

An Investigation of Twenty/20 Vision in Reading

Peter J. Hellyer,^{1*} Zoé V. J. Woodhead,^{2*} Robert Leech,¹ and Richard J. S. Wise^{1,2}

¹Division of Experimental Medicine and ²MRC Clinical Sciences Centre, Imperial College London, Hammersmith Hospital, London, W12 0NN, United Kingdom

One functional anatomical model of reading, drawing on human neuropsychological and neuroimaging data, proposes that a region in left ventral occipitotemporal cortex (vOT) becomes, through experience, specialized for written word perception. We tested this hypothesis by presenting numbers in orthographical and digital form with two task demands, phonological and numerical. We observed a main effect of task on left vOT activity but not stimulus type, with increased activity during the phonological task that was also associated with increased activity in the left inferior frontal gyrus, a region implicated in speech production. Region-of-interest analysis confirmed that there was equal activity for orthographical and digital written forms in the left vOT during the phonological task, despite greater visual complexity of the orthographical forms. This evidence is incompatible with a predominantly feedforward model of written word recognition that proposes that the left vOT is a specialized cortical module for word recognition in literate subjects. Rather, the physiological data presented here fits better with interactive computational models of reading that propose that written word recognition emerges from bidirectional interactions between three processes: visual, phonological, and semantic. Further, the present study is in accord with others that indicate that the left vOT is a route through which nonlinguistic stimuli, perhaps high contrast two-dimensional objects in particular, gain access to a predominantly left-lateralized language and semantic system.

Introduction

A lesion disconnecting or destroying left ventral occipitotemporal cortex (vOT) impairs written word recognition (Binder and Mohr, 1992). This observation has been pursued with functional neuroimaging studies. Since the study of Cohen and colleagues (2000), word recognition has become associated with activity in a region of the left vOT, now known as the visual word form area (VWFA). This area has been demonstrated to respond more to orthographically regular letter strings and written words than consonant strings or false font (Cohen et al., 2000; Baker et al., 2007; Reinke et al., 2008; Woodhead et al., 2011).

The local combination detector (LCD) model (Dehaene et al., 2005) proposes that the left ventral visual stream, from primary visual cortex to the posterior occipitotemporal sulcus in the vOT, forms a feedforward, hierarchical model for written word recognition. While the earlier, bilateral cortical components of this hierarchy extract contours and shape and are sensitive to lower-level visual information such as font and retinal location, the more anterior, left-lateralized part extracts more abstract information about written word form. Glezer et al. (2009) presented evidence in favor of the left vOT containing neuronal assemblies that are tuned uniquely to a specific familiar word form, thereby

equating this region with psychological models of reading that incorporate an orthographical input lexicon. On this account, the response of the vOT exemplifies modular specialization consequent upon experience and not genetic predisposition.

In contrast, other studies have provided evidence that activity in the left vOT is observed during tasks other than written word recognition (Price and Devlin, 2003; Yoncheva et al., 2010). Further, phonological or semantic priming can modulate activity in the left vOT (Devlin et al., 2006), which is incompatible with its predominantly feedforward role in word recognition proposed by the LCD model. These observations are compatible with the interactive account of written word recognition, which proposes bidirectional interactions with semantic and phonological processes (Plaut et al., 1996; Behrmann et al., 1998; Patterson and Ralph, 1999; Devlin et al., 2006; Price and Devlin, 2011). Extrapolating from this computational account to physiology, activity in visual association cortex should be modulated by task-dependent feedback effects.

The present study was designed to address these issues. The hypothesis under investigation was that the left vOT connects various classes of visual stimuli to language processes, and its response is to the task (linguistic) and not to the stimulus category (words). The design made use of the fact that although letters and digits are similar in their visual characteristics, and both have an associated phonology, it is only numbers that exist in both orthographical and digital forms. Varying the task demand on numbers offered the opportunity to determine whether activity in the left vOT was task-specific (linguistic vs numerical) and not stimulus category-specific (orthographical vs digital forms). This hypothesis also addresses the notion that, through familiarity, written words automatically activate language-related feedback, resulting in greater activity in the left vOT than other visual stimuli even when words are viewed passively.

Received June 2, 2011; revised July 28, 2011; accepted July 31, 2011.

Author contributions: Z.V.J.W., R.L., and R.J.S.W. designed research; P.J.H. and Z.V.J.W. performed research; P.J.H. and Z.V.J.W. analyzed data; P.J.H., Z.V.J.W., R.L., and R.J.S.W. wrote the paper.

This work was supported by funding from the Medical Research Council (to Z.W.) and Research Councils UK (to R.L.).

*P.J.H. and Z.V.J.W. contributed equally to this work.

This article is freely available online through the *JNeurosci* Open Choice option.

Correspondence should be addressed to Zoé V. J. Woodhead, MRC Clinical Sciences Centre, Imperial College London, Hammersmith Hospital, Du Cane Road, London, W12 0NN, UK. E-mail: zoe.woodhead06@imperial.ac.uk.
DOI:10.1523/JNEUROSCI.2740-11.2011

Copyright © 2011 the authors 0270-6474/11/3114631-08\$15.00/0

Materials and Methods

Subjects. Nineteen right-handed subjects (9 females, mean age: 28.8 years) participated in the study. All participants spoke English as their first language and had no history of significant neurological or psychiatric illness. Participants gave written consent and were checked for contraindications to MRI scanning. Ethical approval was awarded by the Hammersmith, Queen Charlotte’s and Chelsea research ethics committee.

Functional MRI procedures. Magnetic resonance imaging (MRI) data were obtained with a Phillips Intera 3.0 tesla MRI scanner, using an eight-array head coil, and sensitivity encoding (SENSE) with an undersampling factor of 2. Functional MRI (fMRI) used a T2*-weighted gradient-echo echoplanar imaging sequence with a repetition time of 3 s. Whole-brain volumes (48 axial slices; slice thickness, 5 mm; in-plane resolution, 2.5 × 2.5 mm) were acquired in an interleaved ascending order. T1-weighted whole-brain structural images were also obtained for accurate spatial registration. Functional data were acquired in two runs, one with passive tasks and the other with active tasks. Both runs used a block design with a continuous acquisition protocol. Run order was counterbalanced across the participants. Stimuli were presented using MatLab (MathWorks) and the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007). Earplugs and padded headphones were used to protect participants’ hearing during the scanning procedure.

Functional MRI stimuli and design: passive run. There were six experimental conditions, each presented six times within a single run. Blocks were presented in a pseudorandom order to maximize variability of transitions between tasks. The blocks, each of 16 s duration, were separated by a period of fixation lasting 6 s. Stimuli in all conditions were presented at the horizontal and vertical center of the screen.

As depicted in Figure 1, the six experimental conditions in the passive run were as follows: (1) unconnected words (Single Words), (2) words presented in sentences (Connected Words), (3) unconnected number strings presented as digits (Single Digits), (4) number strings presented as ascending or descending sequences (Connected Digits), (5) a false font baseline, and (6) a checkerboard baseline.

The manipulation of connected versus single stimuli was performed for use in another analysis that will be reported elsewhere. The condition with connected words consisted of blocks of narrative text, presented one word at a time, using the rapid serial visual presentation (RSVP) procedure. During the condition with connected digits, stimuli were presented in numerically ascending or descending sequences that were matched in length to the sentences in the connected words condition. The conditions with single words and digits each used the same stimuli as their associated connected condition, but shuffled into a meaningless order. For the purposes of the present study, the conditions with single and connected stimuli were collapsed to provide contrasts of words or numbers relative to the two baseline conditions. Each block contained 38 stimuli, each presented for 400 ms. Words were three to six letters long. Word, digital, and false font stimuli were presented in black, 50 point font on a white background, using RSVP. False font, checkerboard, and digital stimuli were matched to these words for width, height, and, for false font and digits, grapheme frequency. During all conditions, participants were asked to silently view the stimuli without a task demand.

Functional MRI stimuli and design: active run. The run of active tasks used the same block procedure as the passive run, with six repetitions of each condition and a 6 s gap between each block.

The active run contained blocks of six conditions (Fig. 1), as follows: (1) numbers presented as words with a number decision task (ND-

A – Passive Run

Task	Stimulus	Duration	Repeats	Block Length
Passive Reading	Words (Tree, Bird) *	0.4s	40	16s
	Digit Strings (5377, 2854) *			
	False Font (Δζφφ, Δγσδ) *			
	Checkerboard *			

B – Active Run

Task	Stimulus	Duration	Repeats	Block Length
Phoneme Decision (PD)	Digits (6,7,8)	1.6s	10	16s
Number Decision (ND)				
Phoneme Decision (PD)	Number Words (Six, Seven, Eight) *			
Number Decision (ND)				
Scale Detection	False Font (ΣφΞ, Σγψχθ, θφγηδ) *			
	Checkerboard *			

* Matched for Width, Height and Letter Repetition

Figure 1. Experimental design and stimuli for the passive task (A) and the active task (B).

Words), (2) numbers presented as digits with a number decision task (ND-Digits), (3) numbers presented as words with a phoneme decision task (PD-Numbers), (4) numbers presented as digits with a phoneme decision task (PD-Digits), (5) a false-font baseline with an oddball detection task (False-Font), and (6) a checkerboard baseline with an oddball detection task (Checkerboards).

Conditions 1–4 were modeled as a 2 × 2 repeated-measures ANOVA, with the factors stimulus type (orthographical and digital forms) and task type (number and phoneme decision). In each block, a target stimulus was presented for 2 s at the start of the block, followed by 10 visual stimuli, each presented for 1600 ms, during which a response was recorded using a handheld trigger panel held in the participant’s left hand. Accuracy and reaction time data were recorded for each trial.

In the phoneme decision task, the target stimulus was a particular letter sound (e.g., “n as in night”). Participants were required to decide whether the target letter sound was present within a stimulus. The number decision task was based on whether the number was odd or even. These tasks were chosen to selectively engage higher-order verbal or numerical processes while keeping the low-level stimulus properties constant. The false-font and checkerboard baseline tasks used oddball detection, whereby the visual scale of the stimulus pattern varied between two scales (0.5 and 1), and the target was the smaller stimulus. In all tasks, the participants responded using a button trigger panel in their right hand, with a thumb press to indicate that the target was present and a forefinger press if the target was absent. In the numerical and phonological decisions, the ratio of target present:absent was 4:10. During the oddball detection tasks, the ratio of target present:absent was 1:10. The false-font condition has been shown previously to activate reading-related pathways in the left ventral visual stream (Woodhead et al., 2011), which motivated the inclusion of the checkerboard baseline condition.

Stimuli for the ND-Numbers and PD-Numbers conditions were formed from the word representations of the 19 numbers between 1 and 100 with a maximum length of six letters (mean word length: 4.79 letters). Stimuli for the ND-Digits and PD-Digits conditions were formed from the digit representations of the 19 numbers presented in the Word conditions. Stimuli for the False-Font condition consisted of false-font words, matched for length and repetition of individual graphemes/false fonts to

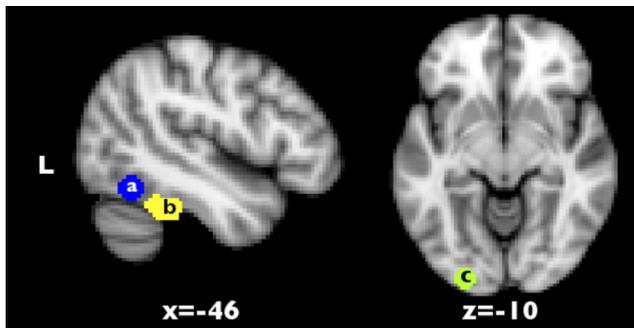


Figure 2. Position of regions of interest within the left ventral visual stream at previously published coordinates for the visual word form area (a), an anterior portion of the ventral visual stream identified by the contrast of reading words (single or connected words) versus false font in the passive task analysis (b), and the left primary visual cortex (c).

the stimuli used in the ND-Words and PD-Words conditions. This was achieved by direct translation into a false font using a custom font generated with GNU FontForge (<http://fontforge.sourceforge.net/>), which executed a direct Roman-to-Greek alphabet correspondence code, as described by Woodhead et al. (2011). Stimuli for the Checkerboards condition consisted of computer-generated checkerboard grids, which were matched in length and height to the stimuli used in the ND-Words and PD-Words conditions.

Functional MRI analysis. Functional MRI data processing was performed using the fMRI Expert Analysis Tool, FEAT (Smith et al., 2004). The following preprocessing was applied to the functional data: motion correction using MCFLIRT (Jenkinson and Smith, 2001), non-brain removal using BET (Smith, 2002), spatial smoothing using a Gaussian kernel of FWHM 7.0 mm, grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with $\sigma = 50.0$ s). Registration of functional images to high-resolution structural images was performed using FLIRT (Jenkinson and Smith, 2001; Jenkinson et al., 2002). Registration from high-resolution structural to standard space was then further refined, using FNIRT nonlinear registration (Andersson et al., 2007a,b).

Time series statistical analysis was performed using FILM with local autocorrelation correction (Woolrich et al., 2001). First-level data analyses matrices were created for each subject in each of the two experiments, using a fixed-effects design. The six conditions within each experiment were modeled as explanatory variables (EVs). The temporal derivative of each EV and the participant's motion parameters were included as covariates of no interest. Contrasts of parameter estimates (COPEs) were calculated for contrasts of interest between EVs, and the COPE data were entered into higher group-level analysis using the FMRIB Local Analysis of Mixed Effects (FLAME) tool. Statistical images were thresholded using a cluster-corrected threshold of $z > 2.3$, $p < 0.05$.

Functional MRI region-of-interest data analysis. Region-of-interest (ROI) analyses were performed using FMRIB's Featquery tool (Smith et al., 2004) to demonstrate the profile of activity across the experimental conditions within the a priori-defined areas of interest.

Spherical ROIs with a diameter of 7 mm were used. The VWFA ROI was located at Montreal Neurological Institute (MNI) coordinates $x = -44$, $y = -58$, $z = -15$, as reported in a meta-analysis by Jobard et al. (2003). An additional ROI in primary visual cortex (V1) was located at the occipital pole and centered on the posterior extent of the calcarine sulcus, $x = -25$, $y = -94$, $z = -8$, consistent with the representation of foveal vision (Serenio et al., 1995; Leff et al., 2000). The anatomical localization of the ROI within V1 was verified using coordinates provided in the Jülich Histological Atlas (Amunts et al., 2000). The locations of the ROIs are shown in Figure 2.

Featquery was used to extract COPE values for each condition versus the resting baseline. For each participant, the 90th percentile voxel value for each ROI was taken rather than the maximum value to exclude from further analysis any voxels with high noise rather than signal.

Results

Passive run

Whole-brain analysis

This analysis was performed to assess the cortical regions of the brain modulated by the type of stimuli in a passive task, that is, when no overt task performance was required of the subjects as they viewed the stimuli. A 2×2 repeated-measures ANOVA, with the levels stimulus type (words and numbers) and continuity (single vs connected stimuli), was performed on the whole-brain fMRI data using FSL software (see Materials and Methods, above). The main effect of continuity was not of interest for the present study and will be presented elsewhere. To interpret the ANOVA results, the main effects were masked with a contrast of all reading conditions versus the checkerboard baseline [(single words + connected words + single digits + connected digits) – checkerboards] or with the opposite contrast of the checkerboard baseline versus the reading conditions. This identified areas that were significantly modulated by task and significantly activated or deactivated by reading relative to the checkerboard baseline.

The main effect of stimulus type is shown in Figure 3A, (top, preferential activity for words; bottom, preferential activity for digits). Unsurprisingly, the distribution of activation was very different for words and numbers. Preferential activity for words over digits was observed in a predominantly left-lateralized network. This included the ventral temporal lobe, first evident ~ 75 mm posterior to the anterior commissure and extending forwards until the signal was lost in susceptibility artifact encompassing the anterior fusiform gyrus (Visser et al., 2010). A peak of activity was observed at MNI coordinates $x = -38$, $y = -42$, $z = -24$ —close (within 19 mm) but not identical to the coordinates published for the usual location of the VWFA (Jobard et al., 2003). Preferential activation for words was also observed in the superior temporal sulcus (STS) and in midline and right frontal cortex. These areas were, with the exception of the right STS, significantly activated relative to the checkerboard baseline (Fig. 3A, orange). The right STS showed no significant difference between reading words and viewing checkerboards (Fig. 3A, yellow), indicating an activation profile of words = checkerboards > digits in this area. A small area of significant deactivation relative to baseline (Fig. 3A, green) was observed in the right primary visual cortex, which can be attributed to the higher levels of visual stimulation in the flashing checkerboard condition.

Preferential activation for digits over words was observed in the bilateral posterior parietal cortices (PPC), the right intraparietal sulcus (IPS) extending forwards toward the postcentral gyrus, and the posterior and anterior extents of the cingulate gyrus. While the more lateral parietal and frontal regions (Fig. 3A, yellow) were similarly activated by the checkerboard baseline, the midline activations mostly represented deactivations relative to baseline. The activation profile of checkerboards > digits > words is consistent with the interpretation that these areas are part of the default mode network (Raichle et al., 2001), which are more strongly active during passive than active tasks. The implication is that the monotony of the checkerboard stimuli compared with reading digits was responsible for this relative difference in activity.

A possible side effect of the use of cluster-based correction for multiple comparisons is that small clusters with a high z -statistic may be deemed insignificant. To confirm that no such omissions had occurred in this analysis, the data were reanalyzed at the group level using a cluster-correction threshold of $z > 3.1$, $p < 0.05$. This did not yield any new significant clusters.

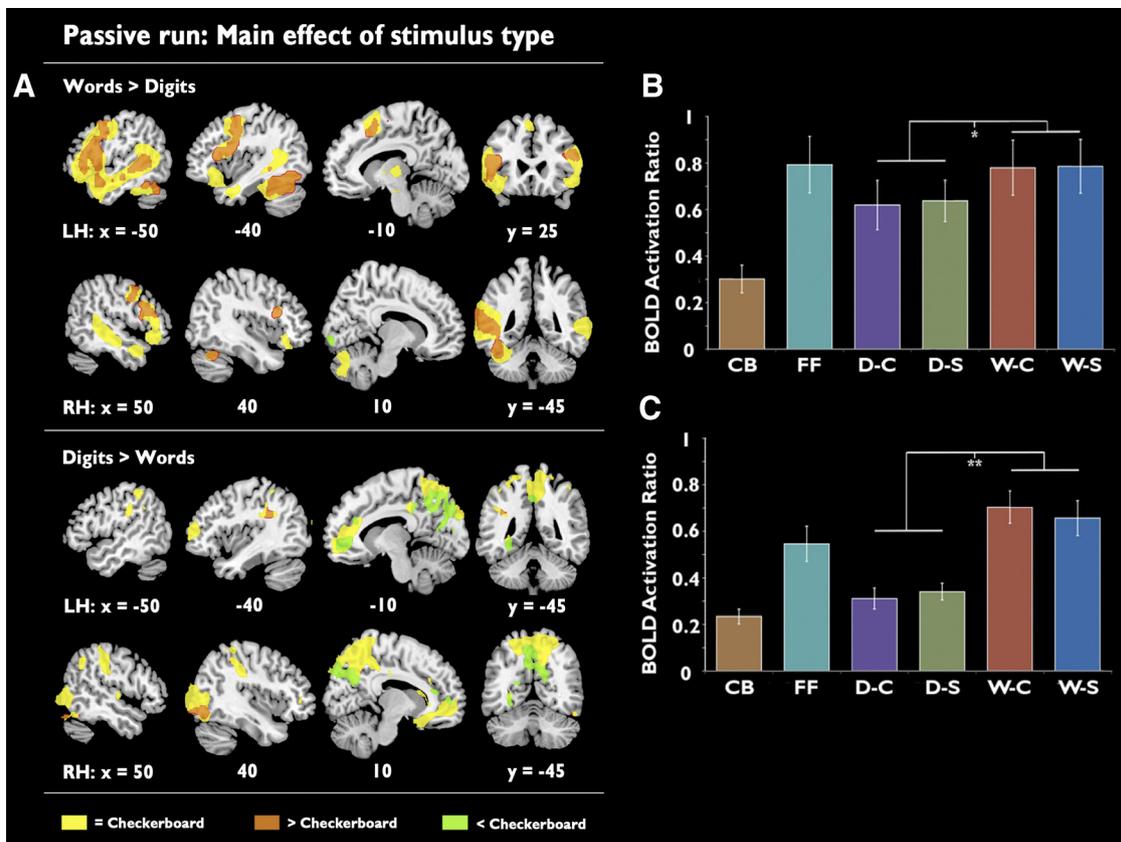


Figure 3. *A*, Activation maps showing the main effect of stimulus type for the passive run. Top, Areas where activation is preferential for words over numbers. Bottom, Areas where activation is preferential for numbers over words. Images are presented using a cluster-corrected thresholded of $z > 2.3, p < 0.01$. *B, C*, Region of interest analyses showing mean activation for each condition relative to the resting baseline in the anatomically localized VWFA (*B*) and the functionally localized VWFA (*C*). $**p = 0.001, *p < 0.05$. LH, Left hemisphere; RH, right hemisphere; CB, checkerboard; FF, false font; D-C, connected digits; D-S, single digits; W-C, connected words; W-S, single words. Error bars are ± 1 SEM. $n = 19$.

Table 1. Mean reaction times (and standard deviations) for the decision tasks in the active run in milliseconds

	PD	ND	All stimuli
Words	702 (94)	735 (71)	718 (77)
Digits	763 (114)	675 (86)	719 (90)
All tasks	732 (93)	705 (72)	719 (79)

$N = 17$.

Region-of-interest analysis

The ROI analysis, using a 7 mm radius sphere in the anatomically defined VWFA (Jobard et al., 2003), showed stronger activity for words than digits ($t_{(19)} = 2.48, p < 0.05$), but there was no significant difference between words and false font ($t_{(19)} = 0.11, p = 0.99$). A cluster of word-preferential activation just anterior to the known coordinates of the VWFA was distinct in the thresholded z -stat image ($z < 2.3, p < 0.01$) for the words versus false-font contrast. This cluster was used as an additional ROI for the subsequent analysis of the active task.

Active run

Behavioral results

Reaction times for the active tasks during the fMRI protocol were recorded for 17 of 19 participants; the data from the remaining two subjects were lost due to a technical error. For each participant, reaction time was calculated for trials where the recorded answer was accurate. Average reaction times for each category are reported in Table 1 and plotted in Figure 4. To reduce outliers (for example, from trials where participants made multiple re-

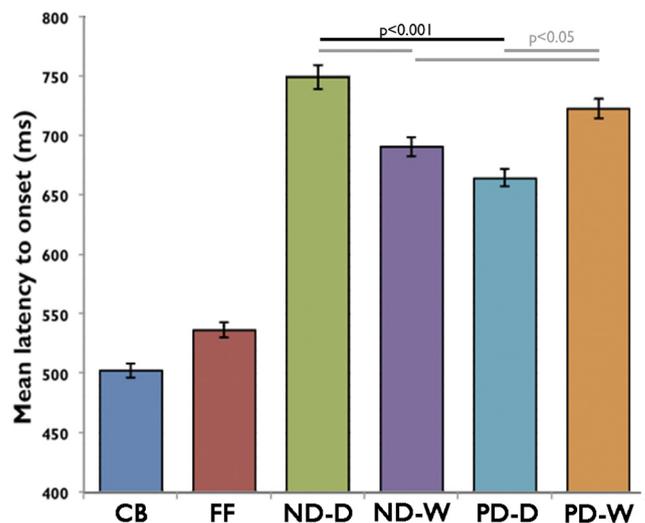


Figure 4. Mean latency to response for each task. Checkerboard (CB) and False-Font (FF) conditions have a significantly shorter latency to response. There is a significant difference between PD-Digits (PD-D) and PD-Words (PD-W) ($p < 0.05$), PD-Digits and ND-Digits (ND-D) ($p < 0.05$), and ND-Digits and ND-Words (ND-W) ($p < 0.01$). Error bars are ± 1 SEM. $n = 17$.

sponses) reaction times that lay > 2 SDs from the mean were excluded from this analysis.

Short reaction times for the baseline conditions were to be expected, as the oddball detection task could be performed using the overall visual characteristics of the stimulus. Reaction time

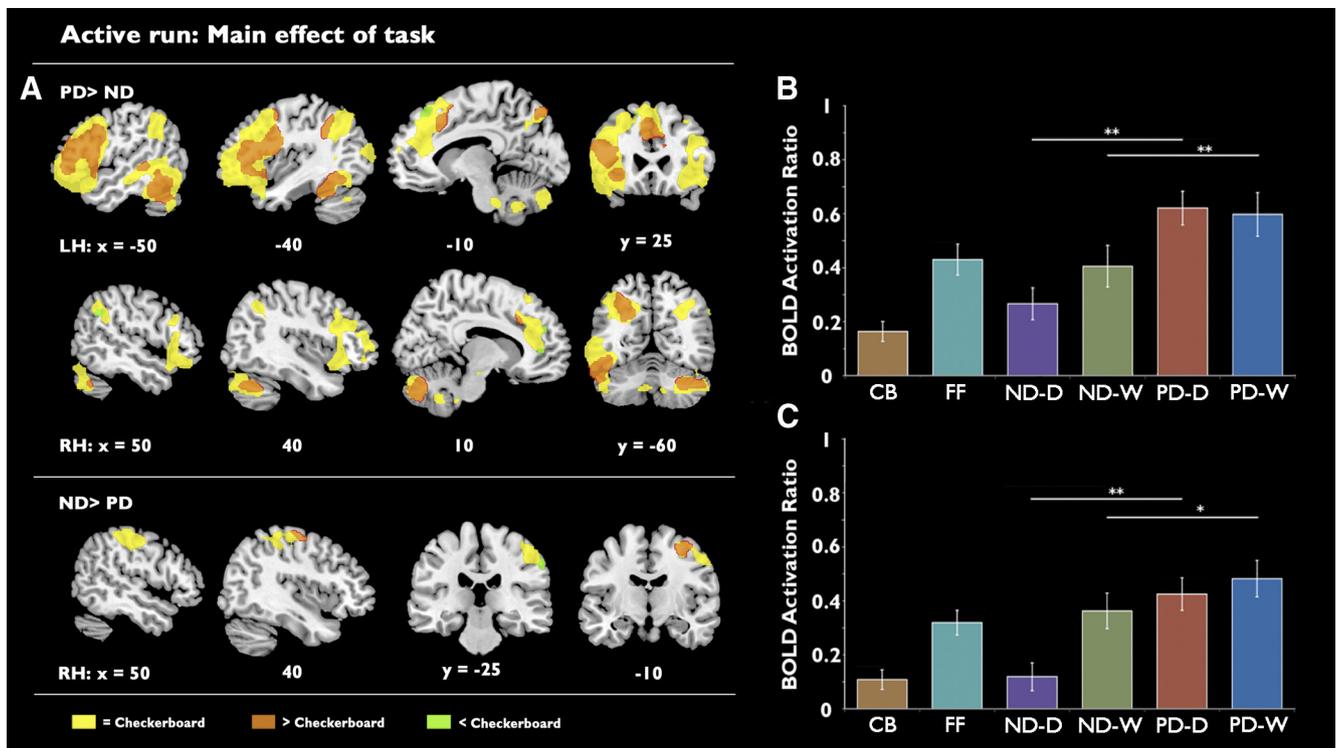


Figure 5. *A*, Maps of activation showing the main effect of task. Top, Preferential activation for the number decision task. Bottom, Preferential activation for the phoneme decision task. Images are presented with a cluster-corrected threshold of $z > 2.3$, $p < 0.01$. *B*, *C*, Region of interest analysis showing mean activation for each of the experimental conditions in the anatomically localized left VWFA (*B*) and the functionally localized VWFA (*C*). Significantly stronger activation was observed for the phoneme decision task in both regions. $**p = 0.001$, $*p < 0.05$. LH, Left hemisphere; RH, right hemisphere; CB, Checkerboard; FF, False Font; ND-D, ND-Digits; ND-W, ND-Words; PD-D, PD-Digits; PD-W, PD-Words. Error bars are ± 1 SEM. $n = 19$.

differences between the four experimental conditions of interest were assessed using a 2×2 repeated-measures ANOVA, with the factors stimulus type (words and digits) and task type (number and phoneme decision). This did not show a significant main effect of stimulus type ($F_{(1,16)} < 0.001$, $p = 0.98$), but a significant main effect of task type was observed ($F_{(1,16)} = 4.9$, $p < 0.05$), with faster reaction times for the numerical decision. A significant interaction between the type of stimulus and the task performed was also observed ($F_{(1,16)} = 16.6$, $p = 0.001$). *Post hoc* paired t tests indicated that the interaction was due to slower reaction times when the decision task was not congruent with the stimulus type. Thus, reaction times were slower for the phonological than the numerical decisions on the digit forms ($t_{(17)} = 3.9$, $p = 0.001$), and for the numerical than the phonological decisions on the word forms ($t_{(17)} = 3.7$, $p = 0.002$). These results indicate a preference, presumably related to prior experience, for making numerical decisions about digits and phonological decisions about words, even when the latter also represents numbers.

Functional MRI whole-brain analysis

To assess the areas of the brain modulated by task demands, a 2×2 repeated-measures ANOVA with the levels stimulus type (words or digits) by task type (phonological or numerical) was performed on the whole-brain fMRI data using FSL software. As with the passive run, the main effects were masked with a contrasts of all reading conditions versus checkerboards [(ND-Words + ND-Digits + PD-Words + PD-Digits) - checkerboards], and by the opposite contrast of checkerboards versus all reading conditions, to identify whether the results represented activations or deactivations relative to the checkerboard baseline.

As shown in Figure 5, the main effect of task demonstrated a predominantly left-lateralized network of activity that was preferential for the phonological decision task.

The results of this contrast (Fig. 5*A*, top) included activation within the ventral temporal lobe, first evident ~ 80 mm posterior to the anterior commissure. Additional task-related activity was observed in the left STS and the left inferior frontal gyrus (IFG), extending into premotor cortex and the anterior insula, with a smaller but similarly distributed cluster in the right frontal lobe. Finally, left lateralized occipitoparietal activation was observed from the occipital cortex toward the dorsal inferior parietal cortex (peak activation of this cluster was centered on stereotactic MNI coordinates $x = 46$, $y = -36$, $z = 54$). While the left hemisphere activations were mostly significantly activated relative to the checkerboard baseline (Fig. 5*A*, orange), the right hemisphere frontal and parietal lobe regions (Fig. 5*A*, yellow) were not preferentially activated or deactivated relative to checkerboards.

Preferential activity for the numerical decision task (Fig. 5*A*, bottom) was observed in right superior parietal cortex and adjacent postcentral gyrus, although the peak lay in the dorsal inferior parietal cortex (peak activation centered on stereotactic MNI coordinates $x = 48$, $y = -34$, $z = 50$). It is unlikely that this activation was primarily due to motor differences between the two conditions, as finger press responses, even though performed with the left hand, were balanced between the individual conditions. Most of the voxels in this cluster did not show a preferential response relative to the checkerboard baseline (Fig. 5*A*, yellow), but activity in a small dorsomedial portion of the cluster represented a significant activation over the baseline (Fig. 5*A*, orange), and a small ventrolateral portion in the vicinity of the supramarginal gyrus represented a significant deactivation relative to baseline (Fig. 5*A*, green).

The main effect of stimulus type demonstrated a left-lateralized cluster within the occipital cortex, driven by stronger activation

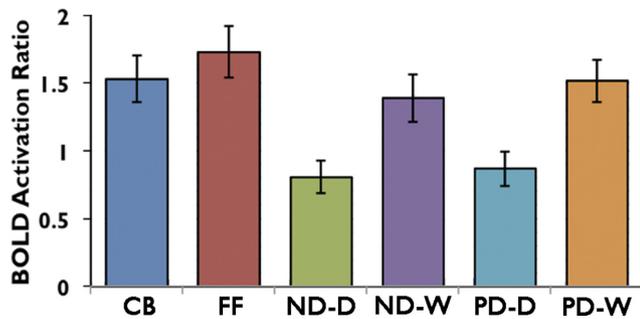


Figure 6. Region of interest analyses showing mean activation for each of the experimental conditions in the left primary visual cortex. CB, Checkerboard; FF, False Font, ND-D, ND-Digits; ND-W, ND-Words; PD-D, PD-Digits; PD-W, PD-Words. Error bars are ± 1 SEM. $n = 19$.

for orthographical than digital forms. This was most likely the consequence of differences in the amount of visual stimulation, as the number words were longer than the number digits (the average stimulus length for the words was 4.8 characters, and 1.6 characters for the digits).

As with the analysis of the passive data, the active data were reanalyzed using a higher z threshold of $z > 3.1$, $p < 0.05$ to ensure that no small but significant clusters had been omitted during the cluster-correction process. This analysis did not yield any new clusters.

Region-of-interest analysis

An ROI analysis was performed to investigate the effects of stimulus type and task demands on activity in the left vOT cortex. As shown in Figure 2, the analysis used an anatomically defined ROI for the VWFA, centered on coordinates cited in the meta-analysis by Jobard et al. (2003), and also a functionally defined ROI resulting from the contrast of words with false font during passive presentation of stimuli. A third ROI was located in the left primary visual cortex to test for early visual effects of stimulus length (number of characters). A 2×2 repeated-measures ANOVA with the levels stimulus type (orthographical or digital forms) and task performed (phonological or numerical) were conducted on the data from each ROI.

As shown in Figure 5B, analysis of the anatomically defined VWFA demonstrated a significant main effect of task ($F_{(1,18)} = 37.7$, $p < 0.001$) due to stronger activation for the phonological decision. There was no main effect of stimulus type. Additionally, a significant interaction between stimulus type and task was observed ($F_{(1,18)} = 15.9$, $p = 0.001$). *Post hoc* paired t tests showed that the interaction could be explained by a significant difference between activity during numerical decisions on orthographical and digital forms ($t_{(19)} = 3.0$, $p < 0.01$), but no difference between phonological decisions on orthographical and digital forms ($t_{(19)} = 0.4$, $p = 0.7$).

The ANOVA at the functionally defined ROI (Fig. 5C) showed a main effect of task ($F_{(1,18)} = 19.3$, $p < 0.001$) and a main effect of stimulus type ($F_{(1,18)} = 13.7$, $p < 0.002$). There was also a significant interaction ($F_{(1,18)} = 14.4$, $p < 0.001$). *Post hoc* paired t tests showed that the interaction could be explained by a significant difference between activity during numerical decisions on orthographical and digital forms ($t_{(19)} = 6.2$, $p < 0.01$).

In addition, paired t tests indicated that no significant difference was observed between activity for false-font scale detection tasks and numerical decisions on orthographical forms at either the functionally ($t_{(19)} = 0.9$, $p = 0.41$) or anatomically ($t_{(19)} = 0.5$, $p = 0.65$) defined VWFA.

A 7 mm spherical ROI was also placed in left primary visual cortex and mean activity was compared across all six conditions. The localization of this ROI is shown in Figure 2 and the results are plotted in Figure 6. A 2×2 repeated-measures ANOVA, with the

levels stimulus type (orthographical or digital forms) and task (phonological or numerical), illustrated that the difference in activation profile at V1 was, as predicted from the whole-brain analysis, driven by the main effect of stimulus type ($F_{(1,18)} = 52.51$, $p > 0.001$). This activity corresponded to stimulus complexity, and the influence of this was further confirmed by the strong activity in this region for false font and checkerboards, which was equivalent to or greater than activity in response to the orthographical stimuli. This profile of activity across conditions is compatible with a strong feedforward response of V1 to visual complexity. It was a profile of activity quite different from the two more anterior ROIs, where activity for checkerboards was lowest and where activity for digital forms (visually the least complex) depended strongly on task. Nevertheless, there was a weak but significant effect of task in V1 ($F_{(1,18)} = 6.47$, $p < 0.05$), indicating some feedback influence from later stages in the ventral visual processing stream on V1, an effect that has been observed by others (Szwed et al., 2011).

Discussion

This study demonstrates that the VWFA, whether defined by published stereotactic coordinates or by functional localization, responds to alphanumeric stimuli according to task and not to stimulus type. We have shown that a linguistic task on digital forms elicits a strong left-lateralized response early on in the left ventral visual stream. In contrast, a numerical decision on number words resulted in a significant reduction in activity compared with a linguistic decision on the same word. When the task performed on words was numerical, activity in both regions was no different from a size discrimination decision on false font. Residual greater activity for words than digits when the task was numerical could have been due either to the greater visual complexity of the orthographical forms or to parallel implicit linguistic processing of word forms even when the explicit task was numerical. Although it might be argued, based on comments by Dehaene et al. (2010) on the study of Yoncheva et al. (2010), that the phonological decision on a digit elicits a stage whereby the digit is transformed internally into a visual image of its word form in the left vOT, before accessing phonology, this interpretation cannot explain why a numerical decision based on perceived word forms results in activity no greater than size decision on false font. The profile of responses we observed across conditions argues strongly that the left vOT responds to different types of objects with high spatial frequency, and the level of activity depends on the task and, perhaps, on automatic access to the language system by words regardless of the type of task. These results, taken in conjunction with the observation that the left vOT shows equivalent or stronger levels of activation for object pictures than for words (Moore and Price, 1999; Price and Mechelli, 2005; Baker et al., 2007; Wright et al., 2008) make the term “visual word form area” misleading.

This conclusion is supported by the observations that phonological structure is strongly associated with the left inferior frontal gyrus (Jobard et al., 2003; Nixon et al., 2004; Price and Mechelli, 2005; Wheat et al., 2010), whereas higher-order number processing depends on parietal cortex (Dehaene et al., 1999; Levy et al., 1999; Mayer et al., 1999; Rickard et al., 2000; Naccache and Dehaene, 2001; Izard et al., 2008; Santens et al., 2010). We speculate that the strong lateralization of phonological processing (and speech production) to the left inferior frontal gyrus exerts a feedback influence that results in the lateralization of the response in vOT to the left when visual stimuli access the language system, either implicitly or explicitly. An alternative interpretation is that the main effect of task observed in the vOT relates to differences in task difficulty. The decision times

observed in the behavioral data indicate that the phonological task was more difficult to perform than the numerical task, particularly when the stimuli were digits. Hence, it is possible that the main effect of task observed in the vOT region of interest analysis was due to task-general top-down effects of attentional control, rather than task-specific top-down modulation from the phonological system. This could be tested by comparing vOT activity during an easy numerical task and a hard numerical task—if the vOT is modulated by attention, a task effect should still be observed. The data presented in the current study cannot discriminate between these two possibilities, but it is clear that, in either case, our results do not support the notion that cortex within the left vOT becomes specialized, mainly or exclusively, only for the perceptual processing of familiar strings of letters.

The importance of the left vOT for reading was originally demonstrated by clinical studies (Damasio and Damasio, 1983; Binder and Mohr, 1992; Cohen et al., 2003; Leff and Behrmann, 2008). A lesion to this region, or to its connections with both ipsilateral and contralateral primary visual cortex, results in a major impairment of reading, with relative preservation of individual letter recognition. This striking deficit in reading, which dominates the symptomatology of affected patients, is accompanied by a less apparent (and less recognized) deficit in recognition of line drawings (Behrmann et al., 1998), and possibly numbers (Starrfelt and Behrmann, 2011), which might be indicative of a more general perceptual impairment. The neuropsychological studies have now been accompanied by a parallel literature on functional neuroimaging studies of reading in normal subjects. From this literature, the most influential inference has been made that the left vOT becomes specialized to respond to word forms (both real words and orthographically regular nonwords) in preference to other two-dimensional objects of high contrast and spatial frequency (such as false font, orthographically irregular nonwords, and line drawings of objects) and that this effect is observed even when visual complexity of the stimuli is carefully controlled (Szwed et al., 2011).

This view has led to a hierarchical model of written word recognition, the LCD model (Dehaene et al., 2005), that describes the processing of word forms from V1 along the left ventral stream as far as the occipitotemporal sulcus adjacent to the posterior fusiform gyrus. This model assumes feedforward processing with lateral inhibition. Neurons are expected to pool information continually along the visual processing stream, progressively integrating simple visual forms such as contours and oriented bars into abstract letter forms. This model suggests that, at the level of what became termed the VWFA, neural representations of abstract words converge to represent orthographically regular combinations of letters. Thus, this model has implicated the VWFA as a localized center with a specialization, acquired through learning, for integrating letter forms into familiar combination units, invariant of retinal position, case, or font of the perceived word (Cohen et al., 2000; Cohen and Dehaene, 2004).

This view has, for some time, been regarded as controversial. Price and Devlin (2003, 2011), in particular, view the left vOT as a part of a more general object-recognition pathway, where activity is modulated by the specific task demand (Starrfelt and Gerlach, 2007; Twomey et al., 2011). Others have also published evidence that the response of the left vOT relates to prior activation of frontal regions that then exert a top-down influence (Bar et al., 2006; Kveraga et al., 2007; Gilaie-Dotan et al., 2009).

Most previous neuroimaging studies of the early reading pathway have used words, orthographically regular nonwords, or line drawings as activating stimuli. The present study took advan-

tage of the fact that numbers exist in two forms, as digits and words, but the associated phonology and meaning are identical. The LCD model, as specified, predicts that the VWFA will be active for words and not digits, regardless of task, as it is only involved in bottom-up orthographical processing (Baker et al., 2007; Reinke et al., 2008). This hypothesis is clearly at odds with the current data, which demonstrates significant local task-dependent modulation of activity for both number words and digits. A significant increase in activity in the left vOT was observed for both number words and digits when the subjects were required to perform a phonological rather than numerical task, which is incompatible with a predominantly feedforward mechanism of visual word recognition.

The phonological task linked numbers to speech production, regardless of the visual form in which the numbers were represented. The correlate of this translation of visual input to speech output was reflected in the strong activity in left posterior frontal cortex, including classic Broca's area. Additional frontal activity in the right IFG, anterior insula, and midline anterior cingulate cortex probably represents executive processing associated with decision making. Although both numerical and phonological decisions required responses, there was a main effect of decision type reflected in reaction times, indicating that, overall, phonological decisions were more difficult. The difference in the two types of decision, with greater time-on-task when the target was phonological structure, accounts for the midline and right frontal activity (Sarter et al., 2006).

The numerical decision task in the context of this study (odd or even) uses number semantics, which are represented in parietal cortex (Cappelletti et al., 2010). A previous study that investigated parietal cortex involvement in a range of different tasks that included converting orthography to phonology and a number task (although, in that instance, a simple subtraction) demonstrated parietal activity for both phonology and numbers (Simon et al., 2002). Although there were differences in tasks and contrasts between that study and this, it would seem reasonable to assume that the odd/even decision was processed in parietal cortex, which was only evident in a more anterior region on the right with masking of activity in more posterior right parietal cortex and left parietal cortex by relatively greater activity in the phoneme detection task. If there is component of the ventral visual stream that processes numbers and is functionally connected to parietal number processing cortex, it was not apparent from this study. However, we did not expect to find a visual number form area, as there is no numerical counterpart to pure alexia in the clinical neuropsychological literature on number recognition.

The results presented here are consistent with other models of reading that incorporate interactive or distributed processing between orthography, phonology, and semantics (Plaut et al., 1996; Patterson and Ralph, 1999; Devlin et al., 2006). These models propose that the balance of functional activity between these three domains leads to the emergence of a stable abstract form. The task-driven modulation of the VWFA demonstrated by the data presented here fits well with such interactive processing models, and argues against the emergence of encapsulated modularity, driven by experience of literacy, in the left vOT.

References

- Amunts K, Malikovic A, Mohlberg H, Schormann T, Zilles K (2000) Brodmann's areas 17 and 18 brought into stereotaxic space—where and how variable? *Neuroimage* 11:66–84.
- Andersson J, Jenkinson M, Smith S (2007a) Non-linear optimisation. FM-RIB technical report TR07JA1.

- Andersson J, Jenkinson M, Smith S (2007b) Non-linear registration, aka spatial normalisation. FMRIB technical report TR07JA2.
- Baker CI, Liu J, Wald LL, Kwong KK, Benner T, Kanwisher N (2007) Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proc Natl Acad Sci U S A* 104:9087–9092.
- Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Schmidt AM, Dale AM, Hämäläinen MS, Marinkovic K, Schacter DL, Rosen BR, Halgren E (2006) Top-down facilitation of visual recognition. *Proc Natl Acad Sci U S A* 103:449–454.
- Behrmann M, Nelson J, Sekuler EB (1998) Visual complexity in letter-by-letter reading: “pure” alexia is not pure. *Neuropsychologia* 36:1115–1132.
- Binder JR, Mohr JP (1992) The topography of callosal reading pathways: a case-control analysis. *Brain* 115:1807–1826.
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
- Cappelletti M, Lee HL, Freeman ED, Price CJ (2010) The role of right and left parietal lobes in the conceptual processing of numbers. *J Cogn Neurosci* 22:331–346.
- Cohen L, Dehaene S (2004) Specialization within the ventral stream: the case for the visual word form area. *Neuroimage* 22:466–476.
- Cohen L, Dehaene S, Naccache L, Lehéricy S, Dehaene-Lambertz G, Hénaff MA, Michel F (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, Martinaud O, Lemer C, Lehéricy S, Samson Y, Obadia M, Slachevsky A, Dehaene S (2003) Visual word recognition in the left and right hemispheres: anatomical and functional correlates of peripheral alexias. *Cereb Cortex* 13:1313–1333.
- Damasio AR, Damasio H (1983) The anatomic basis of pure alexia. *Neurology* 33:1573–1583.
- Dehaene S, Spelke E, Pinel P, Stanescu R, Tsivkin S (1999) Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science* 284:970–974.
- Dehaene S, Cohen L, Sigman M, Vinckier F (2005) The neural code for written words: a proposal. *Trends Cogn Sci* 9:335–341.
- Dehaene S, Pegado F, Braga LW, Ventura P, Nunes Filho G, Jobert A, Dehaene-Lambertz G, Kolinsky R, Morais J, Cohen L (2010) How learning to read changes the cortical networks for vision and language. *Science* 330:1359–1364.
- Devlin JT, Jamison HL, Gonnerman LM, Matthews PM (2006) The role of the posterior fusiform gyrus in reading. *J Cogn Neurosci* 18:911–922.
- Gilaie-Dotan S, Perry A, Bonneh Y, Malach R, Bentin S (2009) Seeing with profoundly deactivated mid-level visual areas: non-hierarchical functioning in the human visual cortex. *Cereb Cortex* 19:1687–1703.
- Glezer LS, Jiang X, Riesenhuber M (2009) Evidence for highly selective neuronal tuning to whole words in the “visual word form area.” *Neuron* 62:199–204.
- Izard V, Dehaene-Lambertz G, Dehaene S (2008) Distinct cerebral pathways for object identity and number in human infants. *PLoS Biol* 6:e11.
- Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration of brain images. *Med Image Anal* 5:143–156.
- Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17:825–841.
- Jobard G, Crivello F, Tzourio-Mazoyer N (2003) Evaluation of the dual route theory of reading: a meta-analysis of 35 neuroimaging studies. *Neuroimage* 20:693–712.
- Kleiner M, Brainard D, Pelli D (2007) What’s new in psychtoolbox-3? *Perception* 36:ECVP Abstract Supplement.
- Kveraga K, Boshyan J, Bar M (2007) Magnocellular projections as the trigger of top-down facilitation in recognition. *J Neurosci* 27:13232–13240.
- Leff AP, Behrmann M (2008) Treatment of reading impairment after stroke. *Curr Opin Neurol* 21:644–648.
- Leff AP, Scott SK, Crewes H, Hodgson TL, Cowey A, Howard D, Wise RJ (2000) Impaired reading in patients with right hemianopia. *Ann Neurol* 47:171–178.
- Levy LM, Reis IL, Grafman J (1999) Metabolic abnormalities detected by 1h-mrs in dyscalculia and dysgraphia. *Neurology* 53:639–641.
- Mayer E, Martory MD, Pegna AJ, Landis T, Delavelle J, Annoni JM (1999) A pure case of Gerstmann syndrome with a subangular lesion. *Brain* 122:1107–1120.
- Moore CJ, Price CJ (1999) Three distinct ventral occipitotemporal regions for reading and object naming. *Neuroimage* 10:181–192.
- Naccache L, Dehaene S (2001) The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cereb Cortex* 11:966–974.
- Nixon P, Lazarova J, Hodinott-Hill I, Gough P, Passingham R (2004) The inferior frontal gyrus and phonological processing: an investigation using rTMS. *J Cogn Neurosci* 16:289–300.
- Patterson K, Ralph MA (1999) Selective disorders of reading? *Curr Opin Neurobiol* 9:235–239.
- Plaut DC, McClelland JL, Seidenberg MS, Patterson K (1996) Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychol Rev* 103:56–115.
- Price CJ, Devlin JT (2003) The myth of the visual word form area. *Neuroimage* 19:473–481.
- Price CJ, Devlin JT (2011) The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn Sci* 15:246–253.
- Price CJ, Mechelli A (2005) Reading and reading disturbance. *Curr Opin Neurobiol* 15:231–238.
- Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. *Proc Natl Acad Sci U S A* 98:676–682.
- Reinke K, Fernandes M, Schwindt G, O’Craven K, Grady CL (2008) Functional specificity of the visual word form area: general activation for words and symbols but specific network activation for words. *Brain Lang* 104:180–189.
- Rickard TC, Romero SG, Basso G, Wharton C, Flitman S, Grafman J (2000) The calculating brain: an fMRI study. *Neuropsychologia* 38:325–335.
- Santens S, Roggeman C, Fias W, Verguts T (2010) Number processing pathways in human parietal cortex. *Cereb Cortex* 20:77–88.
- Sarter M, Gehring WJ, Kozak R (2006) More attention must be paid: the neurobiology of attentional effort. *Brain Res Rev* 51:145–160.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268:889–893.
- Simon O, Mangin JF, Cohen L, Le Bihan D, Dehaene S (2002) Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron* 33:475–487.
- Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17:143–155.
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23:S208–219.
- Starrfelt R, Behrmann M (2011) Number reading in pure alexia: a review. *Neuropsychologia* 49:2283–2298.
- Starrfelt R, Gerlach C (2007) The visual what for areas: words and pictures in the left fusiform gyrus. *Neuroimage* 35:334–342.
- Szwed M, Dehaene S, Kleinschmidt A, Eger E, Valabrègue R, Amadon A, Cohen L (2011) Specialization for written words over objects in the visual cortex. *Neuroimage* 56:330–344.
- Twomey T, Kawabata Duncan KJ, Price CJ, Devlin JT (2011) Top-down modulation of ventral occipito-temporal responses during visual word recognition. *Neuroimage* 55:1242–1251.
- Visser M, Embleton KV, Jefferies E, Parker GJ, Ralph MA (2010) The inferior, anterior temporal lobes and semantic memory clarified: Novel evidence from distortion-corrected fMRI. *Neuropsychologia* 48:1689–1696.
- Wheat KL, Cornelissen PL, Frost SJ, Hansen PC (2010) During visual word recognition, phonology is accessed within 100ms and may be mediated by a speech production code: evidence from magnetoencephalography. *J Neurosci* 30:5229–5233.
- Woodhead ZV, Brownsett SL, Dhanjal NS, Beckmann C, Wise RJ (2011) The visual word form system in context. *J Neurosci* 31:193–199.
- Woolrich MW, Ripley BD, Brady M, Smith SM (2001) Temporal autocorrelation in univariate linear modeling of fMRI data. *Neuroimage* 14:1370–1386.
- Wright ND, Mechelli A, Noppeney U, Veltman DJ, Rombouts SA, Glensman J, Haynes JD, Price CJ (2008) Selective activation around the left occipito-temporal sulcus for words relative to pictures: individual variability or false positives? *Hum Brain Mapp* 29:986–1000.
- Yoncheva YN, Zevin JD, Maurer U, McCandliss BD (2010) Auditory selective attention to speech modulates activity in the visual word form area. *Cereb Cortex* 20:622–632.