

Repetition Suppression in Occipitotemporal Cortex Despite Negligible Visual Similarity: Evidence for Postperceptual Processing?

Aidan J. Horner and Richard N. Henson*

MRC Cognition and Brain Sciences Unit, Cambridge, England



Abstract: The reduced neural response in certain brain regions when a task-relevant stimulus is repeated (“repetition suppression”, RS) is often attributed to facilitation of the cognitive processes performed in those regions. Repetition of visual objects is associated with RS in the ventral and lateral occipital/temporal regions, and is typically attributed to facilitation of visual processes, ranging from the extraction of shape to the perceptual identification of objects. In two fMRI experiments using a semantic classification task, we found RS in a left lateral occipital/inferior temporal region to a picture of an object when the name of that object had previously been presented in a separate session. In other words, we found RS despite negligible visual similarity between the initial and repeated occurrences of an object identity. There was no evidence that this RS was driven by the learning of task-specific responses to an object identity (“S-R learning”). We consider several explanations of this occipitotemporal RS, such as phonological retrieval, semantic retrieval, and visual imagery. Although no explanation is fully satisfactory, it is proposed that such effects most plausibly relate to the extraction of task-relevant information relating to object size, either through the extraction of sensory-specific semantic information or through visual imagery processes. Our findings serve to emphasize the potential complexity of processing within traditionally visual regions, at least as measured by fMRI. *Hum Brain Mapp* 32:1519–1534, 2011. © 2010 Wiley-Liss, Inc.

Key words: priming; object recognition; implicit memory; conceptual



INTRODUCTION

Repetition of a stimulus in a given task often results in a decrease in neural activity within certain cortical regions, a phenomenon known as repetition suppression (RS) [Grill-Spector et al., 2006]. When using fMRI while people cate-

gorise familiar visual objects, for example, RS is normally found in higher-order visual regions within the ventral processing stream [e.g., Koutstaal et al., 2001]. RS may represent a fundamental form of stimulus-specific neural plasticity, reflecting more efficient neural processing. This neural facilitation may also contribute to analogous behavioral phenomena, such as repetition priming [e.g., faster reaction times to make a categorization; see Henson, 2003; Schacter and Buckner, 1998].

Previous research demonstrating RS within ventral stream regions following repetition of familiar visual objects has normally attributed such effects to facilitation of perceptual processes. These perceptual processes are seen as distinct from the RS often seen in more anterior regions, such as inferior prefrontal regions, which is normally attributed to facilitation of phonological, lexical, and/or semantic processing [Poldrack et al., 1999; Wagner et al., 1997; Wagner et al., 2000]. More specifically, lateral

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: UK Medical Research Council; Contract grant number: WBSE U.1055.05.012.00001.01.

*Correspondence to: Dr. Richard N. Henson, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge, CB2 7EF, UK. E-mail: rik.henson@mrc-cbu.cam.ac.uk

Received for publication 14 April 2010; Accepted 15 June 2010

DOI: 10.1002/hbm.21124

Published online 2 September 2010 in Wiley Online Library (wileyonlinelibrary.com).

regions of the occipital cortex including the occipitotemporal sulcus, as well as posterior regions of the fusiform gyrus—corresponding to the posterior and anterior portions respectively of what has been called the “Lateral Occipital Complex” (LOC) [Malach et al., 1995]—have been associated with relatively low-level perceptual processes involved in the extraction of object shape. This is based on findings that RS (or “adaptation” to multiple stimulus repetitions) generalizes across manipulations that maintain object shape, such as changes in retinotopic location and stimulus size [Grill-Spector et al., 1999; Grill-Spector and Malach, 2001] and stimulus format [e.g., from line-drawings to grayscale photographs; Kourtzi and Kanwisher, 2000, 2001] as well as across mirror reflections [Eger et al., 2004]. Conversely, manipulations that disrupt object shape, such as changes in object viewpoint [Andresen et al., 2009; Ewbank et al., 2005; Grill-Spector et al., 1999; Grill-Spector and Malach, 2001; though see James et al., 2002] and occlusion of object parts [Hayworth and Biederman, 2006], have been shown to disrupt RS in these regions.

There has also been a suggestion that RS in posterior fusiform cortex—corresponding to the anterior portion of the LOC—demonstrates greater resilience to such changes, consistent with a posterior-anterior gradient with respect to representational abstraction [Grill-Spector et al., 1999; see Grill-Spector and Malach, 2004 for a review of RS in the functionally defined LOC].¹ Extending such findings, Koutstaal et al. [2001] showed that RS in a left mid-fusiform region generalised over different exemplars of an object with the same name (e.g., pictures of different umbrellas). However, while Vuilleumier et al. [2002] found that RS in this region generalized over different viewpoints of an object, they found no evidence that it generalised over different exemplars, unlike Koutstaal et al. [2001]. This lack of generalization of fusiform RS was also unlike the RS that Vuilleumier et al. [2002] found in left inferior prefrontal cortex, which did generalize across different exemplars with the same name. On the basis of these findings, Vuilleumier et al. [2002] argued that the generalisation of RS found by Koutstaal et al. [2001] reflected visual similarity between the exemplars, rather than an abstract representation of an object identity. This claim was later supported by Chouinard et al. [2008], who failed to find occipitotemporal RS when controlling for visual similarity between exemplars.

Other studies however have continued to implicate mid-fusiform regions—particularly in the left hemisphere—with more abstract processing. For example, Simons et al. [2003] reported a left fusiform region that showed RS to

object pictures that were immediately preceded, and accompanied, by auditory presentation of the name of that object. This implicates this region in lexical or semantic processing. Furthermore, Wheatley et al. [2005] and Gold et al. [2006] even found RS in left fusiform cortex for semantically related versus unrelated word pairs, implicating this region in semantic processing.

Interpretation of these fMRI RS effects in the ventral visual stream is further complicated by recent evidence that RS may also reflect the “by-passing” of processing in such regions, owing to the direct retrieval of task-relevant responses previously associated with a stimulus [Dobbins et al., 2004; Horner and Henson, 2008; Race, Shanker and Wagner, 2009]. The idea behind such stimulus-response (S-R) learning is that the response made on the initial presentation of a stimulus becomes bound to that stimulus, such that when the stimulus is repeated, the response can be retrieved quickly, without needing to repeat any detailed perceptual or semantic analysis of the stimulus. Thus the RS observed in the object categorization tasks used by many of the above fMRI studies may not reflect more efficient perceptual processing per se, but rather substantial attenuation (or even abolishment) of such processing. S-R learning has long been known to exert strong effects on the behavioral priming that is found in speeded categorization tasks [Horner and Henson, 2009; Logan, 1990; Schnyer et al., 2006]. More recently, behavioral effects of S-R learning have been shown to generalize across different object exemplars [Denkinger and Koutstaal, 2009; though see Schnyer et al., 2007]. Thus it is possible that at least some of the RS effects reviewed above that were used to argue for different levels of object representation in the ventral visual stream actually reflect relatively abstract S-R learning, rather than facilitation of processes normally involved in visual object recognition.

Dobbins et al. [2004] suggested one way of testing for such S-R learning: by reversing the categorizations between initial and repeated stimulus presentations. On initial presentation, participants were asked “is the object bigger than a shoebox?”. When the same decision was used for repeated objects, RS was found in regions including the fusiform cortex. When the decision was reversed however (i.e., “is the object smaller than a shoebox?”), RS was no longer significant in fusiform cortex. Although other studies have since found RS in fusiform regions despite such task reversal [Horner and Henson, 2008; Race et al., 2009], the comparison between “Same” and “Reverse” task conditions still offers a way to test for S-R learning, given the large effect it produces on behavioral priming (see Denkinger and Koutstaal, 2009; and Horner and Henson, 2009 for further discussion). We therefore used this Same/Reverse task manipulation in the present experiment.

Given the controversy regarding the degree of abstraction of object processing in ventral visual stream regions, based on evidence from RS paradigms, we asked the following question: can we see RS to visual objects within occipital/temporal regions when repeating an object

¹Similar questions have been asked regarding the role of occipitotemporal cortex in processing of faces [Andrews and Ewbank, 2004; Eger et al., 2005; Eger et al., 2004; Ewbank and Andrews, 2008; Rotstein et al., 2005] and of written words [Cohen and Dehaene, 2004; Cohen et al., 2000; Vinckier et al., 2007]. Given the debate about whether faces and words are special classes of visual stimuli; however, we do not review these findings here.

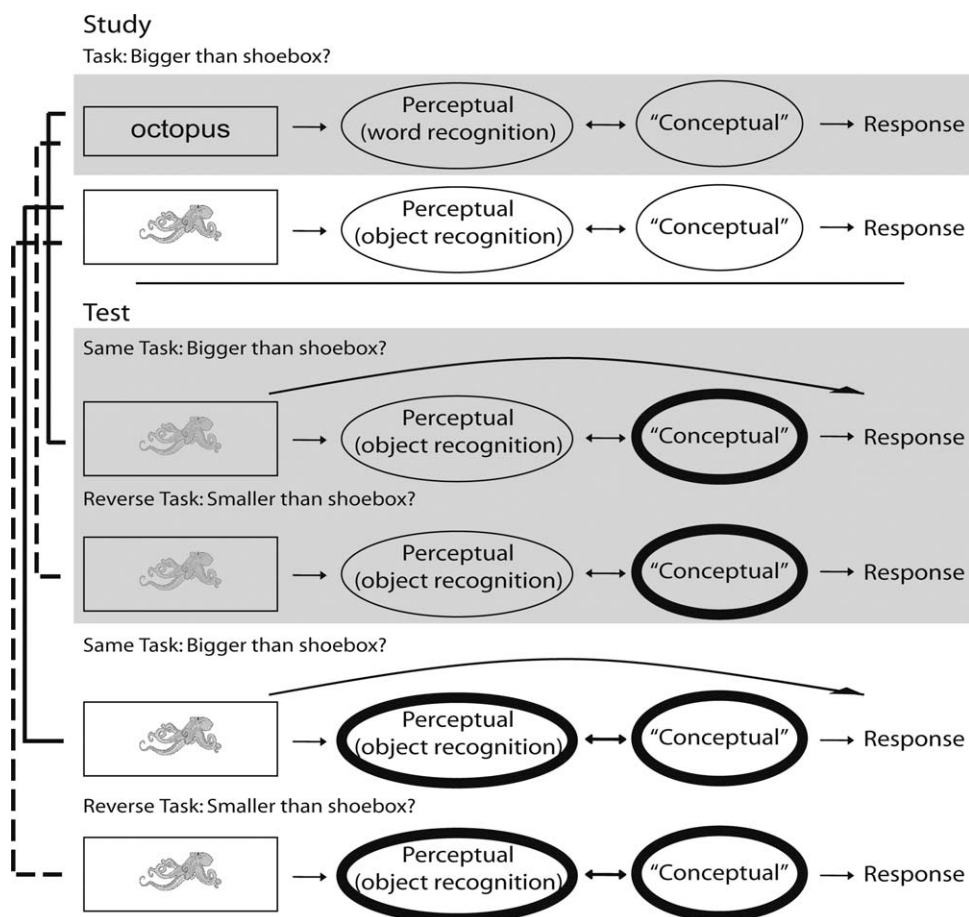


Figure 1.

Basic design of Experiment 1 (shaded rows) and Experiment 2 (all rows). Square boxes indicate stimuli; ovals represent hypothetical cognitive processes with thicker lines indicating facilitated processing from the Study phase. Arrows between ovals represent bottom-up and top-down communication between processes whereas curved arrows between stimuli (squares) and responses represent possible by-passing of cognitive processes through (S-R) retrieval. Solid lines to the left connecting particu-

lar rows highlight the Same Task condition; dotted lines highlight the Reverse Task condition. Note that, though we refer to any RS that is seen in our Word-Picture condition as conceptual, as distinct from perceptual, it could in fact reflect multiple processes operating post object recognition, such as phonological and/or lexical retrieval of object names, as we consider further in the General Discussion.

identity despite negligible visual similarity between its initial and repeated presentation? We tested this by examining the generalization of RS from words (object names) to pictures (of the same objects). The basic study-test design is shown in Figure 1. Color pictures of everyday nameable objects were always used at Test, one half of which (Repeated condition) depicted objects that were previously encountered at Study, and the other half of which depicted objects not encountered at study (Novel condition). RS was defined as the reduction in mean event-related fMRI response to pictures in the Repeated relative to Novel condition in the Test phase. In Experiment 1 (shaded rows of Fig. 1), we only used a word-picture condition; in Experiment 2 we added a picture-picture condition as well, in

order to compare RS from words to that from pictures (i.e., negligible versus full visual similarity).

To test whether RS was affected by S-R learning, we added an orthogonal manipulation of using either the same task at Study and Test, or reversing the task between Study and Test. We chose to use the "bigger-than-shoebox" size-judgment task because it has been used in numerous behavioral [Horner and Henson, 2009; Schnyer et al., 1996, 2007] and fMRI [Dobbins et al., 2004; Horner and Henson, 2008; Koutstaal et al., 2001; Race et al., 2009; Simons et al., 2003] studies. Although S-R learning might not be expected when the visual form of the stimulus (S) changes so dramatically (i.e., from a word to a picture), it is possible that a response (R) can become bound to a

relatively abstract (amodal) representation of an object identity. This would at least mirror our prior behavioral evidence that the response representations in S-R learning can be quite abstract (to the level of a particular semantic label, e.g., “bigger” or “smaller”, independently of the yes/no decision or specific motor action, Horner and Henson, 2009). In any case, if the amount of RS in occipitotemporal regions in the word-picture condition was unaffected by a task reversal, then it is unlikely to reflect S-R learning, and more likely to reflect facilitation of some post-perceptual processing of visual objects.

Experiment 1

Experiment 1 was designed to assess whether significant RS could be seen within occipital/temporal regions once we controlled for visual similarity between Study and Test. At Study, word stimuli were presented (e.g., the word “lion” was presented) with participants performing the “bigger-than-shoebox” task. At Test, the same identities (object referents) seen at Study (along with novel items) were presented as pictures rather than as words (e.g., a picture of a “lion”), with participants either performing the “bigger-than-shoebox” or “smaller-than-shoebox” task (see Fig. 1). Thus, item identity was repeated between Study and Test; however, there was no visual similarity between repetitions. These manipulations resulted in a 2×2 factorial design crossing the factors Repetition (Novel, Repeated) and Task (Same, Reverse).

MATERIALS AND METHODS

Participants

Participants in both experiments were recruited from the MRC-CBU subject panel, or from the student population of Cambridge University. All participants had normal or corrected to normal vision and were right-handed by self-report. Both experiments were of the type approved by a local research ethics committee (LREC reference 05/Q0108/401).

Eighteen participants (eight male) gave informed consent to participate in Experiment 1. The mean age across participants was 23.1 years ($SD = 2.1$). Participants were the same as those reported in Experiment 2 of Horner and Henson [2008].

Materials

Stimuli were 160 colored images of everyday objects (and their name equivalents), taken from a set used by Dobbins et al. [2004]. They were selected so that 50% were bigger than a shoebox and 50% were smaller than a shoebox, according to norms from independent raters [Horner and Henson, 2009]. Each stimulus was randomly assigned to one of 4 groups relating to the four experimental conditions, with each group containing equal numbers of each

stimulus classification, resulting in 40 stimuli per group. The assignment of groups to experimental condition was rotated across participants. The scrambled stimuli used during Study blocks (see Procedure) were created from the same set of objects by randomly redistributing the pixels so that a coherent object was no longer visible. None of the stimuli used in the present experiment were presented in the remainder of the scanning visit (i.e., they were not seen in Horner and Henson, 2008—Experiment 2).

Procedure

Experiment 1 was conducted at the end of the same scanning visit as Experiment 2 of Horner and Henson [2008]. The experiment consisted of two study-test cycles, with each cycle lasting ~ 10 min. During each Study phase 80 stimuli were shown. Forty stimulus identities were presented once as visual words (e.g., the word “lion” was presented rather than a picture of a lion). A further 40 scrambled images (see Materials) were presented once. Words and scrambled images were grouped into mini-blocks of five stimuli, with each mini-block lasting 15 s. During word presentation mini-blocks, participants were required to respond to whether the stimulus was “bigger than a shoebox,” where the comparison referred to the object’s size in real life. During scrambled image mini-blocks, participants were instructed to alternate between right and left key-presses at stimulus onset. During each Test phase, the 40 stimuli from the Study phase were randomly intermixed with 40 novel stimuli. Crucially, object pictures were presented at Test (e.g., a picture of a lion rather than the word “lion”) such that there was no visual overlap between stimuli seen at Study and Test. Participants were either asked the same question to that at Study (e.g., “is the object bigger than a shoebox?” - the Same condition) or the opposite question (e.g., “is the object smaller than a shoebox?”—the Reverse condition). The order of the two test conditions (tasks) was counterbalanced across participants.

Each trial sequence began with a centrally placed fixation cross presented for 500 ms, followed by a stimulus for 2,000 ms, in turn followed by a blank screen for 500 ms. Images subtended $\sim 6^\circ$ of visual angle. Words were presented in black on a white background with the same pixel dimensions as the object picture stimuli. Participants were able to respond at any point up to the start of a new trial (i.e., the presentation of another fixation cross). Participants responded using a “yes” or “no” key with their right or left index finger, respectively. Prior to entering the scanner, participants were asked to perform a practice session using the “bigger-than-shoebox” task.

Behavioral Analyses

Trials in which RTs were less than 400 ms, or two or more standard deviations above or below a participant’s

mean for a given block (i.e., a separate Study or Test phase), were excluded. Subsequent to this exclusion, accuracy was based on prior norms [Horner and Henson, 2009]. For the RT analyses, trials at Test were further excluded if objects were given an incorrect response at Study. Repetition priming was then calculated as the difference in mean RTs between Novel and Repeated stimuli. All statistical tests had alpha set at 0.05, and a Greenhouse-Geisser correction was applied to all *F*-values with more than one degree of freedom in the numerator. *T*-tests were two-tailed, except where stated otherwise.

fMRI Acquisition

Thirty-two T2*-weighted transverse slices ($64 \times 64 \times 3$ mm $\times$ 3 mm pixels, TE = 30 ms, flip-angle = 78°) per volume were taken using Echo-Planar Imaging (EPI) on a 3T TIM Trio system (Siemens, Erlangen, Germany). Slices were 3-mm thick with a 0.75-mm-gap, tilted $\sim 30^\circ$ upward at the front to minimize eye-ghosting, and acquired in descending order. Four sessions of 130 volumes were acquired, with a repetition time (TR) of 2,000 ms. The first five volumes of each session were discarded to allow for equilibrium effects. A T1-weighted structural volume was also acquired for each participant with 1mmx1mmx1 mm voxels using MPRAGE and GRAPPA parallel imaging (flip-angle = 9° ; TE = 2.00 s; acceleration factor = 2).

fMRI Analysis

Data were analyzed using Statistical Parametric Mapping (SPM5, <http://www.fil.ion.ucl.ac.uk/spm5.html>). Preprocessing of image volumes included spatial realignment to correct for movement, followed by spatial normalization to Talairach space, using the linear and nonlinear normalization parameters estimated from warping each participant's structural image to a T1-weighted average template image from the Montreal Neurological Institute (MNI). These re-sampled images (voxel size $3 \times 3 \times 3$ mm³) were smoothed spatially by an 8 mm FWHM Gaussian kernel (final smoothness $\sim 11 \times 11 \times 11$ mm³).

Statistical analysis was performed in a two-stage approximation to a Mixed Effects model. In the first stage, neural activity was modeled by a delta function at stimulus onset. The BOLD response was modeled by a convolution of these delta functions by a canonical Haemodynamic Response Function (HRF). The resulting time-courses were down-sampled at the midpoint of each scan to form regressors in a General Linear Model.

For each Test session (Task), five separate regressors were modeled—the two experimental conditions (Novel, Repeated) were split according to the particular key-press given (left/right), plus an additional regressor for discarded trials (using the behavioral exclusion criteria outlined earlier). To account for (linear) residual artifacts after realignment, the model also included six further regressors representing the movement parameters estimated during

realignment. Voxel-wise parameter estimates for these regressors were obtained by Restricted Maximum-Likelihood (ReML) estimation, using a temporal high-pass filter (cut-off 128 s) to remove low-frequency drifts, and modeling temporal autocorrelation across scans with an AR(1) process.

Images of contrasts of the resulting parameter estimates (collapsed across left/right key-press) comprised the data for a second-stage model, which treated participants as a random effect. In addition to the 18 subject effects, this model had four condition effects, corresponding to a 2×2 (Task $\times$ Repetition) repeated-measures ANOVA. Within this model, Statistical Parametric Maps (SPMs) were created of the *T* or *F*-statistic for the various ANOVA effects of interest, using a single pooled error estimate for all contrasts, whose nonsphericity was estimated using ReML as described in Friston et al. [2002]. Unless otherwise stated, all SPMs were height-thresholded at the voxel-level at $P < 0.05$, corrected for multiple comparisons using Random Field Theory, either across the whole-brain or within regions of interest (ROIs) defined by contrasts from independent data. Stereotactic coordinates of the maxima within the thresholded SPMs correspond to the MNI template.

RESULTS

Behavioral Results

After excluding 5.6% of trials with outlying RTs, the percentages of errors are shown in Table I. A 2×2 (Task $\times$ Repetition) repeated-measures ANOVA on errors revealed no significant main effects or interactions (*F*'s < 1.9 , *P*s > 0.19). A further 3.8% of Repeated trials were excluded from RT analysis due to incorrect responses given at Study (see Methods). Table I displays mean RTs, while Figure 2A shows priming (Novel-Repeated) of RTs across all conditions. Priming was reliable in the Same condition, $t(17) = 1.86$, $P < 0.05$, but did not reach significance in the Reverse condition, $t(17) = 0.48$, $P = 0.32$ (one-tailed). A 2×2 (Task $\times$ Repetition) ANOVA however showed no evidence of an interaction between the Same/Reverse conditions and priming ($F(1, 17) = 0.41$, $P = 0.53$; cf. Experiment 2).

Finally, to check whether priming differed as a function of Test block or Task order, we conducted a 2×2 (Block $\times$ Order) mixed ANOVA, where the within-subject factor Block refers to the Test block 1 or 2 (regardless of Task) and the between-subject factor Order refers to the task order (i.e., Same-Reverse or Reverse-Same). This 2×2 ANOVA failed to reveal any main effects of Block or Order, *F*s < 0.39 , *P*s > 0.54 , suggesting priming did not vary as a function of block or task order.

fMRI Results

We first sought evidence for significant RS (i.e., a Novel—Repeated one-tailed *T*-contrast). Our initial whole-brain analysis revealed no significant effects. Given we

TABLE I. Mean percentage errors and RTs (plus standard deviations) across Task and Repetition for Experiment 1 and Stimulus-modality, Task, and Repetition for Experiment 2

Stimulus-type		Picture-Picture		Word-Picture	
Task		Same	Reverse	Same	Reverse
Errors					
Experiment 1	Novel	Picture-Picture condition not included in Experiment 1		13.5 (5.4)	16.7 (5.4)
	Repeated			13.6 (6.4)	14.4 (4.7)
Experiment 2	Novel	12.1 (5.6)	15.3 (5.7)	11.8 (6.1)	15.7 (6.1)
	Repeated	12.2 (4.7)	12.8 (5.4)	12.8 (5.8)	13.1 (5.4)
RTs					
Experiment 1	Novel	Picture-Picture condition not included in Experiment 1		859 (210)	791 (130)
	Repeated			827 (200)	780 (134)
Experiment 2	Novel	894 (127)	1005 (146)	905 (147)	990 (129)
	Repeated	772 (103)	921 (133)	847 (132)	976 (144)

Results from Experiment 2 are collapsed across Prime-level for clarity. Note that this factor did not interact significantly with Stimulus-modality or Task so is of little theoretical interest.

were interested in regions previously shown to demonstrate significant RS in visual object repetition paradigms, when the same object is shown at Study and Test, we next constrained our search using the “main effect” of RS (i.e., the corrected-thresholded map for the RS T-contrast) from Experiment 2 of Horner and Henson [2008], which contained 1,980 voxels. This T-contrast was derived from independent data taken from the same participants in the same scanning visit as the present experiment; our voxel selection for this small-volume correction (SVC) is therefore not biased in favor of finding a significant RS effect. This T-contrast map covered bilateral occipital/temporal cortex, including lateral occipital and fusiform cortex, as well as distinct clusters in the left inferior prefrontal gyrus. Two regions survived SVC: (1) a region in the left inferior frontal gyrus—*pars opercularis*—henceforth referred

to as the posterior prefrontal cortex (pPFC) (−48, +3, +24) and (2) a region in the left hemisphere on the lateral surface starting in the middle occipital gyrus and descending into the posterior inferior temporal gyrus—henceforth referred to as lateral occipital/inferior temporal cortex (LO-IT) (−51, −66, 0) (see Fig. 3A). RS in the LO-IT region was numerically greater in the Same than Reverse condition (Fig. 3B; also see Table II for mean percentage signal change across all conditions), reflected by a trend for an interaction $F(1, 17) = 4.04$, $P = 0.06$. Nonetheless, residual RS still appeared reliable in the Reverse condition, suggesting that this RS was not dependent on S-R retrieval (error bars in Fig. 3B reflect 95% confidence intervals, though note that these simple effects of repetition are biased by the prior selection of the region to show a main effect of repetition). Assessing Test block and Task order

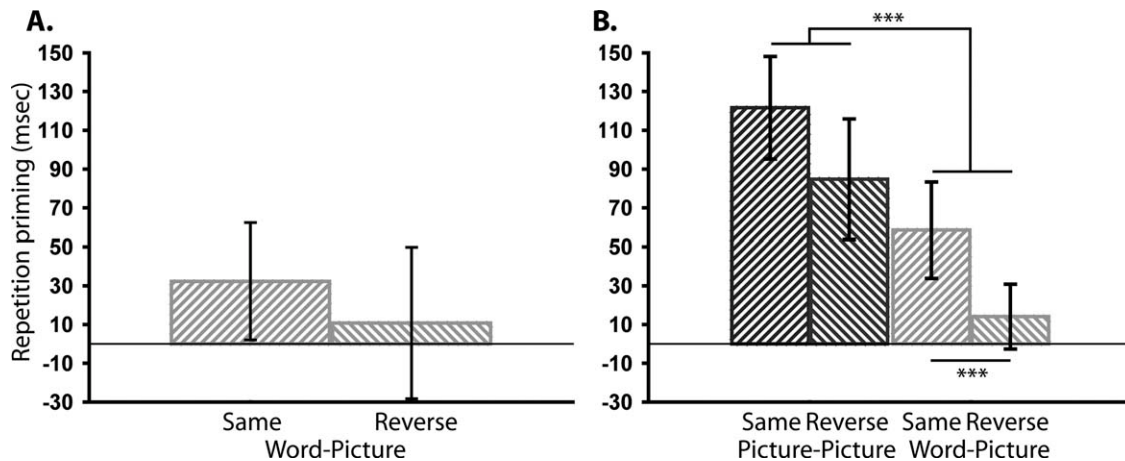


Figure 2.

Behavioral priming across Task (Same vs. Reverse) in Experiment 1 (A) and across Task (Same vs. Reverse) and Stimulus-type (Picture-Picture vs. Word-Picture) in Experiment 2 (B). Error bars represent one-tailed 95% confidence intervals. *** $P < 0.001$.

TABLE II. Mean percentage signal change (and standard deviations) within left LO-IT (-51, -66, 0) across task and repetition for experiment 1 and stimulus-type, task and repetition for experiment 2

Stimulus-type		Picture-Picture		Word-Picture	
Task		Same	Reverse	Same	Reverse
Experiment 1	Novel	Picture-Picture condition not included in Experiment 1			-0.18 (0.66)
	Repeated				-0.09 (0.34)
Experiment 2	Novel	0.13 (0.43)	0.11 (0.45)	0.10 (-0.06)	0.13 (0.39)
	Repeated	-0.05 (0.42)	-0.06 (0.42)	0.05 (0.39)	0.08 (0.46)

Results from Experiment 2 are collapsed across Prime-level for clarity. Percent signal change refers to the peak of the fitted BOLD impulse response, and is relative to the grand mean over all voxels and scans. Note that the baseline level of 0 was not estimated reliably in this design, so only relative patterns across conditions are meaningful.

(see behavioral analysis), a 2×2 (Block $\times$ Order) mixed ANOVA on RS showed no main effects of Block or Order, $F_s < 3.0$, $P_s > 0.10$, suggesting that RS, like behavioral priming, was unaffected by test block or task order.

A whole-brain search for the interaction between Repetition and Task, to further investigate possible effects of S-R learning, did not reveal any regions that survived either whole-brain correction, or small-volume correction for the

main effect of RS. Lastly, searching for regions showing significantly greater activation for repeated than novel items (i.e., repetition enhancement) revealed several clusters that survived whole-brain correction (see Supporting Information Table I). Given our present focus on RS within posterior occipitotemporal regions however we do not discuss these results further (see Horner and Henson, 2008 for a discussion of this issue).

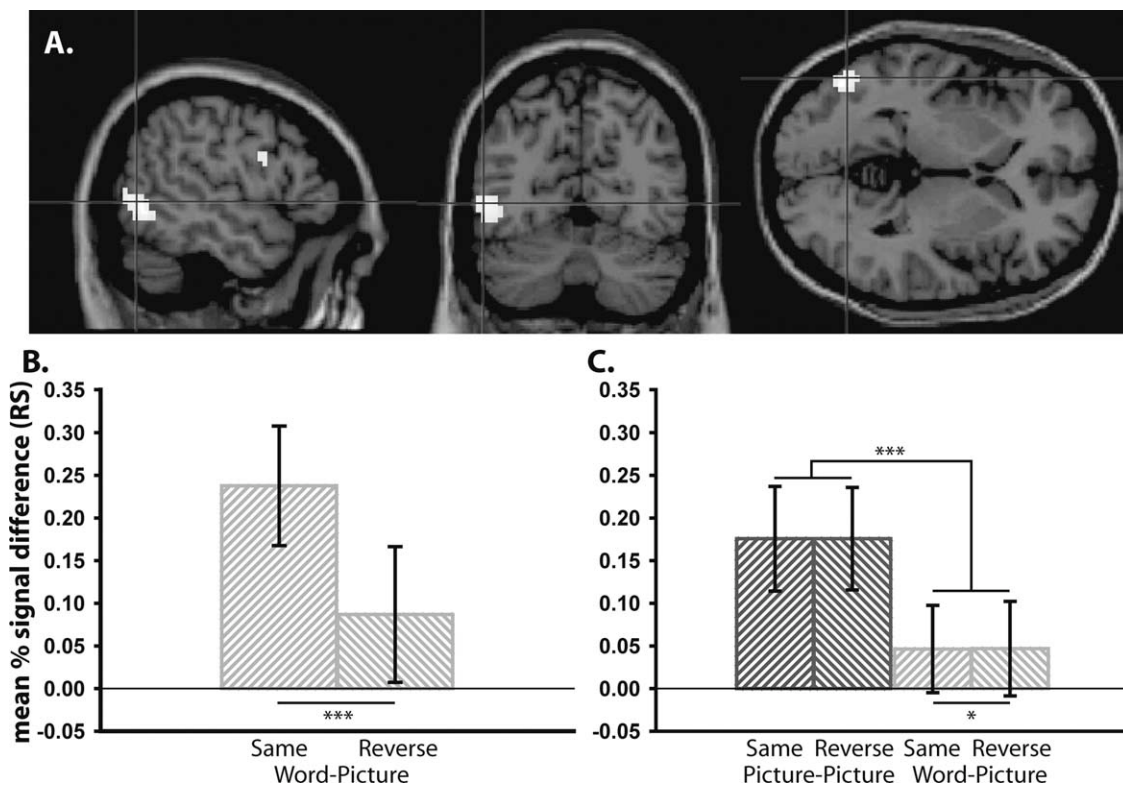


Figure 3.

A: Voxels demonstrating significant repetition suppression (RS) in the Word-Picture condition in Experiment 1 across sagittal, coronal and axial slices; $P < 0.05$ small-volume corrected. RS in left LO-IT (-51, -66, 0) across Task (Same vs. Reverse) in Experiment 1 (**B**) and across Task (Same vs. Reverse) and Stimulus-type (Picture-Picture vs. Word-Picture) in Experiment 2 (**C**). Error bars represent one-tailed 95% confidence intervals. *** $P < 0.001$, * $P < 0.05$.

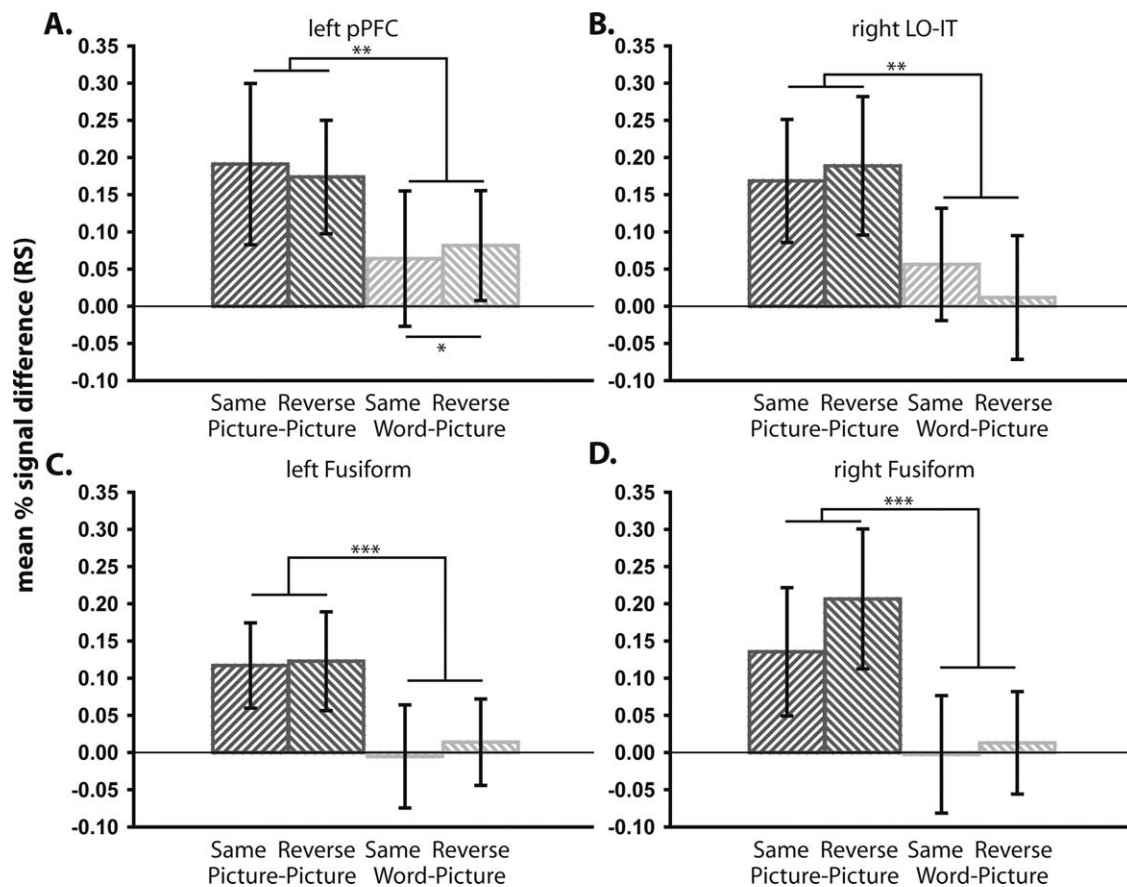


Figure 4.

Repetition suppression (RS) across Task (Same vs. Reverse) and Stimulus-type (Picture-Picture vs. Word-Picture) in Experiment 2 in left posterior prefrontal (pPFC) cortex (**A**), right lateral occipital/inferior temporal (LO-IT) cortex (**B**) and left and right fusiform cortex (**C** and **D**, respectively). Error bars represent one-tailed 95% confidence intervals. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

DISCUSSION

Experiment 1 demonstrated that significant RS can be seen in prefrontal and occipital/temporal regions despite negligible visual similarity between Study and Test stimuli, and without apparent contributions from S-R learning. One possibility is that RS within these regions can reflect facilitation of “higher-level” processes, such as phonological or lexical processes (associated with naming the objects) or possibly semantic processing associated with the conceptual task (see General Discussion).

Our pPFC results support previous research suggesting that RS within left inferior PFC regions reflects improved phonological and/or semantic processing [Poldrack et al., 1999; Wagner et al., 1997]. Indeed, given RS in the present experiment was confined to more dorsal regions of the inferior frontal gyrus—*pars opercularis*—it is likely our results reflect repetition of phonological processes [Poldrack et al., 1999]. Furthermore, the lack of reliable difference between

our Same and Reverse tasks (see also Experiment 2 and Fig. 4A) suggests that PFC RS is not necessarily always related to S-R learning [Dobbins et al., 2004]. One possibility is that, while S-R learning effects might generalize across visually-similar pictures of different exemplars of an object [Denkinger and Koutstaal, 2009], they do not generalize across visually dissimilar stimuli [Schnyer et al., 2007], such as between words and pictures, as in the present study. We return to this issue in Experiment 2.

The significant RS within left LO-IT however was a surprise. Some generalization of RS across stimuli has been found previously in the ventral visual pathway, such as across view-point and/or exemplars of objects [e.g., Koutstaal et al., 2001; Vuilleumier et al., 2002], from names to objects like here [though when the name immediately preceded and was concurrent with the object; Simons et al., 2003] and even for semantically related vs. unrelated words [Wheatley et al., 2005]. However this generalization has been found in more anterior (left) mid-fusiform

regions. RS in more posterior and lateral regions of the occipital cortex, like the LO-IT region here, has tended to be highly sensitive to changes in object view-point [Andersen et al., 2009; Ewbank et al., 2005], suggesting that these regions support relatively low-level shape processing. We return to this point in the General Discussion.

Finally, we found the RS effect only within left and not right LO-IT. This tendency for left-lateralization has been reported previously [Koutstaal et al., 2001; Simons et al., 2003], but more often in fusiform cortex. This finding is consistent with the hypothesis that the left hemisphere processes more abstract visual object representations than does the right hemisphere [e.g., Burgund and Marsolek, 2000; Marsolek, 1999], though it is also consistent with possible linguistic causes of our RS (such as naming), which are known to be left-hemisphere dominant. Note, however that finding a simple effect in the left but not right hemisphere is not sufficient to conclude a difference in laterality. To test for such an effect, RS within homologous regions needs to be contrasted statistically, where those regions are selected in an unbiased manner. We do this in Experiment 2.

Experiment 2

In Experiment 1, we presented only words at study. In Experiment 2, we compared RS from either pictures or words at Study (maintaining only pictures at Test) (see Fig. 1). We could therefore attempt to replicate our surprising RS from words to pictures (the Word-Picture condition) in LO-IT, and furthermore compare the size of this RS with that obtained when repeating pictures (i.e., with perceptual as well as conceptual overlap between Study and Test; the Picture-Picture condition). The presentation of both word and picture stimuli at Study also allowed us to evaluate overall activation levels for each type of stimulus within the region that demonstrated significant word-to-picture RS in Experiment 1. For example, does the left LO-IT region seen in Experiment 1 show greater activation for word than picture stimuli?

Given we were primarily interested in whether the significant RS seen in Experiment 1 was replicable, we used the peak RS co-ordinates from Experiment 1 (i.e., from an independent data set) in an ROI analysis to assess RS in the Word-Picture and Picture-Picture condition. This unbiased ROI selection also allowed us to test for a laterality effect given the RS effect in Experiment 1 was only seen in the left hemisphere. To test for concurrent signs of S-R learning, we again included a Task manipulation (Same vs. Reverse judgment), as well as adding a manipulation of “Prime-level”, whereby stimuli at Study were either seen once (Low-primed) or three times (High-primed), which has previously been demonstrated to modulate the effects of S-R learning on behavioral priming [Horner and Henson, 2009]. This resulted in a $2 \times 2 \times 2 \times 2$ pseudofactorial design with factors Stimulus-type (Picture-Picture, Word-Picture), Task (Same, Reverse), Repetition (Repeated, Novel), and Prime-level (Low-primed,

High-primed); where Novel items were randomly assigned to each Stimulus-type and Prime-level.

MATERIALS AND METHODS

Participants

Twenty-four participants (10 male) gave informed consent to participate in Experiment 2, using the same selection criteria as in Experiment 1. The mean age across participants was 25.4 years ($SD = 3.7$).

Materials

Stimuli were 384 colored images of everyday objects (and their named equivalents)—a superset of those used in Experiment 1. As in Experiment 1, they were selected so that 50% were bigger than a shoebox and 50% were smaller than a shoebox. Each stimulus was randomly assigned to one of 16 groups relating to the 16 experimental conditions, resulting in 24 stimuli per group. Stimuli assigned as Novel were randomly split into four equal-sized groups in order to provide separate unprimed baselines for High- and Low-primed and Word and Picture conditions.

Procedure

The experiment consisted of four study-test cycles, with each cycle lasting ~10 mins. During Study 96 stimuli were shown; 24 stimulus identities were presented as visual words (the Word-Picture condition) and 24 stimulus identities were presented as visual objects (the Picture-Picture condition). Half the word and picture stimuli were presented once (Low-primed) and half were presented three times (High-primed). During both word and picture presentation, participants performed the “bigger-than-a-shoebox” task. Apart from ensuring no immediate repetitions, the stimulus presentation order within each Study block was randomized so that High-primed stimuli and word and picture stimuli were approximately evenly distributed throughout. During each Test phase, the 48 stimuli from the Study phase were randomly intermixed with 48 Novel stimuli. As in Experiment 1, all the stimuli presented at Test were visual objects. As such the stimulus identities presented as pictures at Study (the Picture-Picture condition) ensured exact visual similarity between Study and Test, whereas the stimulus identities presented as words at Study (the Word-Picture condition) minimized visual similarity between Study and Test. The Same and Reverse tasks used in Experiment 1 were each repeated twice at Test, with the order of the two Test tasks counter-balanced across participants in an ABBA/BAAB manner.

fMRI Acquisition and Analysis

fMRI acquisition was identical to Experiment 1. Analysis was also identical, apart from the different design matrices reflecting the different experimental designs, as detailed below.

For each Study session, five separate regressors were modeled. Word and picture stimuli were classified according to whether they were being presented for the first time (Low-primed items and first presentation of High-primed items) or the second or third time. This resulted in four regressors plus an additional regressor for discarded trials (as well as six movement parameters). A second-stage model comprised four condition effects, corresponding to the two word and picture regressors described, as well as a further 24 subject effects. Analyses were restricted to the first presentation of word and picture stimuli to ensure activation levels were not conflated with possible repetition effects.

For each Test session (Task), 17 separate regressors were modeled—the eight experimental conditions (from factorial crossing of Stimulus-type, Repetition, and Prime-level) were split according to the particular key-press given, plus an additional regressor for discarded trials (see Methods of Experiment 1 for exclusion criteria). As in Experiment 1, six further regressors representing movement parameters were included. A second-stage model comprised 16 condition effects, corresponding to a $2 \times 2 \times 2 \times 2$ (Stimulus-type $\times$ Task $\times$ Repetition $\times$ Prime-level) repeated-measures ANOVA, as well as a further 24 subject effects.

RESULTS

Behavioral Results

After excluding 0.7% of trials with outlying RTs, the percentages of errors are shown in Table I. A $2 \times 2 \times 2 \times 2$ (Stimulus-type $\times$ Task $\times$ Repetition $\times$ Prime-level) ANOVA on errors revealed no significant main effect of Repetition, $F(1, 23) = 1.74$, $P = 0.20$, nor any significant interaction involving Repetition (F 's < 2.6 , P s > 0.20). A further 3.5% of Repeated trials were excluded from RT analysis due to incorrect responses given at Study.

Mean RTs are in Table I, and plots of RT priming are in Figure 2B. A $2 \times 2 \times 2 \times 2$ (Stimulus-type $\times$ Task $\times$ Repetition $\times$ Prime-level) ANOVA on RTs revealed several main effects and interactions (see Fig. 2B). Firstly, a significant Stimulus-type $\times$ Repetition interaction was present, $F(1, 23) = 38.75$, $P < 0.001$, showing greater priming in the Picture-Picture than Word-Picture condition. Despite this effect, significant priming was still present in the Word-Picture condition, $t(23) = 4.84$, $P < 0.001$. The latter result therefore replicates Experiment 1 in showing significant behavioral priming despite no visual similarity between Study and Test stimuli; however they extend this finding by demonstrating a significant reduction in priming compared with when the same pictorial stimulus is presented at Study and Test.

Secondly, a Task $\times$ Repetition interaction was present, $F(1, 23) = 6.53$, $P < 0.05$, reflecting greater priming in the Same than Reverse condition. This effect was significant even in the Word-Picture condition alone, $t(23) = 2.26$,

$P < 0.05$.² Together with the numerical trend to the same effect in Experiment 1, these data suggest that behavioral priming is modulated by S-R learning effects even when there is no visual similarity between Study and Test stimuli. In other words, responses become learned to stimulus representations that are sufficiently abstract to transfer from the name of an object to pictures of that object. Finally, a Repetition $\times$ Prime-level interaction was present, $F(1, 23) = 5.14$, $P < 0.05$, reflecting significantly greater priming for High- than Low-primed items. No further factors interacted significantly with Repetition, F s < 2.05 , $P > 0.17$.

In order to test for effects of Test block and Task order, the priming data were entered into a $2 \times 4 \times 2$ (Stimulus-type $\times$ Block $\times$ Order) mixed ANOVA, where the between-subject factor Order referred to the two task orders (Same-Reverse-Reverse-Same and Reverse-Same-Same-Reverse). No main effects of Block or Order reached significance, F s < 1.8 , $P > 0.18$.

fMRI Results

Whole-brain analyses

We first sought evidence for significant RS regardless of Stimulus-type (collapsed across all other conditions). This whole-brain corrected T-contrast revealed bilateral clusters in lateral occipital and inferior temporal cortex, with peak activation in the inferior temporal gyrus (right: +48, -60, -12; left: -48, -69, -6). Two further bilateral clusters were revealed: (1) a region in the posterior inferior frontal gyrus—*pars opercularis*—analogous to the pPFC region seen in Experiment 1 (left: -48, +3, +30; right: +48, +6, +27) and (2) a more anterior region in the inferior frontal gyrus—*pars triangularis*—henceforth referred to as anterior prefrontal cortex (aPFC) (left: -51, +36, +12; right: +51, +39, +15).

Given our interest in replicating the conceptual RS effects in Experiment 1, we next sought evidence for significant RS in the Word-Picture condition alone. When constraining our search within regions that demonstrated significant RS in Experiment 1 (i.e., using the RS T-contrast from that experiment, which contained 100 voxels; see Fig. 3A), two regions survived correction: (1) left pPFC (-48, +3, +27) and (2) left LO-IT (-51, -69, -3). These results therefore replicate those of Experiment 1 in demonstrating significant RS within left PFC and left occipital/temporal regions despite negligible visual similarity between items at Study and Test.

We next sought evidence for a Task $\times$ Repetition interaction. No voxels survived whole brain correction or small-volume correction using the main RS T-contrast from the ²While in the same direction numerically, any difference between Same and Reverse tasks in the Picture-Picture condition failed to reach significance, $t(23) = 1.51$, $P = 0.14$. This is surprising, given that previous research has shown strong signatures of S-R learning using this manipulation [Dobbins et al., 2004; Horner and Henson, 2008, 2009; Schnyer et al., 2007]. For this reason, we think the present lack of Same-Reverse difference in the Picture-Picture condition is most likely to be a Type II error.

previous experiment or present experiment. Thus, unlike the behavioral data, there was no evidence of S-R learning in the fMRI data that survived correction for multiple comparisons at the voxel-level. Searching for regions showing significant repetition enhancement revealed similar clusters to those in Experiment 1 (see Supporting Information Table I). Finally, no voxels showed either a significant effect of Prime-level or a Repetition $\times$ Prime-level interaction in either the whole-brain or SVC analysis.

ROI Analyses

Left LO-IT

To further test for effects of Repetition, Stimulus-type, and Task within left LO-IT, we extracted data from an ROI taken from the peak voxel identified independently in the main effect of RS in Experiment 1 ($-51, -66, 0$; see Table II). These ROI data reflect a weighted average of nearby voxel values by virtue of the Gaussian smoothing of the fMRI images. Subjecting these data to a $2 \times 2 \times 2 \times 2$ (Stimulus-type $\times$ Task $\times$ Repetition $\times$ Prime-level) ANOVA revealed a significant Stimulus-type $\times$ Repetition interaction, $F(1, 23) = 20.72, P < 0.001$, as well as main effects of Stimulus-type, $F(1, 23) = 10.15, P < 0.01$, and Repetition, $F(1, 23) = 39.51, P < 0.001$ (see Fig. 3B). Focusing on the Stimulus-type $\times$ Repetition interaction, post-hoc tests revealed significantly greater RS in the Picture-Picture than Word-Picture condition, $t(23) = 4.57, P < 0.001$. Despite this difference, significant RS was still present in the Word-Picture condition, $t(23) = 2.44, P < 0.05$. Therefore although RS is greater in left LO-IT when the same stimulus is repeated (the Picture-Picture condition), significant RS can still be seen despite minimal visual similarity between Study and Test stimuli (the Word-Picture condition).

Importantly, the main ANOVA failed to reveal a significant Task $\times$ Repetition interaction, $F(1, 23) < 1, P = 0.99$, suggesting RS within left LO-IT was resilient to changes in response between Study and Test. Given that S-R learning effects may be maximal for repetition of the same stimulus between Study and Test [Schnyer et al., 2007], we further limited our analysis to the Picture-Picture condition alone. Post-hoc tests again revealed no significant difference between the Same and Reverse condition, $t(23) = 0.04, P = 0.97$. We could therefore find no evidence for an effect of Task on RS within left LO-IT.

Finally, we conducted a $2 \times 4 \times 2$ (Stimulus-type $\times$ Block $\times$ Order) ANOVA to assess whether RS varied as a function of block or task order. There was no reliable main effect of Order, $F(1, 22) = 0.03, P = 0.87$, but there was a reliable main effect of Block, $F(2.8, 62.0) = 3.71, P < 0.05$. Paired t -tests on RS from successive blocks revealed greater RS in block 1 than 2, $t(23) = 3.10, P < 0.01$ (collapsed across Stimulus-type). No further pairwise comparisons reached significance, $t_s < 1.1, P_s > 0.27$. Thus, RS was maximal in block 1, with a consistent diminution in RS across blocks 2–4. Importantly however, there were no interactions between Order or Block and Stimulus-type, F_s

$< 2.1, P_s > 0.11$, suggesting the difference in RS across blocks did not modulate the main effects of interest.

To further explore the left LO-IT region, we extracted Study phase data from the same ROI ($-51, -66, 0$) (see Methods). This allowed us to directly compare overall activation to word and picture stimuli within this region (see Supporting Information Fig. 1 for a whole-brain analysis comparing picture and word stimuli in the Study phase). This analysis revealed significantly greater activation for picture than word stimuli, $t(23) = 4.36, P < 0.001$. Thus LO-IT responds more to the presentation of pictures compared to words.

One possibility is the Word-Picture RS effect may vary as a function of the size of the object relative to the size of the referent (shoebox). For example, for words like “elephant,” an answer might be produced without needing to perform any mental imagery, because elephants are associated with their large size in semantic memory. For other words like “football,” which are close in size to a shoebox, mental imagery may be necessary. If so, one would expect to see greater RS for objects closer in size to the referent. To test for this possibility, we conducted a further analysis in which we split objects evenly according to whether we deemed them to be “near” or “far” in size to that of a shoebox. Comparing RS in the left LO-IT maxima for near and far objects (collapsed across Task and Prime-level), t -tests found no significant difference in either the Picture-Picture, $t(23) = 1.3, P = 0.21$, or Word-Picture, $t(23) = 0.78, P = 0.44$, condition. Thus, there was no evidence that RS in left LO-IT differed as a function of proximity in size to the task referent.

Laterality Effects: Left vs. Right LO-IT

Experiments 1 and 2 both failed to find any evidence for significant Word-Picture RS in the right LO-IT. This suggests a possible laterality effect, with RS in left LO-IT possibly showing more resilience to changes in stimulus-type between Study and Test. We directly tested this hypothesis by searching for an effect-by-hemisphere interaction [see Koutstaal et al., 2001; Simons et al., 2003; Vuilleumier et al., 2002]. By defining the left and right LO-IT from independent data in Experiment 1, we avoided any bias of this effect-by-hemisphere interaction by prior selection of one hemisphere. More specifically, we defined the left LO-IT by the maxima of the Word-Picture RS contrast in Experiment 1 ($-51, -66, 0$), and the right LO-IT by flipping the x -coordinate ($+51, -66, 0$). Note that defining the left and right hemisphere ROIs in such a way does not ensure the targeting of functionally homologous regions, as the right hemisphere ROI was based on co-ordinates taken from the left hemisphere. Given no RS effects were seen in the right hemisphere in Experiment 1, however, this was the only means with which to assess laterality effects in an unbiased manner. The data in these ROIs from Experiment 2 showed significant RS in left LO-IT (as in the whole-brain analysis), $t(23) = 2.44, P < 0.05$, but not

in right LO-IT, $t(23) = 1.55$, $P = 0.13$ (Fig. 4B; cf. Fig. 3C). However, any interaction between hemisphere and RS failed to reach significance, $t(23) = 0.79$, $P = 0.44$. This lack of hemispheric interaction was also found for the Picture-Picture condition, $t(23) = 0.09$, $P = 0.93$. Thus we have no conclusive evidence for a laterality effect for Word-Picture, or Picture-Picture, RS in LO-IT.

Fusiform

Finally, given previous research has produced conflicting results with regards to conceptually-driven RS within left mid-fusiform regions (see Introduction), we extracted Test data from two ROIs: (1) left Fusiform ($-36, -48, -15$) and (2) right Fusiform ($+36, -48, -15$), with co-ordinates taken from Horner and Henson [2008]. These data failed to reveal significant RS in the Word-Picture condition in either left, $t(23) = 0.21$, $P = 0.84$, or right, $t(23) = 0.19$, $P = 0.85$, Fusiform (Fig. 4C,D) (despite reliable RS for the Picture-Picture conditions in both cases, $t(23)$'s > 6.23 , P s < 0.001). Similar analyses taking co-ordinates from Chouinard et al. [2008], Koutstaal et al. [2001] and Simons et al. [2003] also failed to reveal significant RS in the Word-Picture condition. These data are therefore consistent with the findings of Chouinard et al. [2008], suggesting that RS within left mid-fusiform is only seen if there is visual similarity between Study and Test stimuli.

DISCUSSION

Experiment 2 revealed several important findings. First, we replicated the results of Experiment 1 showing significant RS in left occipital/temporal regions (LO-IT) despite switching from word stimuli at Study to picture stimuli at Test (the Word-Picture condition). Second, compared with the Word-Picture condition, we found significantly greater RS in left LO-IT when the same picture stimulus was presented at Study and Test (the Picture-Picture condition). Third, left LO-IT showed significantly greater activation in the Study phase for picture than word stimuli. These results suggest that although significant RS from word to picture stimuli can be seen in left LO-IT, this region is also sensitive to perceptual similarity between initial and repeated presentations and responds more to picture than word stimuli. Fourth, though RS in the Word-Picture condition was only significant in the left LO-IT in both Experiments 1 and 2, when testing directly for a hemispheric difference, we failed to reveal significantly greater RS in left than right LO-IT. Importantly, these analyses were based on an unbiased ROI selection, taking the peak RS co-ordinates from Experiment 1.

Finally, further analyses failed to reveal significant RS in the Word-Picture condition in mid-fusiform cortex, where prior studies of visual object repetition have suggested some generalisation, at least in left fusiform, across views [e.g., Vuilleumier et al., 2002], exemplars [Koutstaal et al., 2001] and even from words to pictures [Simons et al., 2003]. These fusiform regions did show RS in our Picture-

Picture condition, which is consistent with a general role in visual object recognition, but our data suggest that their processing may not generalize to nonvisual (e.g., conceptual) processing, consistent with the findings of Chouinard et al. [2008].

General Discussion

The main finding of this study was that, across two independent experiments, we found reliable evidence for reduced BOLD responses in a left occipitotemporal region when an object that was initially denoted by its name was later repeated as a picture (our Word-Picture condition). The presence of such repetition suppression (RS), despite negligible visual similarity between the object names and corresponding pictorial objects, suggests that processing in this LO-IT region extends beyond the purely visual processing of objects that is normally associated with such posterior brain regions [e.g., such as the Lateral Occipital Complex, LOC; Grill-Spector, 2003; Kourtzi and Kanwisher, 2001; Malach et al., 1995]. This additional processing may reflect phonological, lexical and/or semantic processing, or it could reflect processing associated with visual imagery, as we discuss further below. Importantly, we found no evidence that this RS reflected Stimulus-Response (S-R) learning, at least as operationalised by reversing the task [Dobbins et al., 2004], despite effects of this manipulation on the behavioural repetition effects (i.e., priming).

As reviewed in the Introduction, some previous fMRI studies of visual object processing have reported generalization of RS across pictures showing different views or exemplars of an object identity [Koutstaal et al., 2001; Simons et al., 2003]; though others have argued that these findings are confounded by visual similarity [Chouinard et al., 2008; Vuilleumier et al., 2002]. RS in occipitotemporal regions has even been reported under conditions with negligible visual similarity, such as for semantically related vs. unrelated word-pairs [Gold et al., 2006; Wheatley et al., 2005]. However, all these "post-perceptual" RS effects were found in left mid-fusiform cortex, considerably more anterior and medial than the present LO-IT region reported here. While we did find RS for repeated pictures of objects in fusiform cortex (in our Picture-Picture condition in Experiment 2), we did not find any generalization from words to pictures [consistent with Chouinard et al., 2008].³

³A related study that implicated the fusiform cortex in post-perceptual processing used priming of object pictures by their names, as here, though using auditory names immediately prior to, and concurrent with, the picture [Simons et al., 2003]. It is not clear why we did not see a similar RS effect, though one possibility is that participants in the Simons et al. study started to perform the bigger/smaller categorization on the name prior to onset of the picture. Then all they need do is decide whether the name matches the subsequent picture: if it does (primed case), they can execute the response prepared to the name; if not (unprimed case), they must reprocess the picture in detail. This may result in greater activity in the unprimed case, i.e., RS.

Although we found significant RS from words to pictures within left LO-IT, we found greater RS still when repeating the same stimulus between Study and Test (i.e., our Picture-Picture condition). RS in this region therefore appears to show some sensitivity to visual similarity between initial and subsequent presentations. This is consistent with its greater overall response to pictures than to words (e.g., for initial presentations of each in the Study phase of Experiment 2).

Postperceptual Processing in LO-IT?

One possible explanation for our findings is that the LO-IT region (and possibly inferior frontal region) is involved in retrieving semantic information, for example about the real-life size of objects (as required by the task). Such an explanation would fit with theories suggesting that sensory-specific semantic information is stored in regions associated with perception of those properties [e.g., Martin, Simmons and John, 2008]. Interestingly, if such semantic processes depended on task demands, then tasks that require retrieval of different object-related information should result in RS in different cortical regions. For example, tasks involving decisions about object manipulability might be expected to produce RS within more middle temporal as well as premotor regions [e.g., Beauchamp and Martin, 2007; Chao et al., 1999]. Such task-dependency may explain why Chouinard et al. [2008] failed to demonstrate significant RS in any occipitotemporal region following a change in exemplar: the naming task used in their study might not have recruited the same post-perceptual task-dependent processes as those engaged by the present “bigger-than-shoebox” task. However, such an explanation is difficult to reconcile with our finding of greater RS in the Picture-Picture relative to the Word-Picture condition, unless one assumes that greater semantic retrieval occurs for pictures than for words, or that LO-IT is recruited in the service of object recognition as well as in the retrieval of information regarding object size.

A further possibility is that the LO-IT activity does relate to visual processing, but processing that can be triggered both directly by a picture and indirectly by visual imagery (triggered by a word, for example) (see Kosslyn, 1994; Kosslyn et al., 1995 for a discussion of the cortical regions recruited during mental imagery). Indeed, mental imaging of the referent of a word is likely to be helpful for our “bigger-than-shoebox” size judgment task. As such, the recruitment of mental imagery may serve as the basis with which participants extract information relating to an object size/shape when a word is presented. Prior studies have reported activity within higher order visual regions associated with visual imagery of objects, faces, and scenes [D’Esposito et al., 1997; Ishai et al., 2000]. These studies have also shown a predominantly left lateralized response during mental imagery, consistent with our findings of RS

in left but not right LO-IT. Furthermore, RS in earlier visual regions (i.e., V3/V4) has been shown to result from the mental imagery of tilted lines [Mohr et al., 2009], suggesting that RS in visual cortex can occur even in the absence of a physical visual stimulus. Finally, RS in an LO-IT region has recently been seen despite switching from tactile to visual presentation of an object, supporting the idea that object representations can be accessed by nonpictorial stimuli [Tal and Amedi, 2009].

Again however, this account would not appear fully satisfactory, because it is generally assumed that the same type of imagery would need to be used at Study and Test to explain RS for a given object. It is unclear whether the image of an object generated by a participant in response to a word at Study should necessarily overlap with the visual image (picture) presented at Test. For example, different participants are likely to generate different images for the word “umbrella” at Study (e.g., an image of an open black umbrella), which could bear little resemblance to the specific picture of an umbrella used at Test (e.g., a red closed umbrella). Although this is an issue, it is noteworthy that objects were primarily depicted from canonical viewpoints [Blanz et al., 1999], plausibly increasing the probability of overlap between mental and pictorial images.

A further though not mutually exclusive possibility is that the RS in LO-IT reflects a modulation of re-entrant processing, subsequent to initial visual processing [Henson, 2003]. Such re-entrant processing may reflect interactions with more anterior temporal and prefrontal regions, for example in the controlled retrieval of semantic information [Badre and Wagner, 2007]. This interpretation is consistent with findings that TMS to left inferior prefrontal cortex can prevent reliable RS during semantic decisions about objects in lateral temporal regions [Wig et al., 2005; though we note that the region targeted by TMS was more anterior to that seen in this study] and with preliminary evidence that patients with a left temporal lobectomy showed no reliable RS in regions posterior to the lesion during visual object naming [Martin, 2007a].

Alternative Explanations

A different explanation for our findings is that LO-IT RS reflects facilitation of processes relating to phonological and/or lexical retrieval associated with covert naming of both picture and word stimuli. Indeed, a recent study proposed that a similar posterior occipital/temporal region is involved in the phonological processing of regular words [Seghier et al., 2008]. Note that this region is posterior to the “Visual Word Form Area,” which has been associated more with the orthographic processing of visual word form [Cohen and Dehaene, 2004; Cohen et al., 2000]. Given we also found significant word-to-picture RS in posterior regions of the inferior prefrontal cortex, a region also thought to be involved in phonological processing [Pol-drack et al., 1999], it is plausible that prefrontal and LO-IT

regions were working in concert to covertly name both word and picture stimuli.

However, if facilitated phonological retrieval were the cause of the LO-IT RS, it is unclear why we should find significantly greater RS in the Picture-Picture (repetition priming) condition relative to the Word-Picture condition. One possibility is that the Picture-Picture condition results in a greater match between phonological encoding and retrieval operations, owing, for example, to greater naming correspondence in the Picture-Picture than Word-Picture condition. However, it is unclear how such an account would explain the greater overall activation for picture than word stimuli during the Study phase, when no repetition has occurred. Indeed, if our LO-IT region is involved in phonological processing one might expect greater activation for word than picture stimuli.

A final possibility is that RS may be incidental to critical object recognition processes, reflecting for example reduced attention to a visual stimulus once a decision has been made. This would be consistent with an fMRI study that artificially slowed-down object recognition, so that the BOLD response pre and postrecognition could be separated [Eger et al., 2007]: this study found that the RS (associated with priming object identification with an object name) in fusiform cortex occurred after the recognition point, not before. This possibility questions the causal role of fMRI RS in object recognition.

Although we have considered several possibilities that would seem to explain our surprising LO-IT word-to-picture RS effect, we believe that such RS most plausibly reflects the extraction of task-relevant information relating to object size, either through the extraction of sensory-specific semantic information or through mental imagery processes. If this is the case, such effects are likely to be sensitive to the particular experimental and task manipulations used. For example, our present “bigger-than-shoe-box” task may have encouraged a greater emphasis on visual imagery processes regardless of picture or word presentation. This emphasis on visual imagery may have been increased further by our use of words (rather than pictures) at Study, given that object names are unlikely to provide as much information about object size as are object pictures. This could be tested by an experiment investigating transfer of RS from pictures at Study to words at Test. Furthermore, the intermixing of word and picture stimuli in the Study phase of Experiment 2, and the completion of four Study-Test cycles, may have encouraged participants to process the stimuli in a similar manner (e.g., increasing the possibility of covert naming following picture presentation, or visual imagery following word presentation), thereby increasing processing overlap in our word-to-picture manipulation. Indeed, the present word-to-picture across-format change may have increased the amount of visual imagery relative to previous research that focused on within-format changes across differing exemplars [Chouinard et al., 2008; Koutstaal et al., 2001; Simons et al., 2003], possibly contributing to the differences

between the present and previous results. Finally, RS effects in occipitotemporal cortex are likely to be dependent on the particular stimuli presented. In the current experiment, a mixture of both animate and inanimate objects were presented (with more inanimate than animate). Given dissociations have been found between more medial and lateral aspects of the fusiform cortex [Martin, 2007b] according to similar categorical groupings, it is possible that stimulus differences might also explain the differences between the results of the present and previous studies. Further research is needed to reveal the extent to which the present RS effects are sensitive to the particular task and stimuli used.

Laterality Effects

As previously noted, prior research has largely implicated left hemisphere occipitotemporal regions in higher order perceptual and/or conceptual processing. Conversely, right hemisphere occipital/temporal regions have been proposed to process more specific object representations [e.g., more viewpoint-dependent representations; Burgund and Marsolek, 2000; Marsolek, 1999]. In these data, using an unbiased selection of LO-IT coordinates, we failed to find a reliable hemisphere-by-RS interaction, in either the Word-Picture or Picture-Picture condition. Nonetheless, both of the present experiments were consistent with prior claims of more abstract processing of objects in the left hemisphere, with RS from words to pictures reaching significance in left but not right LO-IT.

S-R Learning Effects

In the present experiments we attempted to control for S-R learning effects using the Same vs. Reverse contrast introduced by Dobbins et al. [2004]. In both Experiments 1 and 2 we failed to find any significant difference between the Same and Reverse condition in LO-IT (despite a trend for greater RS in the Same condition in Experiment 1). Importantly, in Experiment 2, we found greater behavioural priming in the Same than Reverse condition of the Word-Picture condition. Therefore the lack of any effect of response repetition on RS in LO-IT in Experiment 2 was not due to an absence of an S-R learning effect in general (though this conclusion should be tempered by the fact that we failed to find a significant S-R learning effect on RS in inferior prefrontal cortex). These results support those of Horner and Henson [2008], Wig et al. [2009], and Race et al. [2009] suggesting that RS in occipital/temporal regions is largely resilient to switches in task and/or response. Our finding of significant word-to-picture RS in LO-IT is therefore unlikely to have resulted from S-R learning effects.

This conclusion however must be tempered by the recent finding that the Same vs. Reverse contrast does not sufficiently control for all possible S-R learning effects [Horner and Henson, 2009]. This is because highly abstract

response representations, in this case a task-specific size classification (e.g., “bigger/smaller”), are capable of entering into S-R bindings and affecting subsequent behavioral priming. Although the Same vs. Reverse contrast suitably controls for S-R learning at more specific response representations (e.g., a left/right motor response and yes/no decision), it does not require a reversal of response at the level of classification (e.g., a lion is always classified as “bigger” irrespective of the direction of the question). Therefore, although the present word-to-picture RS effects in LO-IT are unlikely to be driven by S-R learning effects, we cannot conclusively rule out any contribution from such effects. In order to unequivocally rule out the presence of S-R learning effects, future research needs to control for S-R learning in a more effective manner. One possibility is changing the size of the particular referent in the “bigger-than” task (e.g., from a “shoebox” to a “pencil case”). For objects of intermediate size, such a manipulation requires the reversal of response at more specific representational levels as well as the more abstract level of size classification [see Experiment 7 of Horner and Henson, 2009].

CONCLUSION

The present experiments provide evidence for significant repetition suppression within occipitotemporal regions despite negligible visual similarity between initial and subsequent stimulus presentation, as evident by the significant RS in a left lateral occipital/inferior temporal region in our word-to-picture condition. Such RS could reflect the repetition of “post-perceptual” processes, plausibly relating to the extraction of task-relevant information relating to object size, either through the extraction of sensory-specific semantic information or through mental imagery processes. While future experiments will be able to better determine the precise cause, the present findings reinforce the potential complexity of processing within traditional visual regions, at least as indexed by fMRI and repetition suppression.

ACKNOWLEDGMENTS

The authors thank anonymous reviewers for their helpful comments.

REFERENCES

- Andresen DR, Vinberg J, Grill-Spector K (2009): The representation of object viewpoint in human visual cortex. *Neuroimage* 45:522–536.
- Andrews TJ, Ewbank MP (2004): Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage* 23:905–913.
- Badre D, Wagner AD (2007): Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* 45:2883–2901.
- Beauchamp MS, Martin A (2007): Grounding object concepts in perception and action: Evidence from fMRI studies of tools. *Cortex* 43:461–468.
- Blanz V, Tarr MJ, Bulthoff HH (1999): What object attributes determine canonical views? *Perception* 28:575–600.
- Burgund ED, Marsolek CJ (2000): Viewpoint-invariant and viewpoint-dependent object recognition in dissociable neural subsystems. *Psychonomic Bull Rev* 7:480–489.
- Chao LL, Haxby JV, Martin A (1999): Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci* 2:913–919.
- Chouinard PA, Morrissey BF, Köhler S, Goodale MA (2008): Repetition suppression in occipital-temporal visual areas is modulated by physical rather than semantic features of objects. *Neuroimage* 41:130–144.
- 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, Lehericy S, Dehaene-Lambertz G, Henaff M-A, et al. (2000): The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123:291–307.
- D’Esposito M, Detre JA, Aguirre GK, Stallcup M, Alsop DC, Tippet LJ, et al. (1997): A functional MRI study of mental image generation. *Neuropsychologia*, 35:725–730.
- Denkinger B, Koutstaal W (2009): Perceive-decide-act, perceive-decide-act: How abstract is repetition-related decision learning? *J Exp Psychol: Learning Memory Cogn* 35:742–756.
- Dobbins IG, Schnyer DM, Verfaellie M, Schacter DL (2004): Cortical activity reductions during repetition priming can result from rapid response learning. *Nature* 428:316–319.
- Eger E, Henson RN, Driver J, Dolan RJ (2004): BOLD repetition decreases in object-responsive ventral visual areas depend on spatial attention. *J Neurophysiol* 92:1241–1247.
- Eger E, Schyns PG, Kleinschmidt A (2004): Scale invariant adaptation in fusiform face-responsive regions. *Neuroimage* 22:232–242.
- Eger E, Schweinberger SR, Dolan RJ, Henson RN (2005): Familiarity enhances invariance of face representations in human ventral visual cortex: fMRI evidence. *Neuroimage* 26:1128–1139.
- Eger E, Henson RN, Driver J, Dolan RJ (2007): Mechanisms of top-down facilitation in perception of visual objects studied by fMRI. *Cerebral Cortex* 17:2123–2133.
- Ewbank MP, Andrews TJ (2008): Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. *Neuroimage* 40:1857–1870.
- Ewbank MP, Schluppeck D, Andrews TJ (2005): fMR-adaptation reveals a distributed representation of inanimate objects and places in human visual cortex. *Neuroimage* 28:268–279.
- Friston KJ, Penny W, Phillips C, Kiebel S, Hinton G, Ashburner J (2002): Classical and bayesian inference in neuroimaging: Theory. *Neuroimage*, 16:465–483.
- Gold BT, Balota DA, Jones SJ, Powell DK, Smith CD, Andersen AH (2006): Dissociation of automatic and strategic lexical-semantic: Functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *J Neurosci* 26:6523–6532.
- Grill-Spector K (2003): The neural basis of object perception. *Curr Opin Neurobiol* 13:159–166.
- Grill-Spector K, Henson R, Martin A (2006): Repetition and the brain: Neural models of stimulus-specific effects. *Trends Cogn Sci* 10:14–23.
- Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y, Malach R (1999): Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24:187–203.

- Grill-Spector K, Malach R (2001): fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst)* 107(1-3):293–321.
- Grill-Spector K, Malach R (2004): The human visual cortex. *Ann Rev Neurosci* 27:649–677.
- Hayworth KJ, Biederman I (2006): Neural evidence for intermediate representations in object recognition. *Vis Res* 46:4024–4031.
- Henson R (2003): Neuroimaging studies of priming. *Prog Neurobiol* 70:53–81.
- Horner AJ, Henson RN (2008): Priming, response learning and repetition suppression. *Neuropsychologia* 46:1979–1991.
- Horner AJ, Henson RN (2009): Bindings between stimuli and multiple response codes dominate long-lag repetition priming in speeded classification tasks. *J Exp Psychol: Learning Memory Cogn* 35:757–779.
- Ishai A, Ungerleider LG, Haxby JV (2000): Distributed neural systems for the generation of visual images. *Neuron* 28:979–990.
- James TW, Humphrey GK, Gati JS, Menon RS, Goodale MA (2002): Differential effects of viewpoint on object-driven activation in dorsal and ventral streams. *Neuron* 35:793–801.
- Kosslyn SM (1994): *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press.
- Kosslyn SM, Behrmann M, Jeannerod M (1995): The cognitive neuroscience of mental imagery. *Neuropsychologia* 33:1335–1344.
- Kourtzi Z, Kanwisher N (2000): Cortical regions involved in perceiving object shape. *J Neurosci* 20:3310–3318.
- Kourtzi Z, Kanwisher N (2001): Representation of perceived object shape by the human lateral occipital complex. *Science* 293:1506–1509.
- Koutstaal W, Wagner AD, Rotte M, Maril A, Buckner RL, Schacter DL (2001): Perceptual specificity in visual object priming: Functional magnetic resonance imaging evidence for a laterality difference in fusiform cortex. *Neuropsychologia* 39:184–199.
- Logan GD (1990): Repetition priming and automaticity: Common underlying mechanisms? *Cogn Psychol* 22:1–35.
- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, et al. (1995): Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA* 92:8135–8139.
- Marsolek CJ (1999): Dissociable neural subsystems underlie abstract and specific object recognition. *Psychol Sci* 10:111–118.
- Martin A (2007a): Can increased neural synchrony account for repetition-related improvements in processing and behaviour? Evidence from MEG. Paper presented at the 13th Annual Meeting of the Organization for Human Brain Mapping, Chicago.
- Martin A (2007b): The representations of object concepts in the brain. *Ann Rev Psychol* 58:25–45.
- Martin A, Simmons WK, John HB (2008): *Structural Basis of Semantic Memory Learning and Memory: A Comprehensive Reference*. Oxford: Academic Press. pp113–130.
- Mohr HM, Linder NS, Linden DE, Kaiser J, Sireteanu R (2009): Orientation-specific adaptation to mentally generated lines in human visual cortex. *Neuroimage* 47:384–391.
- Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD (1999): Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10:15–35.
- Race EA, Shanker S, Wagner AD (2009): Neural priming in human frontal cortex: Multiple forms of learning reduce demands on the prefrontal executive system. *J Cogn Neurosci* 21:1766–1781.
- Rotshtein P, Henson RNA, Treves A, Driver J, Dolan RJ (2005): Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat Neurosci* 8:107–113.
- Schacter DL, Buckner RL (1998): Priming and the brain. *Neuron* 20:185–195.
- Schnyer DM, Dobbins IG, Nicholls L, Schacter DL, Verfaellie M (2006): Rapid response learning in amnesia: Delineating associative learning components in repetition priming. *Neuropsychologia* 44:140–149.
- Schnyer DM, Dobbins IG, Nicholls L, Davis S, Verfaellie M, Schacter DL (2007): Item to decision mapping in rapid response learning. *Memory Cogn* 35:1472–1482.
- Seghier ML, Lee HL, Schofield T, Ellis CL, Price CJ (2008): Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. *Neuroimage* 42:1226–1236.
- Simons JS, Koutstaal W, Prince S, Wagner AD, Schacter DL (2003): Neural mechanisms of visual object priming: Evidence for perceptual and semantic distinctions in fusiform cortex. *Neuroimage* 19:613–626.
- Tal N, Amedi A (2009): Multisensory visual-tactile object related network in humans: Insights gained using a novel crossmodal adaptation approach. *Exp Brain Res* 198:165–182.
- Vinckier F, Dehaene S, Jobert A, Dubus JP, Sigman M, Cohen L (2007): Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron* 55:143–156.
- Vuilleumier P, Henson RN, Driver J, Dolan RJ (2002): Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nat Neurosci* 5:491–499.
- Wagner AD, Desmond JE, Demb JB, Glover GH, Gabrieli JD (1997): Semantic repetition priming for verbal and pictorial knowledge: A functional mri study of left inferior prefrontal cortex. *J Cogn Neurosci* 9:714–726.
- Wagner AD, Koutstaal W, Maril A, Schacter DL, Buckner RL (2000): Task-specific repetition priming in left inferior prefrontal cortex. *Cereb Cortex* 10:1176–1184.
- Wheatley T, Weisberg J, Beauchamp MS, Martin A (2005): Automatic priming of semantically related words reduces activity in the fusiform gyrus. *J Cogn Neurosci* 17:1871–1885.
- Wig GS, Grafton ST, Demos KE, Kelley WM (2005): Reductions in neural activity underlie behavioral components of repetition priming. *Nat Neurosci* 8:1228–1233.
- Wig GS, Buckner RL, Schacter DL (2009): Repetition priming influences distinct brain systems: Evidence from task-evoked data and resting-state correlations. *J Neurophysiol* 101:2632–2648.