinal analysis of the change in the activated cluster volume in LOTS. For each individual, we calculated the year-to-year difference in cluster volume. We present these volume changes in Figure 6, grouped by age.

The longitudinal analysis reveals positive change in LOTS cluster volume in 7- to 9-year-olds, no change in 10- to 12-year-olds, and a negative change in 13- to 15-year-olds. The difference in volume change between the younger and the older age groups was significant ($t(12) = 2.68$, $p < .05$). These findings provide additional evidence that the cluster activated by word visibility in LOTS first grows and then shrinks, reaching adult size around 15 years.

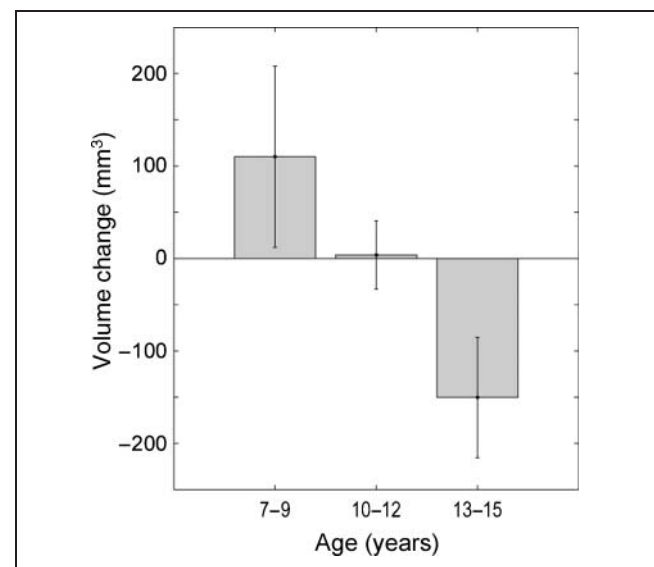


Figure 6. Longitudinal change in LOTS volume. Bars show mean year-to-year change in LOTS cluster volume, calculated as the difference (LOTS volume in year n) – (LOTS volume in year $n - 1$). Activation clusters captured as in Figure 5. Positive change indicates volume growth; negative change indicates volume decrease. Data are grouped by age, with 7- to 9-year-old children showing positive change (year-to-year increase in LOTS volume), 10- to 12-year-olds showing no change, and 13- to 15-year-olds showing negative change (year-to-year decrease in LOTS volume).

Word Visibility Functions in V1 and PPC

We focus on the LOTS, the approximate anatomical location of the VWFA (Cohen, Dehaene, Vinckier, Jobert, & Montavont, 2008; Cohen, Jobert, Le Bihan, & Dehaene, 2004; Cohen et al., 2000, 2002, 2003) because of its significance in the literature. We further examined responses in other cortical regions. For an analysis of adjacent areas along the ventral surface, see Supplementary Appendix I. As another control, we measured responses in foveal V1. There is a response to all of the visual stimuli, but (a) there is no increase in the response with word visibility and (b) no change in the word visibility neurometric function over the years (Supplementary Figure 3). In bilateral PPC (Cohen et al., 2008; Vinckier et al., 2006), we found increasing responses to word visibility (Supplementary Figure 4). In left PPC, but not in right PPC, there are cross-sectional differences between the youngest and the older age groups. None of the correlations between cortical and behavioral changes in PPC pass significance of conventional statistical tests (Supplementary Figure 5). A nonparametric bootstrap test reveals a nonspecific pattern of nonzero correlations, including correlations with phonological and speeded word reading measures (Supplementary Figure 5, error bars). This cognitive profile, together with the early developmental difference found in LPPC, fit well with recent accounts viewing this dorsal region as a serial attention spotlight applied to the letter-by-letter analysis of degraded written words (Cohen et al., 2008). It is possible that this serial strategy develops early and then phases out as more adult recognition strategies are acquired.

DISCUSSION

Cortical sensitivity to written words increases with age in 7- to 15-year-old children within the posterior LOTS, a region which matches the anatomical characterization of the VWFA. Furthermore, a longitudinal analysis of the change in a child's LOTS word sensitivity quantitatively matches the change in an individual's ability to quickly read common "sight words." This developmental pattern is functionally and anatomically specific: it is not found with other cognitive measures, such as rapid letter naming or phonological skills, nor in the ROTS.

Our findings fill in the developmental trajectory for the enhanced sensitivity to visual word forms previously described in adult good readers (Ben-Shachar et al., 2007b; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999). We demonstrate rising response curves to words in children as young as 7 years, in agreement with recent reports in German readers (Brem et al., 2010). Whereas many previous developmental studies focus on **response selectivity** for words compared with other categories (Brem et al., 2006, 2009, 2010; Turkeltaub et al., 2003), our study follows up on the development of **perceptual sensitivity** to word visibility. Developmental changes in LOTS show up as increased responses to word stimuli embedded in noise.

This finding supports earlier claims on the basis of data from adult dyslexics, suggesting that successful reading relies on the ability to exclude noise effectively (Sperling, Lu, Manis, & Seidenberg, 2005).

The correlation with the development of SWE supports the role of the LOTS in fast and efficient extraction of visual word forms from noisy background. This aspect of our findings complements a recent fMRI study (Bruno et al., 2008), which demonstrates a similar correlation between LOTS sensitivity to word frequency and SWE in adults. Our finding is important in both its positive and negative implications. On the positive side, this finding supports a role for the LOTS in fast extraction of visual word forms but not in phonological decoding. Many frequent "sight words" included in the SWE subtest have irregular orthography, which may not be read by piece-wise grapheme to phoneme conversion. Thus, this finding can be viewed as supporting the idea that the LOTS learns to process larger orthographic units with time and exposure. On the negative side, the specific correlation with SWE works against an explanation in terms of general cognitive maturation, which would predict an across-the-board correlation with reading and intelligence.

Our findings provide a quantitative benchmark for the typical development of responses in the ventral reading pathways. Although our study mostly focused on typically developing readers, it shows some evidence that poor readers differ in their LOTS word visibility curves (Figure 4). The individual subject approach we presented here can be applied to clinical imaging-based assessment of developing English readers compared with the typical development profile charted here. The longitudinal perspective we offer provides a dynamic and continuous view of reading development beyond a single snapshot provided by child-adult group comparisons. Group comparisons of brain responses preintervention and postintervention are very informative for the neuroscientist, because they imply plasticity within a specific system (Brem et al., 2010; Keller & Just, 2009; Meyler, Keller, Cherkassky, Gabrieli, & Just, 2008; Temple et al., 2003). But for the parent and the educator, the important question is whether an individual's brain can change its properties following intervention and whether this is predictable on the basis of specific anatomical or functional properties. For this purpose, the individual perspective offered by our measurements may be effective. A natural extension of our measurements can trace longitudinal changes in individuals undergoing different types of intervention methods to study the effect of specific training on specific subsystems of reading. For example, a training study combined with longitudinal measurements can decide whether correlations between MT+ contrast responsivity and reading reflect a causal relation or stem from a third variable (Ben-Shachar et al., 2007a).

Our data also reveal developmental changes in the volume of visibility-sensitive voxels within LOTS. The active region within LOTS first grows (between ages 7 and 9 approximately) and then shrinks (after the age of 13) until it

reaches adult levels around age of 15 years (see Figures 5 and 6). These changes differ from the reports in face-sensitive cortex, which present smaller activated clusters in younger age groups compared with adults, with adolescents somewhere in between (Golarai et al., 2007, 2010; Scherf et al., 2007). We interpret the size reduction observed in older children as reflecting the ability to process the same visual stimulus with fewer resources, that is, increased efficiency, along the same lines explaining the differences between good and poor readers (Figure 4). Future studies will examine the generality of this finding beyond a specific task and stimulus paradigm.

Although we believe that longitudinal designs are important in the study of cognitive development, there are several precautions to take before embarking on a new longitudinal imaging study. First, it has been shown that cross-sectional comparisons between age groups provide a good approximation for longitudinal group comparisons results (Coalson, Petersen, & Schlaggar, 2007). Given the technical complexity and slow nature of longitudinal designs, we suggest that a longitudinal design should only be adopted for charting individual cortical change. Second, on the basis of our experience, we suspect that annual measurements are spaced too widely. More measurements at shorter time lags would likely allow us to reach similar conclusions at a shorter overall time, which may also reduce attrition.

Limitations

A limitation of fMRI experiments in children is the restricted time available for effective scanning. A parametric design necessarily includes multiple conditions to adequately map the response function. We explained the advantages of this design choice in providing a robust, stimulus-referred, threshold measure of cortical sensitivity, which may be reliably followed over time. A clear disadvantage is that this manipulation does not allow testing hypotheses about the category selectivity of the response or about developmental changes in reading strategies. Furthermore, although we believe the individual perspective is essential in a longitudinal study, this analysis approach necessarily focuses on the changes in response properties of a specific brain region and may miss on more global changes in the spatial distribution of signals or the connectivity between regions (Bitan et al., 2007). Finally, this study focused on the development of reading in typically developing children with a wide range of reading skills. We did not recruit a large group of dyslexics; in fact, only about 10% of our sample can be considered poor readers, similar to the distribution in the overall population. Our results, therefore, provide information about reading development in the general population, with some hints regarding impaired reading (see Figure 4).

Reading requires the participation of a network of neural systems, each dominating different reading processes (Ben-Shachar, Dougherty, & Wandell, 2007; Schlaggar & McCandliss, 2007; Noble & McCandliss, 2005; Shaywitz

et al., 2002). It is likely that the developmental changes we document in the LOTS reflect increased sensitivity to visual inputs delivered through neighboring ventral visual areas such as hV4, VO1, and VO2 (Ben-Shachar et al., 2007; Wandell, Dumoulin, & Brewer, 2007). The process of shaping the responses in LOTS is likely to be guided, initially, by phonological and multisensory integration regions in posterior STS and inferior parietal cortex (Blau et al., 2010; Blau, van Atteveldt, Ekkebus, Goebel, & Blomert, 2009). However, at the time frame measured in this article, the changes in LOTS sensitivity appear to be related to efficient reading of sight words, not to phonological decoding. This suggests that in the later school year LOTS disengages from phonological modulation and becomes efficient at extracting perceptual patterns of words. The proper development of this system is surely relevant in English (Shaywitz et al., 2002), but it may be even more important in other writing systems (van der Mark et al., 2009; Share, 2008; Ziegler, 2006; Siok, Perfetti, Jin, & Tan, 2004; Holopainen, Ahonen, & Lyytinen, 2001; Paulesu et al., 2001). The growth of signals in the LOTS of individual children provides an interesting glimpse of how culturally guided education couples with experience-dependent plasticity to shape both cortical processing and reading development.

Acknowledgments

We thank the participating children and their families for their long lasting commitment and enthusiasm. We are grateful to Jan Ruby, Adelle Behn, Arvel Hernandez, Polina Potanina, Sweta Patnaik, Michael Perry, and Jessica Tsang for their help along the 4 years of recruitment, data collection, and analysis. This research was supported by NIH grant EY015000 and by the Schwab Foundation for Learning. M. B. S. was partly supported by a Marie Curie International Reintegration Grant (DNLP 231029) from the European Commission.

Reprint requests should be sent to Michal Ben-Shachar, The Gonda Multidisciplinary Brain Research Center, Bar Ilan University, Ramat Gan 52900, Israel, or via e-mail: michalb@mail.biu.ac.il.

REFERENCES

- Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *104*, 9087–9092.
- Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K., & Wandell, B. A. (2007a). Contrast responsivity in MT+ correlates with phonological awareness and reading measures in children. *Neuroimage*, *37*, 1396–1406.
- Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K., & Wandell, B. A. (2007b). Differential sensitivity to words and shapes in ventral occipito-temporal cortex. *Cerebral Cortex*, *17*, 1604–1611.
- Ben-Shachar, M., Dougherty, R. F., & Wandell, B. A. (2007). White matter pathways in reading. *Current Opinion in Neurobiology*, *17*, 258–270.
- Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., & Buchanan, L. (2006). Tuning of the human left fusiform

- gyrus to sublexical orthographic structure. *Neuroimage*, 33, 739–748.
- Bitan, T., Cheon, J., Lu, D., Burman, D. D., Gitelman, D. R., Mesulam, M. M., et al. (2007). Developmental changes in activation and effective connectivity in phonological processing. *Neuroimage*, 38, 564–575.
- Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., et al. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: A functional magnetic resonance imaging study of dyslexic children. *Brain*, 133, 868–879.
- Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Current Biology*, 19, 503–508.
- Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience*, 16, 4207–4221.
- Brem, S., Bach, S., Kucian, K., Guttorm, T. K., Martin, E., Lyytinen, H., et al. (2010). Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proceedings of the National Academy of Sciences, U.S.A.*, 107, 7939–7944.
- Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., et al. (2006). Evidence for developmental changes in the visual word processing network beyond adolescence. *Neuroimage*, 29, 822–837.
- Brem, S., Halder, P., Bucher, K., Summers, P., Martin, E., & Brandeis, D. (2009). Tuning of the visual word processing system: Distinct developmental ERP and fMRI effects. *Human Brain Mapping*, 30, 1833–1844.
- Brown, T. T., Lugar, H. M., Coalson, R. S., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2005). Developmental changes in human cerebral functional organization for word generation. *Cerebral Cortex*, 15, 275–290.
- Bruno, J. L., Zumberge, A., Manis, F. R., Lu, Z. L., & Goldman, J. G. (2008). Sensitivity to orthographic familiarity in the occipito-temporal region. *Neuroimage*, 39, 1988–2001.
- Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: What have we learned about cognitive development? *Trends in Cognitive Sciences*, 9, 104–110.
- Church, J. A., Coalson, R. S., Lugar, H. M., Petersen, S. E., & Schlaggar, B. L. (2008). A developmental fMRI study of reading and repetition reveals changes in phonological and visual mechanisms over age. *Cerebral Cortex*, 18, 2054–2065.
- Coalson, R. S., Petersen, S. E., & Schlaggar, B. L. (2007). Data derived from a large cross-sectional fMRI study of language development predict longitudinal measurements. Paper presented at the Society for Neuroscience, November 3–7, 2007, San Diego, CA.
- Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., et al. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, 123, 291–307.
- Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *Neuroimage*, 40, 353–366.
- Cohen, L., Jobert, A., Le Bihan, D., & Dehaene, S. (2004). Distinct unimodal and multimodal regions for word processing in the left temporal cortex. *Neuroimage*, 23, 1256–1270.
- Cohen, L., Lehericy, 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, 1054–1069.
- Cohen, L., Martinaud, O., Lemer, C., Lehericy, S., Samson, Y., Obadia, M., et al. (2003). Visual word recognition in the left and right hemispheres: Anatomical and functional correlates of peripheral alexias. *Cerebral Cortex*, 13, 1313–1333.
- Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56, 384–398.
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, 9, 335–341.
- Dejerine, J. (1892). Contribution à l'étude anatomo-pathologique et clinique des différentes variétés de cécité verbale. *Mémoires de la Société de Biologie*, 4, 61–90.
- Dougherty, R. F., Ben-Shachar, M., Deutsch, G., Potanina, P., Bammer, R., & Wandell, B. A. (2005). Occipital-callosal pathways in children: Validation and atlas development. *Annals of the New York Academy of Sciences*, 1064, 98–112.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
- Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the “visual word form area.” *Neuron*, 62, 199–204.
- Glover, G. H. (1999). Simple analytic spiral K-space algorithm. *Magnetic Resonance in Medicine*, 42, 412–415.
- Golarai, G., Ghahremani, D. G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J. L., Gabrieli, J. D., et al. (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nature Neuroscience*, 10, 512–522.
- Golarai, G., Liberman, A., Yoon, J. M., & Grill-Spector, K. (2010). Differential development of the ventral visual cortex extends through adolescence. *Frontiers in Human Neuroscience*, 3, 80.
- Holopainen, L., Ahonen, T., & Lyytinen, H. (2001). Predicting delay in reading achievement in a highly transparent language. *Journal of Learning Disabilities*, 34, 401–413.
- Keller, T. A., & Just, M. A. (2009). Altering cortical connectivity: Remediation-induced changes in the white matter of poor readers. *Neuron*, 64, 624–631.
- Kronbichler, M., Hutzler, F., Staffen, W., Mair, A., Ladurner, G., & Wimmer, H. (2006). Evidence for a dysfunction of left posterior reading areas in German dyslexic readers. *Neuropsychologia*, 44, 1822–1832.
- Maurer, U., Brem, S., Bucher, K., Kranz, F., Benz, R., Steinhausen, H. C., et al. (2007). Impaired tuning of a fast occipito-temporal response for print in dyslexic children learning to read. *Brain*, 130, 3200–3210.
- Maurer, U., Brem, S., Kranz, F., Bucher, K., Benz, R., Halder, P., et al. (2006). Coarse neural tuning for print peaks when children learn to read. *Neuroimage*, 33, 749–758.
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, 7, 293–299.
- Meyler, A., Keller, T. A., Cherkassky, V. L., Gabrieli, J. D., & Just, M. A. (2008). Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: A longitudinal study of neuroplasticity. *Neuropsychologia*, 46, 2580–2592.
- Moore, C. J., & Price, C. J. (1999). Three distinct ventral occipito-temporal regions for reading and object naming. *Neuroimage*, 10, 181–192.
- Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19, 304–330.
- Noble, K. G., & McCandliss, B. D. (2005). Reading development and impairment: Behavioral, social, and neurobiological

- factors. *Journal of Developmental and Behavioral Pediatrics*, 26, 370–378.
- Paulesu, E., Demonet, J. F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al. (2001). Dyslexia: Cultural diversity and biological unity. *Science*, 291, 2165–2167.
- Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage*, 19, 473–481.
- Price, C. J., & Devlin, J. T. (2004). The pro and cons of labelling a left occipito-temporal region: “The visual word form area.” *Neuroimage*, 22, 477–479.
- Scherf, K. S., Behrmann, M., Humphreys, K., & Luna, B. (2007). Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Developmental Science*, 10, F15–F30.
- Schlaggar, B. L., & McCandliss, B. D. (2007). Development of neural systems for reading. *Annual Review of Neuroscience*, 30, 475–503.
- Semel, E., Wiig, E. H., & Secord, W. A. (1995). *Clinical Evaluation of Language Fundamentals 3 (CELF-3)*. San Antonio, TX: The Psychological Corporation.
- Share, D. L. (2008). On the anglocentricities of current reading research and practice: The perils of overreliance on an “outlier” orthography. *Psychological Bulletin*, 134, 584–615.
- Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipito-temporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry*, 55, 926–933.
- Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry*, 52, 101–110.
- Shaywitz, B. A., Skudlarski, P., Holahan, J. M., Marchione, K. E., Constable, R. T., Fulbright, R. K., et al. (2007). Age-related changes in reading systems of dyslexic children. *Annals of Neurology*, 61, 363–370.
- Singer, J. D., & Willett, J. B. (1996). Methodological issues in the design of longitudinal research: Principles and recommendations for a quantitative study of teachers’ careers. *Educational Evaluation and Policy Analysis*, 18, 265–283.
- Siok, W. T., Perfetti, C. A., Jin, Z., & Tan, L. H. (2004). Biological abnormality of impaired reading is constrained by culture. *Nature*, 431, 71–76.
- Sperling, A. J., Lu, Z. L., Manis, F. R., & Seidenberg, M. S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience*, 8, 862–863.
- Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipito-temporal cortex. *Brain*, 122, 2119–2132.
- Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proceedings of the National Academy of Sciences, U.S.A.*, 100, 2860–2865.
- Teo, P., Sapiro, G., & Wandell, B. A. (1997). Creating connected representations of cortical gray matter for functional MRI visualization. *IEEE Transactions on Medical Imaging*, 16, 852–863.
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency*. Austin, TX: PRO-ED Publishing, Inc.
- Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A., & Eden, G. F. (2003). Development of neural mechanisms for reading. *Nature Neuroscience*, 6, 767–773.
- van der Mark, S., Bucher, K., Maurer, U., Schulz, E., Brem, S., Buckelmüller, J., et al. (2009). Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *Neuroimage*, 47, 1940–1949.
- Vigneau, M., Jobard, G., Mazoyer, B., & Tzourio-Mazoyer, N. (2005). Word and non-word reading: What role for the visual word form area? *Neuroimage*, 27, 694–705.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron*, 55, 143–156.
- Vinckier, F., Naccache, L., Papeix, C., Forget, J., Hahn-Barma, V., Dehaene, S., et al. (2006). “What” and “where” in word reading: Ventral coding of written words revealed by parietal atrophy. *Journal of Cognitive Neuroscience*, 18, 1998–2012.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: Pro-Ed Publishing, Inc.
- Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, 56, 366–383.
- Warrington, E., & Shallice, T. (1980). Word-form dyslexia. *Brain*, 103, 99–112.
- Wechsler, D. (2003). *Wechsler intelligence scale for children-4th edition*. San Antonio, TX: Psychological Corporation.
- Wiederholt, J. L., & Bryant, B. (2001). *Gray oral reading test-4th edition (GORT-4)*. Austin, TX: Pro-Ed Publishing, Inc.
- Willett, J. B., Singer, J. D., & Martin, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and Psychopathology*, 10, 395–426.
- Woodcock, R. W. (1987). *Woodcock reading mastery tests-revised*. Circle Pines, MN: American Guidance Service.
- Ziegler, J. C. (2006). Do differences in brain activation challenge universal theories of dyslexia? *Brain and Language*, 98, 341–343.