Development of sensitivity versus specificity for print in the visual word form area

Tracy M. Centanni *, Livia W. King, Marianna D. Eddy, Susan Whitfield-Gabrieli, John D.E. Gabrieli

Massachusetts Institute of Technology, McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences, 43 Vassar St., 46-4033 Cambridge, MA 02139, USA

ABSTRACT

An area near the left lateral occipito-temporal sulcus that responds preferentially to print has been designated as the visual word form area (VWFA). Research suggests that specialization in this brain region increases as reading expertise is achieved. Here we aimed to characterize that development in terms of *sensitivity* (response to printed words relative to non-linguistic faces) versus *specificity* (response to printed words versus line drawings of nameable objects) in typically reading children ages 7–14 versus young adults as measured by functional magnetic resonance imaging (fMRI). Relative to adults, children displayed equivalent sensitivity but reduced specificity. These findings suggest that sensitivity for print relative to non-linguistic stimuli develops relatively early in the VWFA in the course of reading development, but that specificity for printed words in VWFA is still developing through at least age 14.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

The visual word form area (VWFA), a functionally defined region of left occipito-temporal cortex typically near the left occipito-temporal sulcus (Cohen et al., 2002; McCandliss, Cohen, & Dehaene, 2003), is conceptualized as an important component of the reading pathway in the adult brain (Cai, Lavidor, & Brysbaert, 2008; Cai, Paulignan, & Brysbaert, 2010; Cohen et al., 2002; Jobard, Crivello, & Tzourio-Mazoyer, 2003; Pinel & Dehaene, 2010). Specialized visual regions for processing distinct categories of stimuli such as faces (Kanwisher, McDermott, & Chun, 1997), places (Epstein & Kanwisher, 1998), and body parts (Downing, Jiang, Shuman, & Kanwisher, 2001) have been discovered, but the visual word form area is unique in that written words are a recent cultural invention originating only a few thousand years ago. Visual word processing, therefore, must take advantage of existing portions of the object recognition system, perhaps recycling visual regions predisposed to visual stimuli that share visual properties with words (Dehaene & Cohen, 2007). The exact characteristics of this region, however, remain debated in adults and the

nature of its developmental trajectory in children is largely unknown. We addressed both topics in the present study.

There is ongoing debate with regards to both the nature of specialization in the adult VWFA and how to define this region. The definition of the VWFA's location is often based on a contrast of words over non-linguistic stimuli (e.g., words > various checkerboards, Cohen et al., 2002). If this region is truly specialized for visual print, it should also show activation for words over other stimulus types, such as nameable objects that, like words, have semantic associations. Findings in regard to preferential activation for words relative to nameable objects have been notably mixed. Several studies using high-resolution fMRI imaging have found left fusiform activation for words relative to objects in the majority of adult participants (Baker, Liu, & Wald, 2007; Kanwisher, 2010; Szwed, Dehaene, & Kleinschmidt, 2011). Another study found that when stimuli were controlled for luminance, contour length, and number of features, word-selective regions were found in not only the VWFA, but also in early visual regions (Szwed et al., 2011).

On the other hand, several studies have failed to observe a difference in regional activation for visual words relative to other meaningful stimuli, such as objects, colors, braille, and spoken words (Kherif, Josse, & Price, 2011; Price & Devlin, 2003; Vogel, Petersen, & Schlaggar, 2014). One study found that activation for words relative to objects did not replicate within participants across different scans, and were interpreted as reflecting false

* Corresponding author at: 43 Vassar Street, 46-4037C, Cambridge, MA 02139, USA.

E-mail address: tmcentanni@gmail.com (T.M. Centanni).

positives (Wright et al., 2008). Further, cross-modal (picture/word) priming reduced activation in the left occipital temporal region by the same magnitude as word/word priming (Kherif et al., 2011).

Importantly, the method used to identify the region of interest in these studies may partially explain the discrepancy in results. The location of the VWFA varies across individuals, and the VWFA is a relatively small region. Therefore, the pattern of results observed when the ROI is defined on a group level or based on prior literature often varies from the pattern obtained when the ROI is defined separately for each individual. When individual ROIs are used, specificity is commonly observed in adult, expert readers (Baker et al., 2007; Glezer & Riesenhuber, 2013). In the current study, we identified regions of interest for each participant individually to characterize VWFA responses that may be blurred when group activations are analyzed.

Broadly, differences across studies may reflect paradigms that highlight *sensitivity* for print (activations relative to lower-order, meaningless visual stimuli or nonverbal stimuli like faces) versus *specificity* for print (activations relative to potentially meaningful stimuli such as pronounceable non-word letter strings or line drawings of recognizable objects). The evidence for differential *sensitivity* is consistent, such that illiterate adults learning to read develop a VWFA (increased response to words relative to checkerboards and faces; Dehaene et al., 2010). Further, there is evidence that the VWFA is *specific* for print, especially when individual regions of interest are utilized (Baker et al., 2007; Glezer & Riesenhuber, 2013; Vinckier, Dehaene, Jobert, & Dubus, 2007).

Although several studies have investigated either VWFA sensitivity or specificity in children learning to read, the relative developmental trajectories of these two response patterns and whether they develop separately or in tandem are currently unknown. Studies comparing sensitivity to print in left occipito-temporal cortex in children versus adults have reported greater activation in 10-year-old children than adults for words relative to fixation (Olulade, Flowers, Napoliello, & Eden, 2013) and equivalent activation in 4-year-old children and adults for individual letters versus faces (Cantlon, Pinel, Dehaene, & Pelphrey, 2011). These studies suggest that sensitivity to print in brain regions in or around the VWFA matures early in development.

Conversely, specificity in response to print in the same region appears to have a protracted development through at least adolescence and perhaps into young adulthood. Adults, but not young children (4 years old), exhibited specific responses to letters relative to numbers (Cantlon et al., 2011). Older children actually exhibited greater specificity than adults for words than false fonts, but this finding is difficult to interpret in that adults exhibited no specificity in VWFA-region response to print (Olulade et al., 2013). Orientation specificity in response to letters also suggests a protracted developmental trajectory for specificity in children. Objects can be identified regardless of orientation, but letter orientation can be essential to its identity (e.g. "b" vs "d") and novice readers often reverse letters when writing. Thus, orientation specificity likely develops over the course of reading experience. Indeed, there is evidence that adult experience with letter orientation can even alter object perception: literate adults, but not illiterate adults, display behavioral sensitivity to the orientation of familiar objects (Kolinsky & Fernandes, 2014). In regards to reading development, children ages 5–12, despite years of reading experience, show reduced specificity for normal letters relative to mirror-reversed letters by both fMRI measures (including in the VWFA) and ERP measures (Blackburne et al., 2014).

The prior studies overall suggest that *sensitivity* for print in the VWFA develops relatively early in the process of learning to read, whereas *specificity* in for print in the VWFA only emerges with adult expertise. No prior study, however, has examined both sensitivity and specificity in a single group of children with individually defined VWFAs. This precludes evidence for or against a possible dissociation in the development of sensitivity and specificity in the VWFA. Therefore, we tested two related hypotheses. First, we asked whether adults exhibit *specificity* in the VWFA for printed words relative to nameable line drawings. Second, we tested the hypothesis that the development of the VWFA's responsiveness to printed words occurs differentially for *sensitivity* versus *specificity* for print.

In the present study, typically reading children ages 7–14 and young adults viewed words, line drawing of nameable objects (with object names matched to the words), faces, and visual noise displays while undergoing functional magnetic resonance imaging (fMRI). In order to minimize the influence of higher-order language abilities that increase with age, participants performed an incidental task in which they responded only to occasional green-colored stimuli. VWFA was defined as cortical area within fusiform gyrus that significantly responded to the words > visual noise contrast. *Sensitivity* for print was operationalized as greater activation for words than faces. *Specificity* for print was operationalized as greater activation for words than line drawings of nameable objects.

1.1. Methods

1.1.1. Participants

Participants were 17 children (10 male; mean age 10.27, age range 7–14) and 22 adults (11 male; mean age 22.34; age range 18–31) with no history of reading difficulty and were recruited from the university and surrounding community. Participants were required to have been exposed to English from birth, and not to have been exposed to any other language before the age of two. In order to verify that reading ability was typically developing, participants were required to have scored above an 85 standard score on at least three of the following four tests: the Word Identification (Word ID) and Word Attack (WA) subtests from the Woodcock Reading Mastery Test (WRMT; Woodcock, 1998), and the Sight Word Efficiency (SWE) and Phonetic Decoding Efficiency (PDE) subtests from the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999). Participants also had to score above a standard score of 85 on the Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 1990). All participants were right-handed. Age and reading score measures are reported in Table 1.

Informed consent for participation in the study, approved by the Massachusetts Institute of Technology Institutional Review Board, was obtained from all adult participants. For children, parental consent was obtained and children provided verbal assent. Adults were compensated for their participation and children received gift cards to a bookstore for participating.

Table 1
Behavioral characteristics of both groups (means ± standard deviation).

	Adults	Children	p value
Age	22.34 ± 3.26	10.27 ± 1.78	4.44e−16[***]
WRMT Word Attack (raw)	38.41 ± 2.87	35.71 ± 4.58	0.03[*]
WRMT Word Attack (SS)	104.09 ± 9.16	117.12 ± 14.18	0.001[**]
WRMT Word ID (raw)	99.41 ± 2.84	82.47 ± 9.94	4.25e−9[***]
WRMT WORD ID (SS)	105.82 ± 5.73	117.00 ± 10.23	1.08e−4[***]
TOWRE SWE (raw)	97.27 ± 6.71	78.12 ± 10.87	5.62e−8[***]
TOWRE SWE (SS)	102.09 ± 10.29	117.53 ± 12.09	1.18e−4[***]
TOWRE PDE (raw)	55.45 ± 5.98	45.35 ± 7.09	2.42e−5[***]
TOWRE PDE (SS)	102.77 ± 10.13	117.65 ± 9.31	3.45e−5[***]
KBIT Nonverbal (raw)	41.27 ± 2.88	37.47 ± 4.94	0.005[*]
KBIT Nonverbal (SS)	113.68 ± 10.74	123.53 ± 12.54	0.01[*]

[*] p < 0.05.
[**] p < 0.01.
[***] p < 0.001.

1.1.2. Stimuli

Stimuli consisted of words, faces, line drawings of objects, and visual noise/scribbles (196 each). To control for low-level visual characteristics such as contour structure and spatial frequency, stimuli were run through a custom Matlab script that reconstructed the images as dot patterns (Fig. 1A). Specifically, this script removed the clean corners and line connections characteristic of letters so that the contour of the print stimuli more closely matched the contours of the visual noise stimuli (comprised of randomly drawn line segments). Further, by recreating the stimuli as dots, differences in spatial frequency caused by line weights or proximity were eliminated. This script is unpublished but available upon request.

Words were nouns ranging from 3 to 8 letters long (avg = 4.6). Average hyperspace analog to language (HAL) frequency according to the English lexicon project was 27,670 ± 124,497 (Lund & Burgess, 1996). Statistics for two words, 'yoyo', and 'bagel' were not available and thus were not included in the average. All stimuli were divided into two matched lists that were counterbalanced between participants. The words in one list were the names of the line drawings presented in the other list, and vice versa. Black and green versions of all stimuli were created for the task (described below). Stimuli were presented in a box that subtended roughly 4 degrees visual angle.

1.1.3. Procedure

In each trial, participants were presented with the stimulus for 200 ms, followed by 800 ms of a blank screen (Fig. 1B). Stimuli were presented in a block design fashion, each consisting of 14 trials (blocks of 14 s). To control for attention effects on activation in the VWFA (Blackburne et al., 2014), participants were instructed to press the response button anytime a stimulus was green instead of black, which occurred one or two times per block. Between each

block, a cartoon alien flashed on screen for two seconds. Because this paradigm was also run with children, participants were told that the experiments were an attempt to teach the alien about color. Participants were scanned in this experiment for two runs of 4 min and 26 s each. In the two runs combined, there were seven blocks of each stimulus type, plus six fixation blocks.

1.1.4. FMRI acquisition parameters

FMRI scanning took place at the Athinoula A. Martinos Imaging Center at McGovern Institute for Brain Research at MIT. Imaging was performed using a Siemens 3T MAGNETOM Trio, A Tim System, (Siemens Medical Solutions, Erlangen, Germany) and a commercial Siemens 32 channel head coil. High-resolution structural whole-brain images were acquired using a T1-weighted anatomical scan with motion correction (176 slices per slab; 1 mm isotropic voxel size; TR = 2530 ms; TE = 1.64 ms, FoV = 224 mm, flip angle = 9°; Tisdall et al., 2012).

Functional data were collected using a gradient echo T2*-weighted EPI sequence sensitive to the BOLD contrast (2 mm isotropic voxel size; TR = 2; TE = 30 ms; 216 slices, FoV = 256 mm, flip angle = 90°). The gradient-echo EPI images were acquired with PACE (Thesen, Hied, Mueller, & Schad, 2000), an online motion correction algorithm that minimizes movement related artifact by adjusting the system gradients and the acquisition field of view between one whole brain acquisition and another based on participant movement. Slices were placed at an oblique orientation parallel to the AC-PC line. Because whole brain coverage was not always possible with this scan resolution, we made sure that the lowest part of the occipital lobe and the bottom part of the temporal lobe in the left-hemisphere (including the temporal pole) were covered. We collected images from 39 participants for a total of 8424 functional images.
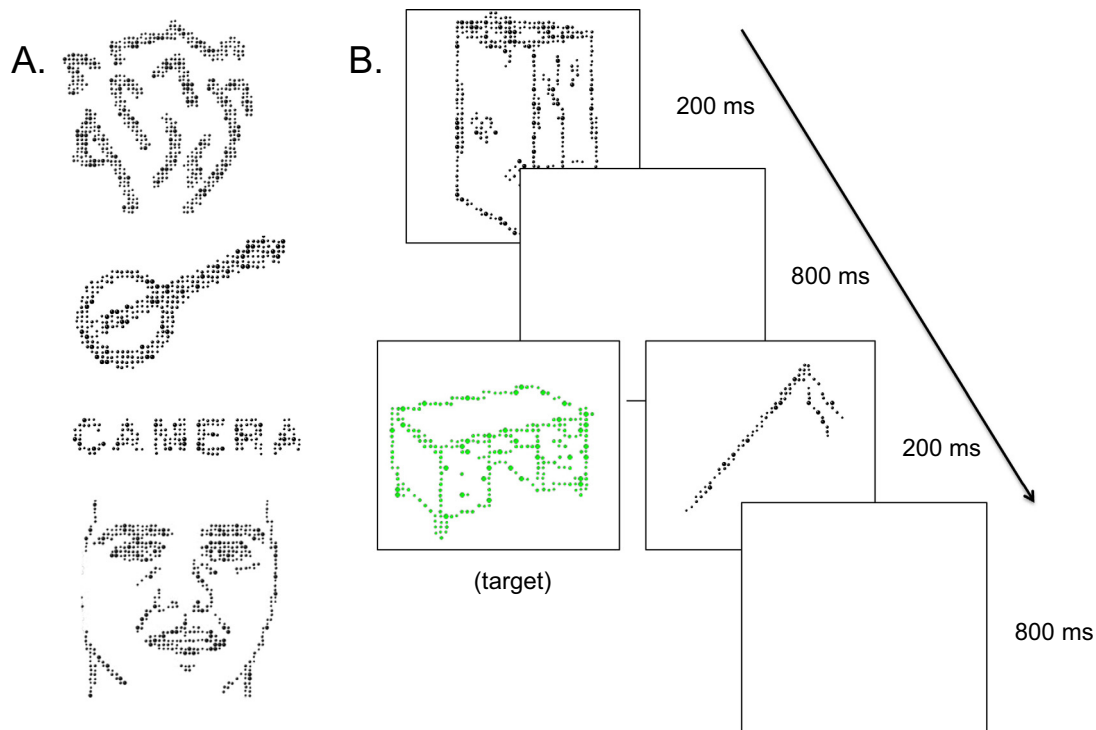

Fig. 1. *Task schematic and stimulus examples.* A. Examples of each perceptually matched stimulus type: noise, objects, words, and faces (top to bottom). B. Task Schematic. Stimuli were presented in blocks of 14 trials each consisting of a single stimulus type. Stimuli appeared for 200 ms each, followed by a blank screen for 800 ms. Participants were instructed to press a button anytime the stimulus was green instead of black (approximately 10% of trials). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

1.1.5. FMRI analysis

The analysis was performed with SPM12 (Wellcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm/software/spm12/), FreeSurfer (Fischl et al., 2002, 2004), Artifact Rejection Toolbox (ART; http://www.nitrc.org/projects/artifact_detect) and Advanced Normalization Tools (ANTS; Klein, Andersson, & Ardekani, 2009) using Nipype and bash scripts for workflow design and execution (Gorgolewski et al., 2011). Functional images were realigned to the mean image and then smoothed with a 4 mm FWHM Gaussian kernel. The functional image co-registration to the 3D anatomical was performed in Freesurfer using a surface based registration algorithm (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999). Structural and functional images were normalized to the Montreal Neurological Institute (MNI) space using ANTS (Klein et al., 2009). Data were high pass filtered with 128 seconds (128 s) cutoff. In the first level analysis, each condition was convolved with a canonical HRF. A one-lag autoregression (AR(1)) model was used to correct for serial (i.e., temporal) autocorrelations.

The ART toolbox was used to detect motion outliers and calculate the degree of stimulus-correlated motion. Movement parameters were included in the individual participant's design matrix for a total of 7 measurements: x, y, z, roll, pitch, and yaw and a composite motion estimate. An image was defined as an outlier if head displacement deviated from the previous time point by more than 1 mm or whose average signal intensity deviated from the series average by more than 3 standard deviations (Blackburne et al., 2014). These outlier measurements were then included as nuisance regressors (consisting of "0s" for all time points and a "1" for the artifactual time point) in the subject-level General Linear Models. In addition, to ensure the number of outliers did not affect the results, analyses were run with all participants as well as with a subset of participants who were matched on the number of outliers.

1.1.6. Stimulus correlated motion and number of artifacts

Children averaged 18.06 ± 4.08 rejected time points as defined in the previous section across both runs of the task, while adults averaged 3.77 ± 0.62 rejected time points (unpaired two-tailed t-test, t(37) = 4.04, p < 0.001). Children and adults did not differ significantly in the amount of stimulus-correlated motion (children had 0.126 ± 0.03 versus 0.111 ± 0.01 in adults; unpaired two-tailed t-test; t(37) = 0.64, p = 0.52).

1.1.7. Characterization of activations and statistical analyses

A mask image (created in WFU PickAtlas; http://fmri.wfubmc.edu/software/PickAtlas) was used to restrict analysis to the left VWFA, defined by previous research on activation to words > checkerboards (region of interest/ROI was a sphere with a 15 mm radius, centered around MNI coordinates [−41 −57 −16]) (Cohen et al., 2002). Each individual's VWFA was defined using words > noise (voxel-wise uncorrected; p < 0.05) and was the closest cluster to the center of the VWFA search area (Fig. 2, white sphere). In participants who did not have significant activation within the VWFA search area, we conducted separate analyses restricted to left fusiform gyrus (Fig. 2, gray) to maximize statistical power in the brain regions of *a priori* interest (as in Brem, Bucher, Halder, & Summers, 2006; Brem, Halder, & Bucher, 2009; Olulade et al., 2013; Van der Mark, Bucher, Maurer, & Schulz, 2009; Vinckier et al., 2007). For participants without a left hemisphere cluster, we identified a right hemisphere cluster (using the same words > noise contrast and voxel-wise threshold of p < 0.05) in each participant within right fusiform gyrus. Three children and one adult had no significant clusters at this threshold, so the p-value was raised until a word-responsive cluster could be identified for each participant (p < 0.10 for one child and p < 0.15 for two children and one adult). Thus, 33 participants (19 adults and 14 children) had significant clusters in the left hemisphere and 6 participants (3 adults and 3 children) had significant clusters in the right hemisphere. The distribution of left versus right loci did not differ significantly between adults and children (chi-square test; $x^2 = 0.12$, p = 0.73; Table 3).

There was no difference in the strength of the response to word stimuli in individuals with a left hemisphere VWFA vs. individuals with a right hemisphere VWFA (adults; t(20) = 0.35, p = 0.73, children; t(15) = 1.35, p = 0.20). In addition, previous studies have documented the presence of right hemisphere VWFA clusters in small numbers of participants (Baker et al., 2007; Cohen et al., 2002). Therefore, we combined data across hemispheres from final sample of 22 adults (19 with a left VWFA, 3 with a right VWFA) and 17 children (14 with a left VWFA and 3 with a right VWFA).
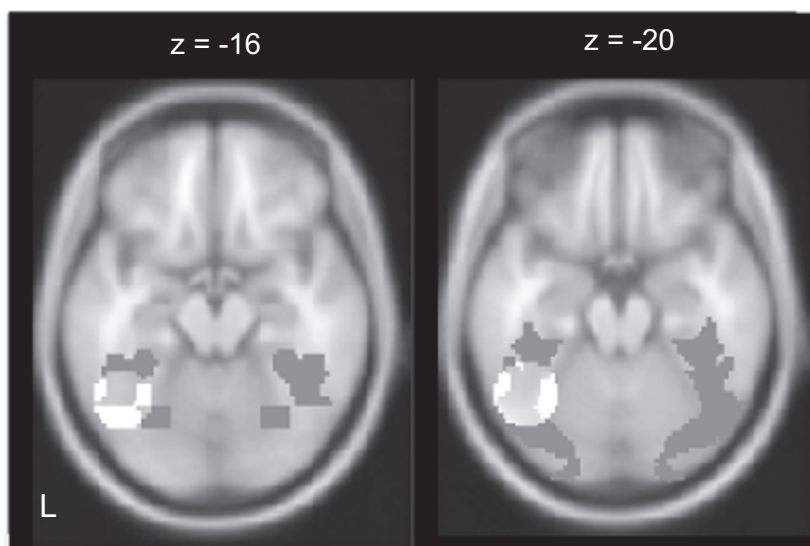

Fig. 2. *Search regions used to define VWFA and right hemisphere regions of interest (ROIs).* VWFA search area was a 15 mm radius sphere centered at [MNI: −41 −57 −16], which is the average peak of VWFA as defined by Cohen et al., 2002 (white circle). Left and right fusiform gyri were used to locate VWFA in individuals without a cluster in the VWFA search space and to evaluate lateralization of word responses (gray areas).

Table 2
Accuracy of in scanner task (means ± standard deviation).

	Adults	Children	p value
Words			
Reaction time (ms)	536.82 ± 59.02	645.56 ± 45.83	2.7e−7[***]
% Probes detected (hits)	96.46 ± 9.93	88.24 ± 14.94	0.05
Overall accuracy	99.30 ± 1.11	97.60 ± 3.36	0.03[*]
Objects			
Reaction time (ms)	534.00 ± 72.55	639.36 ± 54.98	1.5e−5[***]
% Probes detected (hits)	95.96 ± 8.07	88.24 ± 13.30	0.03[*]
Overall accuracy	99.49 ± 0.88	98.08 ± 1.97	0.005[**]
Noise			
Reaction time (ms)	522.08 ± 73.47	635.69 ± 60.87	1.0e−5[***]
% Probes detected (hits)	95.45 ± 6.56	90.20 ± 12.35	0.09
Overall accuracy	99.44 ± 0.82	98.32 ± 3.16	0.12

[*] $p < 0.05$.
[**] $p < 0.01$.
[***] $p < 0.001$.

Table 3
Location and size of words > noise cluster for each participant. Coordinates are reported in MNI space. All ROIs were defined using a voxel-wise threshold of $p < 0.05$ unless noted as follows: *voxel-wise $p < 0.1$, +voxel-wise $p < 0.15$.

Group	Participant	# voxels	x	Y	z	Hemisphere
Adults	1	21	−31	−54	−25	Left
	2	27	−41	−25	−30	Left
	3	5	−42	−50	−16	Left
	5	122	−36	−59	−11	Left
	6	17	−37	−45	−30	Left
	7	157	−39	−45	−24	Left
	8	25	−29	−60	−13	Left
	9	168	−40	−43	−23	Left
	10	142	−44	−21	−33	Left
	11	77	−36	−51	−19	Left
	12+	6	−29	−56	−19	Left
	13	16	−37	−63	−22	Left
	14	6	−33	−54	−10	Left
	16	42	−48	−58	−22	Left
	17	21	−44	−46	−22	Left
	19	48	−42	−25	−23	Left
	20	180	−30	−67	−26	Left
	21	138	−39	−55	−28	Left
	22	13	−35	−34	−19	Left
	Average:	64.79	−37.47	−47.95	−21.84	
	4	22	33	−37	−15	Right
	15	8	36	−39	−24	Right
	18	10	40	−18	−28	Right
	Average:	13.33	36.33	−31.33	−22.33	
Children	1	134	−42	−62	−24	Left
	3	5	−35	−39	−19	Left
	4	6	−50	−58	−15	Left
	5	10	−44	−49	−10	Left
	6	6	−47	−56	−11	Left
	7	20	−25	−52	−13	Left
	8	143	−33	−50	−23	Left
	9	9	−41	−65	−22	Left
	11	111	−41	−48	−26	Left
	12	10	−47	−50	−33	Left
	13	169	−50	−54	−17	Left
	14	12	−44	−55	−25	Left
	16	48	−34	−48	−23	Left
	17	12	−45	−66	−21	Left
	Average:	49.64	−41.29	−53.71	−20.14	
	2*	7	40	−36	−25	Right
	10+	9	36	−36	−22	Right
	15+	8	36	−45	−12	Right
	Average:	8	37.33	−39	−19.67	

We extracted beta-values from within each individual's ROI (iROI) using FSL (Jenkinson et al., 2012). Beta values were extracted for each stimulus type (using the stimulus > fixation contrasts) and

comparisons across groups were accomplished using repeated measures ANOVA and unpaired *t*-tests. Paired *t*-tests were used for comparisons within groups and were two-tailed unless noted otherwise.

1.2. Results

1.2.1 In-scanner behavioral measures

Overall accuracy (both correct hits and correct rejections), percentage of probes detected (hits), and reaction time for children and adults are reported in Table 2. For reaction time, adults were faster than children main effect of group ($F(1,1) = 32.24$, $p < 0.001$), and there was a trend towards responses varying by stimulus type ($F(1,2) = 2.73$, $p = 0.07$). This trend reflected faster responses, across all participants, for the word condition over the noise condition (paired, two-tailed *t*-test, $t(38) = 2.34$, $p < 0.05$), but no reliable differences between the object and the noise conditions (paired, two-tailed *t*-test, $t(38) = 1.45$, $p = 0.16$), or between the word and the object conditions (paired, two-tailed *t*-test, $t(38) = 0.82$, $p = 0.42$). Critically, there was no interaction between groups and stimulus type ($F(1,2) = 0.2$, $p = 0.82$). For accuracy, although both groups were highly accurate, adults were more accurate than children **main effect of group ($F(1,1) = 5.86$, $p < 0.05$). There were no effects of stimulus type ($F(1,2) = 1.29$, $p = 0.28$) and no interaction between group and stimulus type ($F(1,2) = 0.62$, $p = 0.54$).

1.2.2. Sensitivity and specificity for words within VWFA

There was no main effect of group on the magnitude of the response to the four stimulus types; words, objects, faces, or visual noise ($F(1,37) = 0.01$, $p = 0.96$) but there was a significant main effect of stimulus type ($F(1,37) = 33.01$, $p < 0.0001$; Fig. 3). We evaluated the individual responses to the *sensitivity* contrast by extracting beta values to words compared to faces within each participant's iROI. In a repeated measures ANOVA with factors of group (adults, children) and stimulus type (words, faces), there was a main effect of stimulus type ($F(1,33) = 28.72$, $p < 0.0001$), but no main effect of group ($F(1,33) = 0.15$, $p = 0.70$) and no interaction between group and stimulus type ($F(1,33) = 0.03$, $p = 0.86$). *Post hoc* *t*-tests confirmed the finding that both groups showed greater activation to words compared to faces (paired *t*-tests; adults: $t(21) = 5.28$, $p = 3.1e−5$ and children: $t(16) = 2.78$, $p = 0.01$; Fig. 3).

In contrast to adult-like sensitivity (words > faces), children exhibited reduced specificity (words > objects) compared to adults. In a repeated measures ANOVA with factors of group (adults, children) and stimulus type (words, objects), there was a significant main effect of stimulus type ($F(1,33) = 4.48$, $p = 0.04$), no main effect of group ($F(1,33) = 0.04$, $p = 0.84$), and a significant interaction between group and stimulus type ($F(1,37) = 4.26$, $p = 0.046$). This interaction reflected that adults exhibited significant specificity (paired *t*-test $t(21) = 3.76$, $p = 0.001$), but children failed to exhibit reliable specificity (paired *t*-test; $t(16) = 0.12$, $p = 0.90$; Fig. 3). The degrees of sensitivity and specificity were not significantly different in adults (paired *t*-test; $t(21) = 1.52$, $p = 0.14$), but were significantly different in children ($t(16) = 2.54$, $p = 0.02$). Finally, a sign test for specificity (with the null hypothesis that the median was 0) was significant in adults ($p = 0.004$) but not in children ($p = 0.63$).

The children spanned ages 7–14, a period over which there is considerable growth in reading ability. We examined the relation between age and the measure of VWFA specificity (words > objects) among the children, and there was no significant relation ($r = −0.28$, $p = 0.27$), suggesting a long trajectory for development of word specificity in VWFA.
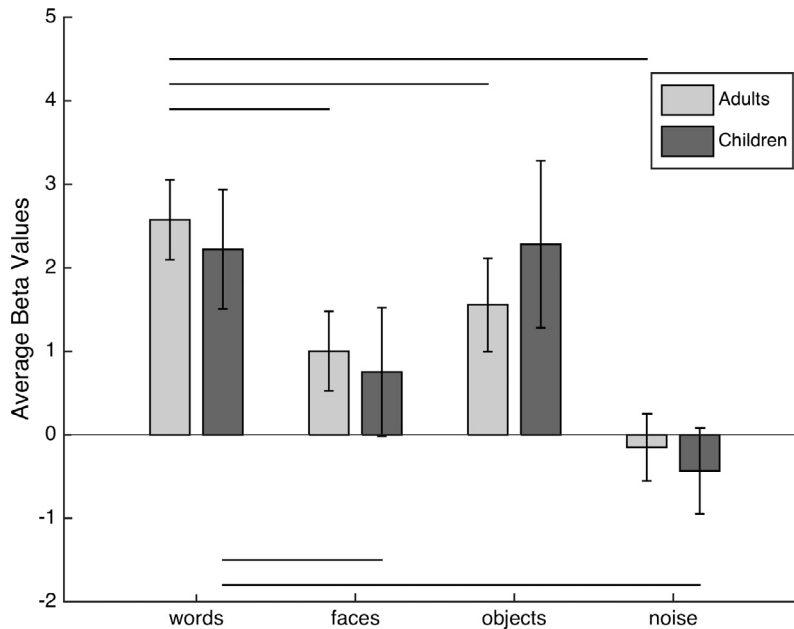
Fig. 3. *Beta values to each stimulus type and each contrast in VWFA iROI.* Horizontal lines represent significant *t*-test at *p* < 0.05; lines above the bars refer to significant *t*-tests for adults, lines below the bars refer to significant *t*-tests for children. Beta values to each stimulus type versus rest. There were no significant group differences in response to any single stimulus type (*ps* > 0.49). Within each group, responses to words were significantly greater than responses to faces. In adults only, word responses were significantly greater than objects.

1.2.3. Relation of VWFA specificity and reading abilities

Because VWFA specificity for words did not arise until adulthood in our sample, we evaluated whether the degree of specificity for words was related to behavioral tasks of reading and pseudoword decoding. We compared *specificity* for words to performance on two untimed measures of reading; the Word Identification and Word Attack subtests of the WRMT. We evaluated both raw and standard scores, as the use of standard scores in these correlations automatically covaries for age. For children, there was a significant and positive relation between specificity for words and standard scores on the decoding measure (WRMT Word Attack; *r* = 0.59, *p* = 0.01), and a trend towards such a relation for the sight word reading measure (WRMT Word ID; *r* = 0.42, *p* = 0.09). Using raw scores, only the WRMT Word Attack measure exhibited a trend towards a positive relation between VWFA specificity and pseudoword decoding ability (*r* = 0.43, *p* = 0.09). For adults, neither Word ID (raw scores; *r* = 0.09, *p* = 0.69, standard scores; *r* = 0.09, *p* = 0.68) nor Word Attack (raw scores; *r* = −0.11, *p* = 0.62, standard scores; *r* = −0.12, *p* = 0.59) was related to specificity.

The absence of any correlation among adults may reflect a more restricted range of standardized scores among the adults. The range of scores was narrower in adults compared to children on Word ID (raw scores: adult range = 10 versus child range = 34; standard scores: adult range = 20 versus child range = 34). Adults also had a narrower range than children on Word Attack (raw scores: adult range = 11 versus child range = 14; standard scores: adult range = 35 versus child range = 52).

1.2.4. Relation of In-Scanner motion on VWFA responses

The adults had significantly less in-scanner motion than the children (*p* < 0.001). To determine whether differential motion affected the results, we repeated the analyses on subgroups of children (N = 10) and adults (N = 11) who were matched on motion. The mean number of motion outliers in these children was 7.72 ± 5.08 vs. 5.82 ± 2.52 in adults (unpaired *t*-test; *t*(20) = 1.12, *p* = 0.28). Consistent with the full sample, there was evidence that

both adults (*t*(9) = 3.26, *p* = 0.005) and children (*t*(8) = 2.26, *p* = 0.03) exhibited significantly higher activation for words compared to faces. Also consistent with the full sample, only adults exhibited a significantly stronger response to words than objects (*t*(9) = 1.94, *p* = 0.04), and there was no significant difference in the children (*t*(8) = 0.39, *p* = 0.35). These results suggest that in-scanner motion did not drive the developmental differences in word specificity seen in the current study.

1.2.5. Relation of IQ and word reading on VWFA responses

The children scored significantly higher than the adults on age-adjusted measures of IQ and word reading, which likely reflects the nature of families who volunteer for their children to participate in neuroimaging research. In order to determine whether this IQ discrepancy could have biased our results, we identified a subset of 18 participants (N = 8 children and N = 10 adults) who were matched on nonverbal IQ (KBIT standard score, 114.73 ± 8.49 in adults vs. 117.44 ± 13.66 in children; *t*(16) = −0.33, *p* = 0.74), untimed word reading (WRMT Word ID standard score, 108.82 ± 3.66 in adults vs. 113.33 ± 7.76 in children; *t*(16) = −1.24, *p* = 0.23) and untimed pseudoword decoding (WRMT Word Attack standard score, 105.27 ± 8.09 in adults vs. 111.00 ± 11.24 in children; *t*(16) = −1.17, *p* = 0.26). Standard scores were used to match participants due to the large age range across groups. Consistent with the full sample, adults exhibited sensitivity to words over faces (paired, one-tailed *t*-tests: *t*(9) = 4.53, *p* = 7.1e−4) and specificity to words over objects (t(9) = 1.92, *p* = 0.04; Fig. 4). Also consistent with the full sample, children exhibited a trend towards sensitivity to words over faces (t(7) = 1.75, *p* = 0.06), but no reliable specificity to words over objects (t(7) = −0.14, *p* = 0.56). These results suggest that the significant differences in assessment scores across groups in the full sample did not bias the findings.

1.3. Discussion

We found evidence that adults exhibited specificity in their VWFA response to printed words relative to semantically and
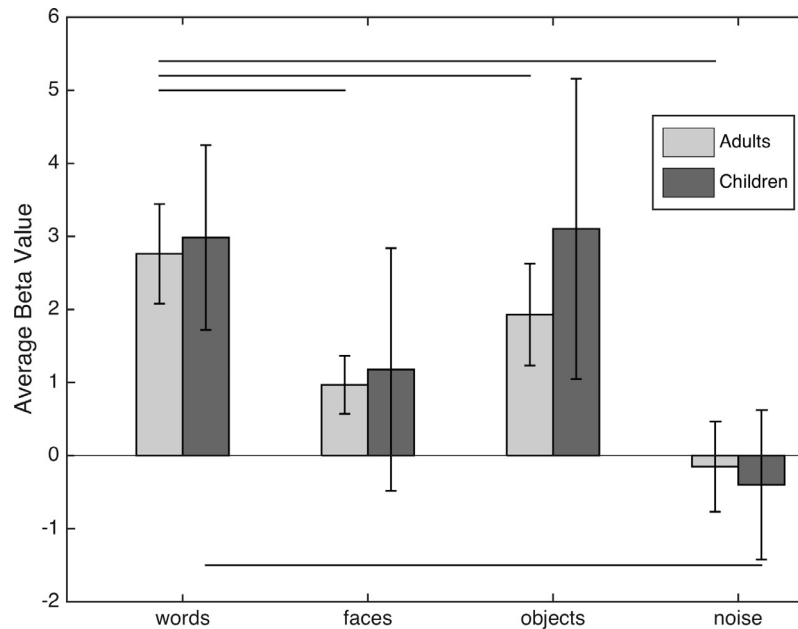
Fig. 4. *Beta values to each stimulus and each contrast in IQ and reading matched sample.* Horizontal lines represent significant *t*-test at *p* < 0.05; lines above the bars refer to significant *t*-tests for adults, lines below the bars refer to significant *t*-tests for children. Beta values to each stimulus type versus rest. Adults showed a significant difference between words and faces, while this did not reach significance in children. Only adults had significantly higher word responses compared to objects.

perceptually equated nameable line drawings. We also found evidence in children ages 7–14 for a dissociation between early development of adult-like VWFA sensitivity for words (words > faces), but a later development of adult-like specificity for words (words > objects). These results suggest that by age 14, VWFA *sensitivity* for words is mature, while VWFA *specificity* for words is still developing.

1.3.1. Specificity for words in the adult VWFA

There have been mixed findings about whether the VWFA shows differential response to words over pictures of objects, with some studies finding such print selectivity (Baker et al., 2007; Kanwisher, 2010; Szwed et al., 2011) and other studies not finding such print selectivity (Kherif et al., 2011; Price & Devlin, 2003; Vogel et al., 2014). A complexity in this field of research is that studies vary considerably in regards to stimuli, tasks, and the process by which VWFA is defined. Different findings may therefore reflect different designs.

In the present study, specificity for words relative to objects in the adult VWFA was clear-cut. Words and objects names were matched on multiple dimensions and counterbalanced across participants. Perceptual properties of words and objects were matched in several ways. Perhaps for these reasons, participants did not differ significantly for reaction time or accuracy for words versus objects, so activation differences cannot be accounted for by behavioral differences. Individual definitions of the VWFA, rather than use of group normalized data, may have contributed to more accurate definition of the VWFA (as suggested by (Glezer & Riesenhuber, 2013)). It is therefore possible that some of the previous controversy regarding the specificity of the VWFA can be attributed to individual differences in VWFA location and differing degrees of overlap with the region of interest used.

1.3.2. Developmental dissociation between sensitivity versus specificity in the VWFA

In the VWFA, both children and adults exhibited similar *sensitivity* for words (versus faces), whereas only adults showed *specificity* for words (versus objects). By examining both sensitivity

and specificity of VWFA responses to words in a single group of children, the present study provides direct evidence for a dissociation between early maturation of sensitivity and protracted maturation of specificity (where maturation is defined as adult-like magnitudes of activation).

This distinction between sensitivity and specificity has been suggested in children as young as age 4, who exhibited adult-like sensitivity to symbols (letters and numbers) relative to faces and objects (shoes), but not adult-like specificity between letters and numbers (Cantlon et al., 2011). Similarly, older children failed to exhibit adult-like specificity for normal versus mirror-reversed letters (Blackburne et al., 2014). The protracted development of specialization of the VWFA parallels evidence for protracted development of the fusiform face area (FFA) (Dundas, Plaut, & Behrmann, 2013, 2014; Golarai et al., 2007; Golarai & Liberman, 2009; Li et al., 2013).

In the present study, we observed a significant and positive correlation between the degree of word specificity and aspects of word reading, especially decoding ability, in children but not in adults. Although specificity correlated with age-adjusted reading skills (standard scores), it did not correlate with age per se, suggesting that specificity was more related to reading skill than age.

1.3.3. Analysis of participant groups

The in-scanner performance and behavioral characterization of the child and adult groups supported the validity of comparing these two groups. As regularly occurs, adults were faster and more accurate than children during task performance (although children were over 97% accurate), but there were no interactions between groups and words versus objects. Therefore, the critical activation difference between words and objects could not be accounted for by in-scanner performance differences.

In regards to behavioral characterization, the adults had higher raw (absolute) scores on all reading and nonverbal IQ measures than children, showing the expected developmental gains from childhood to adulthood. When scores were age-normalized (standardized), both groups were above the population means of 100 on most measures, and the children had significantly higher scores,

relative to their peers, than the adults. The high standardized scores of the children likely reflect the high education level of families around MIT who chose to participate in research. The scores on the standardized tests indicate that the children and adults were average to above average readers who could be appropriately compared to one another (and that any group differences were, if anything, minimized in those comparisons). In any case, these differences did not account for the critical word specificity group difference, as this pattern was also observed in the sub-groups of participants who were matched on IQ, word reading, and decoding.

A critical issue in developmental neuroimaging is the observation that head motion during fMRI is usually highly correlated with age (Satterthwaite, Wolf, & Loughead, 2012). Therefore, care needs to be taken such that fMRI differences are neither manufactured nor masked by differences in head motion (Chai, Ofen, Gabrieli, & Whitfield-Gabrieli, 2014). In the present study, children did move more than adults, but stimulus-correlated motion did not differ across groups. To account for the differences in overall movement, outlier time points were accounted for the general linear model (GLM) in first-level analysis, and we ran additional analyses in two groups of motion-matched individuals to ensure motion did not drive the results. In this motion-matched analysis, we observed the same pattern of similar word sensitivity in children and adults, but word specificity only in adults. This analysis indicates that developmental differences in word specificity are unlikely to reflect developmental differences in motion.

1.3.4. Limitations

Several limitations are evident in the present study. First, development is inferred by a cross-sectional comparison between children and adults. A longitudinal study of children across years would more definitively characterize developmental changes in VWFA specificity and the relation of that specificity to word reading and decoding. A second limitation is that children and families volunteering to come to the MIT campus for imaging research reflects mostly highly educated families, as evidenced by the high standardized word reading scores of the children in the current study. Although we saw a comparable pattern of results in children and adults matched for IQ and two aspects of reading, it is possible that a larger sample size may reveal more complex relationships between IQ, reading skills, age, and brain responses. Future studies should also evaluate word specificity with respect to nameable objects from a sample with a more diverse socio-economic distribution to better reflect the developmental trajectory of the general population. A third limitation is that some participants (4 out of 39 participants) required a more liberal threshold for identification of a VWFA.

1.3.5. Conclusions

In the current study, we found evidence that the adult VWFA exhibits specialization for printed words relative to matched, nameable objects. Children ages 7–14 exhibited adult-like VWFA *sensitivity* for words (relative to faces), but these same children exhibited an absence of the adult-like VWFA *specificity* for words (relative to objects). These findings suggest an age-related dissociation between early-maturing sensitivity and late-maturing specificity for print in the VWFA that may extend into early adulthood.

Acknowledgments

References

Baker, C., Liu, J., & Wald, L. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences, 104*(21), 9087–9092. Retrieved from http://www.pnas.org/content/104/21/9087.short

Blackburne, L. K., Eddy, M. D., Kalra, P., Yee, D., Sinha, P., & Gabrieli, J. D. E. (2014). Neural correlates of letter reversal in children and adults. *PLoS ONE, 9*(5), e98386. http://dx.doi.org/10.1371/journal.pone.0098386

Brem, S., Bucher, K., Halder, P., & Summers, P. (2006). Evidence for developmental changes in the visual word processing network beyond adolescence. *Neuroimage, 29*(3), 822–837. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811905006889

Brem, S., Halder, P., & Bucher, K. (2009). Tuning of the visual word processing system: Distinct developmental ERP and fMRI effects. *Human Brain Mapping, 30*(6), 1833–1844. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/hbm.20751/full

Cai, Q., Lavidor, M., & Brysbaert, M. (2008). Cerebral lateralization of frontal lobe language processes and lateralization of the posterior visual word processing system. *Journal of Cognitive Neuroscience, 20*(4), 672–681. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/jocn.2008.20043

Cai, Q., Paulignan, Y., & Brysbaert, M. (2010). The left ventral occipito-temporal response to words depends on language lateralization but not on visual familiarity. *Cerebral Cortex, 20*(5), 1153–1163. Retrieved from http://cercor.oxfordjournals.org/content/20/5/1153.short

Cantlon, J. F., Pinel, P., Dehaene, S., & Pelphrey, K. A. (2011). Cortical representations of symbols, objects, and faces are pruned back during early childhood. *Cerebral Cortex (New York, N.Y.: 1991), 21*(1), 191–199. http://dx.doi.org/10.1093/cercor/bhq078

Chai, X., Ofen, N., Gabrieli, J., & Whitfield-Gabrieli, S. (2014). Selective development of anticorrelated networks in the intrinsic functional organization of the human brain. *Journal of Cognitive Neuroscience, 26*(3), 501–513. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/jocn_a_00517

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: A Journal of Neurology, 125*(Pt 5), 1054–1069.

Dale, A., Fischl, B., & Sereno, M. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage, 9*(2), 179–194. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811998903950

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384–398. Retrieved from http://www.sciencedirect.com/science/article/pii/S0896627307007593

Dehaene, S., Pegado, F., Braga, L., Ventura, P., Nunes Filho, G., Jobert, A., ... Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science (New York, N.Y.), 330*(6009), 1359–1364. http://dx.doi.org/10.1126/science.1194140

Downing, P., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science, 293*(5539), 2470–2473. Retrieved from http://www.sciencemag.org/content/293/5539/2470.short

Dundas, E. M., Plaut, D. C., & Behrmann, M. (2013). The joint development of hemispheric lateralization for words and faces. *Journal of Experimental Psychology: General, 142*(2), 348–358. http://dx.doi.org/10.1037/a0029503

Dundas, E., Plaut, D., & Behrmann, M. (2014). An ERP investigation of the co-development of hemispheric lateralization of face and word recognition. *Neuropsychologia, 61*, 315–323.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392*(6676), 598–601. Retrieved from http://www.nature.com/nature/journal/v392/n6676/abs/392598a0.html

Fischl, B., Sereno, M., & Dale, A. (1999). Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. *Neuroimage, 9*(2), 195–207. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811998903962

Fischl, B., Salat, D., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., ... Montillo, A. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron, 33*(3), 341–355.

Fischl, B., Salat, D. H., van der Kouwe, A. J. W., Makris, N., Ségonne, F., ... Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage, 23*, S69–S84.

Glezer, L., & Riesenhuber, M. (2013). Individual variability in location impacts orthographic selectivity in the "Visual Word Form Area". *The Journal of Neuroscience, 33*(27), 11221–11226.

Golarai, G., & Liberman, A. (2009). Differential development of the ventral visual cortex extends through adolescence. *Frontiers in Human Neuroscience, 43*. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2831628/>.

Golarai, G., Ghahremani, D., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J., ... Grill-Spector, K. (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. *Neuroscience, 10*(4), 512–522.

Gorgolewski, K., Burns, C., Madison, C., Clark, D., Halchenko, Y., ... Ghosh, S. (2011). Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in Python. *Frontiers in Neuroinformatics, 5*.

Jenkinson, M., Beckmann, C., Behrens, T., Woolrish, T., Woolrish, M., & Smith, S. (2012). FSL. *Neuroimage, 62*(2), 782–790.

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *Neuroimage, 20*(2), 693–712. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811903003434.

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences, 107*(25), 11163–11170. Retrieved from http://www.pnas.org/content/107/25/11163.short.

Kanwisher, N., McDermott, J., & Chun, M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience, 17*(11), 4302–4311. Retrieved from http://www.jneurosci.org/content/17/11/4302.short.

Kaufman, A., & Kaufman, N. (1990). K-BIT: Kaufman brief intelligence test. Retrieved from <http://scholar.google.com/scholar?q=kaufman+brief+intelligence+test&btnG=&hl=en&as_sdt=0%2C22#1>.

Kherif, F., Josse, G., & Price, C. (2011). Automatic top-down processing explains common left occipito-temporal responses to visual words and objects. *Cerebral Cortex, 21*(1), 103–114. Retrieved from http://cercor.oxfordjournals.org/content/21/1/103.short.

Klein, A., Andersson, J., & Ardekani, B. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage, 46*(3), 786–802. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811908012974.

Kolinsky, R., & Fernandes, T. (2014). A cultural side effect: Learning to read interferes with identity processing of familiar objects. *Frontiers in Psychology, 5*, 1224. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4215613/.

Li, S., Lee, K., Zhao, J., Yang, Z., He, S., & Weng, X. (2013). Neural competition as a developmental process: Early hemispheric specialization for word processing delays specialization for face processing. *Neuropsychologia, 51*(5), 950–959. Retrieved from http://www.sciencedirect.com/science/article/pii/S0028393213000481.

Lund, K., & Burgess, C. (1996). Hyperspace analogue to language (HAL): A general model semantic representation. *Brain and Cognition, 30*(3), 5. Retrieved from https://scholar.google.com/scholar?hl=en&q=lund+burgess+1996+hal&btnG=&as_sdt=1%2C30&as_sdtp=#0.

McCandliss, B., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences, 7*(7), 293–299. Retrieved from http://www.sciencedirect.com/science/article/pii/S1364661303001347.

Olulade, O., Flowers, D., Napoliello, E., & Eden, G. (2013). Developmental differences for word processing in the ventral stream. *Brain and Language, 125*(2), 134–145. http://dx.doi.org/10.1016/j.bandl.2012.04.003.

Pinel, P., & Dehaene, S. (2010). Beyond hemispheric dominance: Brain regions underlying the joint lateralization of language and arithmetic to the left hemisphere. *Journal of Cognitive Neuroscience, 22*(1), 48–66. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/jocn.2009.21184.

Price, C., & Devlin, J. (2003). The myth of the visual word form area. *Neuroimage, 19*(3), 473–481. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811903000843.

Satterthwaite, T., Wolf, D., & Loughead, J. (2012). Impact of in-scanner head motion on multiple measures of functional connectivity: Relevance for studies of neurodevelopment in youth. *Neuroimage, 60*(1), 623–632. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811911014650.

Szwed, M., Dehaene, S., & Kleinschmidt, A. (2011). Specialization for written words over objects in the visual cortex. *Neuroimage, 56*(1), 330–344. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811911001261.

Thesen, S., Heid, O., Mueller, E., & Schad, L. (2000). Prospective acquisition correction for head motion with image-based tracking for real-time fMRI. *Magnetic Resonance in Medicine, 44*(3), 457–465.

Tisdall, M., Hess, A., Reuter, M., Meintjes, E., Fischl, B., & van der Kouwe, A. (2012). Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. *Magnetic Resonance in Medicine, 68*(2), 389–399. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/mrm.23228/full.

Torgesen, J., Wagner, R., & Rashotte, C. (1999). Test of word reading efficiency. Retrieved from <https://scholar.google.com/scholar?hl=en&as_sdt=0,22&q=test+of+word+reading+efficiency#0>.

Van der Mark, S., Bucher, K., Maurer, U., & Schulz, E. (2009). Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *Neuroimage, 47*(4), 1940–1949. Retrieved from http://www.sciencedirect.com/science/article/pii/S1053811909005035.

Vinckier, F., Dehaene, S., Jobert, A., & Dubus, J. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron, 55*(1), 143–156. Retrieved from http://www.sciencedirect.com/science/article/pii/S0896627307004503.

Vogel, A. C., Petersen, S. E., & Schlaggar, B. L. (2014). The VWFA: It's not just for words anymore. *Frontiers in Human Neuroscience, 8*, 88. http://dx.doi.org/10.3389/fnhum.2014.00088.

Woodcock, R. W. (1998). *Woodcock reading mastery tests, revised.* MN: American Guidance Service Circle Pines.

Wright, N., Mechelli, A., Noppeney, U., Veltman, D., Rombouts, S., Glensman, J., ... Price, C. (2008). Selective activation around the left occipito-temporal sulcus for words relative to pictures: Individual variability or false positives? *Human Brain Mapping, 29*(8), 986–1000. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/hbm.20443/full.