NeuroImage 49 (2010) 1919–1932

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Identifying objects impairs knowledge of other objects: A relearning explanation for the neural repetition effect

Chad J. Marsolek [a,*], Rebecca G. Deason [a], Nicholas A. Ketz [a], Pradeep Ramanathan [a], Edward M. Bernat [a], Vaughn R. Steele [a], Christopher J. Patrick [a], Mieke Verfaellie [b], David M. Schnyer [c]

[a] Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455, USA
[b] Boston VA Healthcare System, and Boston University School of Medicine, MA, USA
[c] University of Texas at Austin, and Athinoula A. Martinos Center for Biomedical Imaging, MGH/MIT/HMS, MA, USA

ARTICLE INFO

Article history:
Received 20 July 2009
Revised 25 August 2009
Accepted 31 August 2009
Available online 8 September 2009

ABSTRACT

Different items in long-term knowledge are stored in the neocortex as partially overlapping representations that can be altered slightly with usage. This encoding scheme affords well-documented benefits, but potential costs have not been well explored. Here we use functional magnetic resonance imaging (fMRI), neurocomputational modeling, and electrophysiological measures to show that strengthening some visual object representations not only enhances the subsequent ability to identify those (repeated) objects—an effect long known as repetition priming—but also impairs the ability to identify other (non-repeated) objects—a new effect labeled antipriming. As a result, the non-repeated objects elicit increased neural activity likely for the purpose of reestablishing their previously weakened representations. These results suggest a novel reevaluation of the ubiquitously observed repetition effect on neural activity, and they indicate that maintenance relearning may be a crucial aspect of preserving overlapping neural representations of visual objects in long-term memory.

Introduction

The neocortex encodes different visual objects using distributed, partially overlapping representations (Haxby et al., 2001; Ishai et al., 1999; Tanaka, 1993; Tanaka, 2003) and these representations undergo learning-related changes after their access and usage (Kobatake et al., 1998; Sigala and Logothetis, 2002). This coding scheme affords benefits such as efficient storage of many items and enhanced generalization abilities (McClelland et al., 1995; McClelland and Rumelhart, 1985). However, potential costs of this scheme have not been well examined (McClelland et al., 1995). Here we provide evidence that visual object representations are strengthened after they are used to identify objects, and the changes to those representations also entail weakening of other representations with which they overlap. Strengthening one set of visual object representations not only enhances the subsequent ability to identify those (repeated, or "primed") objects but also impairs the subsequent ability to identify other (non-repeated, or "antiprimed") objects. We use functional magnetic resonance imaging (fMRI) to show that a consequence of the weakening of representations of non-repeated objects is that those objects elicit increased neural activity in order to reestablish their weakened representations. Neurocomputational models provide a clear mechanistic hypothesis for the origins of this effect, and event-related potentials (ERPs) corroborate the interpretation of the fMRI results as reflecting relearning of previously weakened representations. These findings indicate that the ubiquitously observed neural repetition effect (lower neural activation for repeated items than for non-repeated items; for reviews, see Grill-Spector et al., 2006; Henson, 2003) may be reevaluated. It may not reflect changes solely to the neural representations of the repeated stimuli (e.g., adaptation or suppression) but rather the reestablishment of previously weakened representations of the non-repeated stimuli.

For example, the distributed visual representations for desk and piano would overlap if they share some visual features (e.g., moderately complex features like shared front legs). If so, identifying a desk and strengthening its representation should cause the overlapping features to be more diagnostic for identifying a desk as well as less diagnostic for identifying a piano. Hence, subsequent identification of a desk would be enhanced (primed), and subsequent identification of a piano would be impaired (antiprimed). As a result, a piano would elicit increased neural activity after identification, which is needed to reestablish its previously weakened neural representation.

We measured both enhanced object identification and impaired object identification relative to a common, unbiased baseline condition (see Fig. 1A). Baseline object identification was measured in an early phase of the experiments, during which participants

* Corresponding author. Fax: 612 626 2079.
E-mail address: chad.j.marsolek-1@umn.edu (C.J. Marsolek).
URL: http://levels.psych.umn.edu (C.J. Marsolek).

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.08.063

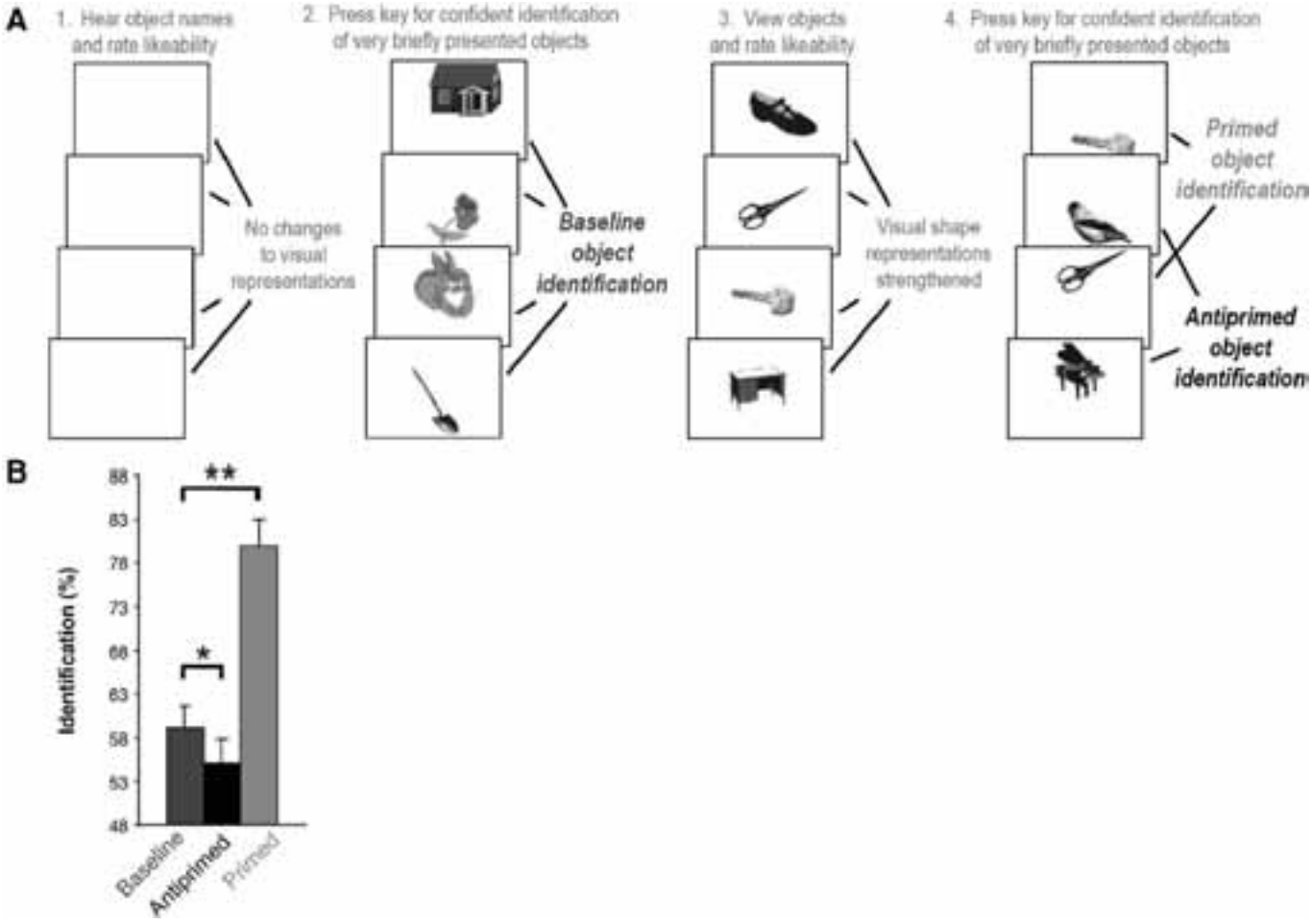

Fig. 1. fMRI experiment design and behavioral results. (A) The experiment had four phases, and example visual stimuli are shown for each phase. First, participants stared at a blank visual display while they heard the names of 50 objects, each of which they rated for likeability. Second, participants attempted to visually identify 100 very briefly presented pictures of familiar objects that were neither primed nor antiprimed. They pressed one of two buttons when they confidently identified an object. Baseline object identification and baseline neural activity were measured during these trials, because they occurred after a period of time in which no changes in visual representations of objects should have taken place (during the first phase). Third, participants viewed and rated the likeability of each of a new set of 50 objects that were presented visually for long exposures, enabling strengthening of their visual representations. Fourth, participants attempted to visually identify 100 very briefly presented objects, 50 of which were repeated from the preceding phase (primed objects) and 50 of which were not presented in any form previously in the experiment (antiprimed objects), pressing one of two buttons when they confidently identified an object. Primed and antiprimed object identification and neural activity were measured during these trials. (B) The behavioral results from the fMRI experiment are depicted. Mean identification rates (+ standard errors of the mean) are plotted as a function of test presentation condition (baseline, antiprimed, and primed). Primed objects were identified significantly more often than baseline objects, as effect known as priming. More importantly, antiprimed objects were identified significantly less often than baseline objects, a new impairment effect we label antipriming. *$p<.05$. **$p<.001$.

identified very briefly presented visual objects (e.g., shovel, rabbit, flower, etc.). This condition served as baseline because identification occurred when little or no recent changes to visual object representations had taken place (participants stared at a blank display and made judgments of spoken words for a period of time before baseline performance was measured). Then, participants viewed and rated a new set of visual objects (e.g., desk, key, scissors, etc.) each displayed for a long time, enabling their visual shape representations to be strengthened. Finally, participants identified very briefly presented visual objects, half of which were repeated from the preceding phase (primed objects; e.g., scissors, etc.) and half of which were new, non-repeated objects (antiprimed objects, according to our theory; e.g., piano, etc.). Enhanced identification of the repeated objects relative to the baseline measure reflected the benefit long known as priming (Schacter, 1987; Tulving and Schacter, 1990), and impaired identification of the non-repeated objects relative to the baseline measure reflected a new impairment effect that we label antipriming (Marsolek, 2008; Marsolek et al., 2006). The baseline condition was new in this study; typically the primed condition is compared against what we label the "antiprimed" condition to measure effects of repetition. No neuroimaging study to date has included the unbiased baseline condition, which is necessary to tease apart priming and antipriming, thus we used this baseline condition to examine both priming and antipriming in an fMRI experiment, a control experiment, neurocomputational models, and an ERP experiment.

Two virtues of our procedure helped to assure that the results were due to changes in the visual object representations of interest. First, utilizing different tasks between the first presentations (rating objects) and the second presentations (visually identifying briefly presented objects) assured that any rapid learning of stimulus–response associations (Dobbins et al., 2004; Horner and Henson, 2008; Schnyer et al., 2006) could not be responsible for repetition effects. Second, having participants visually identify very briefly presented objects increased the degree to which purely visual processes were engaged, relative to tasks in which stimuli are presented for longer exposures or participants must access additional information about the objects (Buckner et al., 1998; Sayres and Grill-Spector, 2006; van Turennout et al., 2003).

Materials and methods

Participants

For the fMRI experiment, 20 participants were scanned at the Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital on a 3T Siemens Trio system using a Siemens 8-channel head coil. All participants gave written, informed consent in accordance with procedures and protocols approved by the human subjects review committee of Boston University Medical Center, Massachusetts General Hospital and Boston VA Healthcare System.

For the texture control experiment, 40 participants were tested individually. All participants gave written, informed consent in accordance with procedures and protocols approved by the human subjects review committee of the University of Minnesota.

For the ERP experiment, 40 participants were tested individually in a sound-attenuated EEG laboratory chamber. All participants gave written, informed consent in accordance with procedures and protocols approved by the human subjects review committee of the University of Minnesota. Behavioral data from one participant were lost due to programming error, thus behavioral effects were tested on 39 participants.

Stimuli

Two hundred fifty gray-scaled depictions of familiar visual objects were used as visual stimuli, and 250 auditory recordings of the entry-level names (Jolicœur et al., 1984) of those objects were used as auditory stimuli. None of the visual objects was highly visually similar to another in the set. Within a participant, the visual objects used to represent baseline, primed, and antiprimed conditions were balanced on rated frequency of encountering the objects in everyday life, the typicality of those object images with respect to their object categories, and on agreement of the best names for the objects. Full counterbalancing in each experiment assured that each visual object represented each experimental condition an equal number of times across participants.

fMRI experiment procedure

An experimental session had four different phases. Fig. 1A depicts examples of visual stimuli presented during the four phases and the conditions that they represented.

In the first phase, participants heard auditory recordings of the names of 50 objects while staring at a blank screen. They judged whether they liked or disliked that kind of thing, considering only the meanings of the objects, not how they may sound or look. They pressed one button for like and a different button for dislike.

In the second phase, baseline object identification performance was measured. Participants attempted to visually identify 100 very briefly presented visual objects. The objects were presented slightly above or below the center of the display, in order to keep identification ability lower than the 100% ceiling rate. All of the objects presented in this phase were different from the names of objects that were heard in the preceding phase. In each object presentation trial, a fixation point appeared for 500 ms, then an object appeared for 15 ms, then a blank display appeared for 2485 ms. (In control trials, the 500-ms fixation point was followed by another 15 ms of fixation point display, followed by a blank display for variable times, 985–3985 ms, which helped to achieve optimal deconvolution of the hemodynamic response.) Participants pressed one button when they confidently identified an object and another button when they could not confidently identify an object. This task was used to avoid motion artifacts that would have occurred from speaking aloud identification responses.

In the third phase, participants viewed 50 visual objects that were presented centrally for a relatively long time. This enabled easy, successful visual identification, and the visual shape representations for these objects presumably were strengthened. In each trial, a fixation point appeared for 500 ms, and then an object appeared for 3 s. All of these objects were different from the visual objects and auditory names of objects presented in the preceding phases. Participants judged whether they liked or disliked the kind of thing depicted, as they did in the first phase, considering only the meanings of the objects, not how they may look or sound.

In the fourth phase, antiprimed and primed object identification performance was measured. Participants attempted to visually identify 100 very briefly presented visual objects, and the presentation procedure was the same as used in the second phase. Fifty trials in this phase were used to measure priming. They involved the same 50 objects as presented visually during the third phase, thus they should have shape representations that were strengthened during the preceding phase. The other 50 trials in this phase (intermixed with primed object trials) were used to measure antipriming. They involved 50 objects that were different from the visual objects presented previously (and they did not correspond with the auditory names heard in the first phase). Thus, when their shape representations overlapped with shape representations strengthened by visual object processing in the preceding phase(s), then by hypothesis, their shape representations should have been weakened by the previous strengthening of other representations. As in the second phase, participants pressed one button when they confidently identified an object and another button when they could not confidently identify an object.

With this paradigm, a behavioral effect of priming can be measured as significantly enhanced identification of primed objects relative to the baseline condition, and antipriming can be measured as significantly impaired identification of antiprimed objects relative to the baseline condition. Given that the task was to press a button after making a judgment about whether an object was confidently identified, response times would have been ambiguous reflections of processing, and they were not collected or analyzed. Two-tailed, paired *t*-tests were conducted to test identification effects, with an alpha value of .05.

To enable fMRI scanning with relatively high spatial resolution ($2\times2\times2.5$ mm) in visual shape processing regions, we oriented the MRI slice prescriptions on the ventral visual pathway. All participants were scanned during phases 2 and 4 to measure activity in control (non-object), baseline, primed, and antiprimed conditions. Object identification was covert (signaled by a button press) instead of overt to avoid motion artifacts from speaking aloud (Chao et al., 2002; van Turennout et al., 2000). Participants viewed pictures projected onto a screen at the end of the bore through a mirror mounted on the head coil, and the procedures were programmed and presented with E-Prime (Psychology Software Tools, Inc. Pittsburgh, PA). As is standard, four dummy volumes were collected to allow for T1 equilibration, and head motion was restricted using pillows and foam inserts. Control trials of a fixation point followed by a blank presentation were interspersed throughout the phases 2 and 4 in an order and timing duration determined to allow for optimal deconvolution of the hemodynamic response (optseq2, MGH NMR Center, Charlestown, MA). At the beginning of the scan session and then again between the first and second phases, a high-resolution anatomical scan (3D T1-weighted MPRAGE volumes) was acquired for anatomical co-registration with fMRI data (TR = 7300 ms, TE = 3.2 or 3.0 ms, flip angle = 7°, slice thickness = 1.3 mm, 128 slices, FOV = 256×256 mm). Functional EPI images were acquired during each of the identification phases consisting of 17 axial slices oriented to best cover the temporal–occipital cortex (TR = 1500 ms, TE = 30 ms, $2\times2\times2.5$ mm voxels with a 0.2 mm gap and 96×96 acquisition matrix). The temporal and spatial resolutions of the

voxels during the functional runs did not allow for whole brain coverage.

fMRI analyses

The approach to the fMRI data analyses was based in part on work using masked priming with words (Dehaene et al., 2001, 2004) and pictures (Eddy et al., 2007). Data were processed using FSL (Smith et al., 2004) and Freesurfer (MGH NMR Center, Charlestown, MA). For fMRI analysis, images were motion-corrected, smoothed with a 5 mm Gaussian filter, high-pass filtered, and "prewhitened" before event-related responses were estimated using event-related convolution with an ideal hemodynamic response represented by a gamma function and its temporal derivative. Explanatory variables were modeled for each of the following five conditions: identified baseline objects, nonidentified baseline objects, identified primed objects during phase 4, identified antiprimed objects during phase 4, and nonidentified objects during phase 4. A second-level analysis was conducted on each individual subject by registering each subject to the standard MNI152 template and combining all identified object trials from phases 2 and 4 (i.e., identified baseline, identified primed, and identified antiprimed object trials) and contrasting them against the mean activation measured after all modeled conditions were removed (i.e., the mean activation for non-object trials). A group-level map of this contrast was calculated by examining these individual lower-level contrasts in a higher-order statistical parametric map utilizing the FLAME (FMRIB's Local Analysis of Mixed Effects) technique implemented through FSL and thresholded at $p<.05$, cluster corrected. This contrast revealed the regions involved in object identification.

Active voxels identified in the group-level analysis were then searched for significant priming and antipriming effects utilizing a functional ROI approach. First, anatomical regions (defined by the Freesurfer parcellation) within the ventral temporal stream that overlapped with the functional activation associated with object identification were defined by taking an anatomically labeled average brain (Fischl et al., 2004) that was aligned with the standard space functional images from the current study and overlaying the group functional map. Only active voxels within each of the 10 critical object identification regions were examined for between-condition differences (Table 1).

Average time courses for each voxel and for each condition were computed for each of these ROIs utilizing a selective averaging technique (http://www.poldracklab.org/software). These responses were baselined by subtracting the average of the first 3 pre-stimulus time points and then an average response was calculated for the time period from 4.5 to 7.5 seconds post stimulus onset. This average response was examined for between-condition differences using paired *t*-tests. Three regions revealed significant effects of the typical neural repetition effect (unprimed>primed, which according to our paradigm is antiprimed>primed): left lateral occipital–temporal cortex (−44, −79, 6), left ventral temporal cortex including the fusiform gyrus (−38, −56, −20), and right inferior occipital cortex (48, −72, −12). Two-tailed, paired *t*-tests (antiprimed>baseline; baseline>primed) were used to examine whether the typical neural repetition effect in these regions reflected antipriming or priming, with an alpha value of .05.

Table 1
Object identification ROIs.

Brain region	BA	MNI coordinates (*x*, *y*, *z*)	Size (voxels)	Max *z* stat
Left lateral occipital–temporal	37	−44, −79, 6	181	4.90
Left ventral occipital–temporal	19	−38, −56, −20	77	5.71
Right inferior occipital	19	48, −72, −12	93	5.40
Right ventral occipital–temporal	37	40, −58, −20	56	4.13
Left insular gyrus	47	−35, 9, −7	22	3.50
Left anterior insula	47	−28, 18, −12	36	3.46
Right inferior temporal	37	48, −37, −15	16	5.35
Right lateral occipital–temporal	37	44, −54, −16	106	5.21
Right medial occipital–temporal	20	40, −22, −26	28	4.47
Left calcarine	17	0, −84, −6	432	4.31

Texture control experiment procedure

The materials and procedure in the behavioral control experiment were identical to those in the fMRI experiment with the following exceptions. First, participants spoke aloud an acceptable name for each briefly presented object during phases 2 and 4 (to avoid motion artifacts, speaking aloud did not occur in the fMRI experiment). Second, no fixation-plus-blank control trials were used (these were needed in fMRI only to help achieve optimal deconvolution of the hemodynamic responses). Most importantly, all of the stimuli presented in the third phase were unfamiliar and non-nameable visual texture patterns (instead of familiar visual objects). Because these patterns were unfamiliar, they should not have been identifiable and should not have instigated relearning changes to visual object representations that are hypothesized to cause antipriming. Because these patterns were not identifiable, none of them were presented again in the fourth phase for (familiar) visual object identification; thus no priming effects could be measured in this experiment. To test whether antipriming occurred, identification accuracy and response times for correctly identified objects (using a voice-activated key) were the dependent variables in separate analyses of variance. We were able to use name responses in this experiment, thus the independent variables were test presentation condition (baseline vs. fourth phase), frequency of the names of the objects (high vs. low), and frequency of encountering the objects in everyday life (high vs. low). The test presentation condition main effect was used to test whether identification differed between baseline and phase 4, with an alpha value of .05.

Neurocomputational models

Neurocomputational models were used to provide a concrete specification of the theory of priming and antipriming in visual object identification systems, to simulate behavioral expressions of priming and antipriming effects, and to examine potential mechanistic predictors for the pattern of fMRI results obtained in visual object identification areas. These models used the Local, Error-driven and Associative, Biologically Realistic Algorithm (LEABRA++; O'Reilly, 2001; O'Reilly and Munakata, 2000), which is biologically inspired but simplified with respect to some of the neural details. Twenty networks (simulating 20 participants) were first trained to identify the same set of visual object images used in the fMRI experiment, using small changes in the weights on the connections to accomplish learning, and using partially overlapping internal representations of different object shapes. Then, each network was put through a simulated experiment following the same procedure as used in the fMRI experiment. Identification performance in the models was compared against human performance, and mechanistic variables in the models were compared against neural activity in the fMRI experiment.

The models had four layers of units (Fig. 2). The input layer was a two-dimensional 128×128 array on which bitmapped gray-scaled images of visual objects were centered to simulate visual presentations of objects in a retinotopic format.

The second layer had 1000 units fully interconnected with the first layer. The connections between layers 1 and 2 were hardwired to perform Gabor-jet transformations of the input images, capturing filter properties of brain area V1 (Lades et al., 1993). Using Gabor-type wavelets, each unit of layer 2 represented one combination of eight

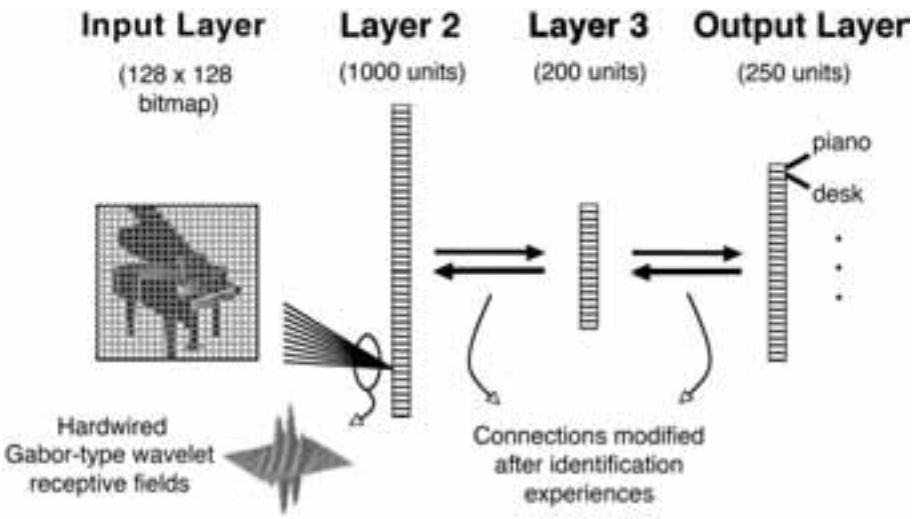

Fig. 2. Architecture of the neurocomputational models. Four-layer models were first trained to identify the 250 object images used in the fMRI experiment, and then the models were used in a simulation of the fMRI experiment. Weights on the connections between layers 2 and 3 and between layers 3 and 4 were modified during original training to enable identification of visual object images, and these weights were modified in the same way during the simulated experiment. The models exhibited both priming and antipriming effects in object identification performance. Perhaps more importantly, two variables in the models that are involved in changes of connection weights (simulating synaptic modifications) predicted the pattern of fMRI results (greater activation for antiprimed objects than for baseline or primed objects). A different variable in the models, magnitudes of unit activations (simulating neuronal spiking activity), did not predict the pattern of fMRI results.

orientations (π/8, π/4, 3π/8, π/2, 5π/8, 3π/4, 7π/8, π), five sine-wave-cycle scales (4, 6, 8, 12, and 16 pixels/cycle), and 25 tessellated locations (5×5) on the input array. The resulting layer 2 vectors were normalized within pattern and thresholded to 25%-activated binary vectors, enhancing LEABRA++ system performance.

The third layer had 200 units fully interconnected with layer 2, using sparse-distributed (partially overlapping) representations to process the 250 input object shapes. The fourth layer was the output layer, with 250 units fully interconnected with the units in the third layer, and the models were trained to activate one of these output units for each of the 250 input images. Layers 2 to 4 were trained to identify the input patterns using the three-layer LEABRA++ algorithm (O'Reilly, 2001; O'Reilly and Munakata, 2000), as summarized next.

In each trial, bidirectional activation flowed across weighted connections, enabling both bottom-up and top-down information to affect activations of units. Each trial involved iterations over minus and plus phases of settling. In the minus phase, the pattern of incoming activation on layer 2 was clamped, and in the plus phase, the pattern of incoming activation on layer 2 and the target pattern of output activation on layer 4 were clamped. In each cycle of settling, each non-clamped unit calculated an excitatory net input, a k winners-take-all (kWTA) inhibition over each layer, and then a point-neuron activation value that combined the excitatory input and the kWTA inhibition. After settling was completed in one trial, the final activation values for both minus and plus phases were recorded for each unit. These activation values were used to guide changes to the weights on all of the connections between layers 2 and 3 and between layers 3 and 4. Both error-driven learning weight changes (using both minus- and plus-phase activations to enable calculations of error) and Hebbian learning weight changes (from just plus-phase activations) were computed for each connection. The net weight change was a weighted sum of the error-driven and Hebbian changes.

All parameters were the default LEABRA++ parameters, except kWTA for layer 3 was set to 18 units (out of 200 units total) and the membrane potential, dt_{vm}, was set to 0.03. These exceptions enhanced original training according to pilot investigations. A base version of the networks that includes input (layer 2) patterns and target output (layer 4) patterns can be found at http://levels.psych.umn.edu/stimuli/A++.proj.gz, for use with the LEABRA++ (v3.0) software package available at http://grey.colorado.edu/emergent/index.php/PDP++PDP++.html.

Neurocomputational simulations

First, the 20 networks were trained to identify the 250 visual object images. Fifty epochs of 250 training trials each (one trial for each image per epoch) were used during original training. Trials within epochs were presented in a permuted random order. During original training, the number of settling cycles per trial was set at 70.

Then, each of the 20 networks was put through a simulated experimental session, mimicking the procedure used in the fMRI experiment. Simulating phase 2, the networks were presented with 100 input images for identification. To simulate identification of very briefly presented objects, the number of cycles to settle per trial during this phase was set to 27 (instead of 70, which was used during original training). Baseline object identification was measured in this simulation of phase 2. Because all objects were equally recently trained during the last original training epoch, little priming or antipriming can be measurable. Identification of an image was accurate when the most activated output unit (in layer 4) was the target output unit for that input.

Simulating phase 3, the networks followed the original training procedure to relearn or strengthen the representations for a set of 50 input images, none of which were presented in the preceding phase. During this phase, the number of cycles to settle per trial was set to 70 (as in original training).

Finally, simulating phase 4, the networks were presented with 100 object images for identification. As in the simulated phase 2, the number of cycles to settle per trial during this phase was set to 27 to simulate identification of very briefly presented objects. Fifty of the presented images were the same images whose representations were strengthened during the preceding phase; these were primed images. The other 50 were new in the simulated experiment (originally trained images that were not presented in the preceding phases of a simulated experiment); these were antiprimed images. Primed images and antiprimed images were presented in intermixed orders. As in the simulated phase 2, identification of an image was accurate when the most activated output unit (in layer 4) was the target output unit associated with that input. Identification rates were calculated for the primed and antiprimed trials, and these rates were compared against the baseline identification rates obtained in the simulated phase 2. Using two-tailed, paired *t*-tests, a simulation of behavioral priming was measured as significantly greater identification of

primed images compared against baseline, and a simulation of behavioral antipriming was measured as significantly poorer identification of antiprimed images compared against baseline (alpha value = .05).

Neurocomputational predictors for fMRI results

The networks were examined to test whether plausible aspects of the models predicted the pattern of fMRI results obtained in cortical visual object identification areas. Three main predictors were plausible a priori and hence investigated directly. Because fMRI analyses were restricted to trials in which participants gave identification responses, similarly the following analyses were restricted to correct identification trials.

The theory that the typical neural repetition effect (measured in our paradigm as antiprimed > primed) reflects increased signal for antiprimed objects (rather than decreased signal for primed objects) stems from considering that antiprimed objects have representations that have been recently weakened. Because of this weakening, antiprimed objects should elicit large magnitudes of error between target and actual activations of units. Thus, the squared error per unit was a metric investigated after every baseline, primed, and antiprimed trial during the simulated experiments. Another result of the weakening of representations of antiprimed objects should be that the magnitudes of weight changes on the connections between units should be large for antiprimed objects (in part because the error measure factors into the magnitudes of the weight changes). Thus, the magnitudes of weight changes on the connections between layers 2 and 3 (the connections feeding into the internal representations of objects) were recorded after every baseline, primed, and antiprimed trial during the simulated experiments. The results from both of these predictors are presented in Table 2, and they were compared against the obtained fMRI signal effects.

A different possibility stems from consideration that sharpening of unit activations or reduced firing rates may be responsible for the typical neural repetition effect as measured with fMRI (Desimone, 1996; Wiggs and Martin, 1998). Thus, summed output unit (layer 4) activations were recorded after every baseline, primed, and antiprimed trial during the simulated experiments (the activations of layer 3 units were forced to be sparse-distributed at 18 units activated by the kWTA algorithm, so they were not examined). The results from this potential predictor are presented in Table 2, and they also were compared against obtained fMRI effects.

Table 2
Successful and unsuccessful predictors of fMRI results from neurocomputational models.

Condition comparison	Mean (SEM)	Mean (SEM)	*t* value (*df*)	*p* value
Magnitude of weight changes per unit				
Antiprimed > primed	0.0005754 (0.0000080)	0.0005579 (0.0000071)	2.82 (19)	***.02***
Antiprimed > baseline	0.0005754 (0.0000080)	0.0005555 (0.0000051)	2.88 (19)	***.01***
Baseline > primed	0.0005555 (0.0000051)	0.0005579 (0.0000071)	−0.42 (19)	.68
Squared error per unit				
Antiprimed > primed	0.00399540 (0.00000047)	0.00398875 (0.00000212)	3.17 (19)	***.01***
Antiprimed > baseline	0.00399540 (0.00000047)	0.00399119 (0.00000153)	2.40 (19)	***.03***
Baseline > primed	0.00399119 (0.00000153)	0.00398875 (0.00000212)	0.86 (19)	.40
Summed unit activations				
Antiprimed > primed	0.00342 (0.00078)	0.02474 (0.01788)	−1.19 (19)	.25
Antiprimed > baseline	0.00342 (0.00078)	0.01723 (0.01165)	−1.18 (19)	.25
Baseline > primed	0.01723 (0.01165)	0.02474 (0.01788)	−0.34 (19)	.74

ERP experiment procedure

The materials and procedure in the ERP experiment were identical to those in the fMRI experiment with the following exception. In the object identification phases (phases 2 and 4), participants spoke aloud an acceptable name for the briefly presented object. (Speaking aloud was prohibited in the fMRI experiment to avoid motion artifacts.) Therefore, to examine behavioral effects, identification accuracy and response times for correctly identified objects were the dependent variables in separate analyses of variance. Because we were able to use name responses in this experiment, the independent variables were test presentation condition (baseline, primed, or antiprimed), frequency of encountering the objects in everyday life (high or low), frequency of the names of objects (high or low), and which of 10 counterbalanced assignments of objects to conditions was used for a participant. Simple effect contrasts were calculated to test for enhanced identification of primed objects relative to baseline and impaired identification of antiprimed objects relative to baseline, with an alpha value of .05.

The theory that motivates this study posits that the greater the visual similarity between two object images, the greater the overlap of their visual shape representations, and hence the greater that the strengthening of one representation should weaken the other representation. We took advantage of the large amount of behavioral data in the ERP experiment (from 40 participants) to test whether the visual similarities between object images predicted the magnitudes of antipriming effects. A correlation analysis was performed in an object-based manner. For each of the 250 objects used in the experiment, we first calculated the average visual similarity between that object and the 50 other objects that could have antiprimed it in the preceding phase when that object was used in the antiprimed condition. The similarity measure was the most basic and commonly used, involving the mean point-wise Euclidean distance between two images (Moses et al., 1994). The point-wise distance, d_{jk}, is calculated as:

$$d_{jk} = \frac{1}{n}\sqrt{\sum_{x=1}^{n}\left(I_j(x) - I_k(x)\right)^2}, \quad j,k = 1 \ldots p,$$

with n = number of pixels in an image, $I_j(x)$ = gray-scale value of the pixel in location x in the image I_j, and p = number of images. Because d_{jk} is a dissimilarity measure, the similarity measure (s_{jk}) between two images is:

$$s_{jk} = 1 - d_{jk}$$

Second, we also calculated the magnitude of antipriming observed for each test object (average poorer identification of that object when it was used in the antiprimed condition than when it was used in the baseline condition). A Pearson r was calculated to test for a significant correlation between the two measures, with an alpha value of .05.

Electrophysiological data were collected using two IBM-compatible computers and a 64-channel Neuroscan Synamps amplifier (Acquire versions 4.2 and 4.3). One computer used E-Prime to deliver stimuli, accept responses, and send event triggers to a second computer that used Neuroscan Acquire software for data acquisition with the attached Neuroscan amplifier. A 64-channel cap collected EEG activity from 54-scalp sites using sintered Ag–AgCl electrodes placed in accordance with the 10-20 International System (Jasper, 1958). EEG activity was recorded using CPZ as the reference, and then re-referenced offline to averaged mastoids. All EEG signals were digitized at 1000 Hz during data collection. In

addition to the scalp sites, electrodes were placed on each participant's face to measure electro-oculogram (EOG) data, which were used later in reduction of eye blink artifacts. Horizontal EOG was measured with electrodes placed on either canthus and vertical EOG was measured with electrodes placed sub- and supra-orbital ridge of the left eye.

ERP analyses

ERP epochs, 1000 ms pre-object onset and 2000 ms post-object onset, were locked to the onset of object stimuli during phases 2 and 4. Eye movement artifacts were reduced using an algorithm (Semlitsch et al., 1986), as implemented in the Neuroscan Acquire version 4.2 software. All impedances were kept below 8 kΩ. Epochs were baseline corrected for the 150 ms preceding the beginning of the trial (−650 to −500 before object onset). Within each trial, individual electrodes in which activity exceeded ±100 μV were omitted from analysis. Additionally, only correct identification trials were considered for ERP data analysis. After filtering and removing artifacts, data were averaged across trials and within test presentation conditions: primed, antiprimed, and baseline.

To examine an area similar to the most important regions measured in the fMRI experiment (left lateral occipital–temporal and left ventral occipital–temporal), the five electrodes that were closest to the average of the fMRI coordinates (−44, −79, 6 and −38, −56, −20) were selected for the analysis of ERP waveforms (TP7, CP3, P7, P5, and P3; http://wwwneuro03.uni-muenster.de/ger/t2tconv/). A sliding window analysis was performed to assess condition differences in mean ERP amplitude for every 100-ms time window post stimulus. The within-participants independent variables in an analysis of variance were test presentation condition (primed, antiprimed, or baseline) and time window (0–100, 100–200, …, 1500–1600 ms post stimulus). To examine a significant interaction between test presentation condition and time window, two-tailed, paired *t*-tests were used to test condition differences at each time window, with an alpha value of .05.

Results

fMRI experiment results

Behavioral effects of both enhanced and impaired object identification were observed in the fMRI experiment (Fig. 1B). Primed objects (80.0%) were identified significantly more often than baseline objects (59.1%) [$t(19) = 7.95$, $p < .0001$], replicating the benefit long known as priming (Schacter, 1987; Tulving and Schacter, 1990). More importantly, antiprimed objects (55.0%) were identified significantly less often than baseline objects (59.1%) [$t(19) = 2.20$, $p < .05$], reflecting a new impairment that we label antipriming (Marsolek et al., 2006).

It is important to note that some degree of impaired identification of baseline objects may have occurred due to real-world object identification that took place before the experimental session began. But if so, such reduced baseline performance would have worked against the finding of significantly poorer identification of antiprimed objects relative to the baseline measure within the experiment, and yet this significant difference was observed. Any uncontrollable pre-experimental changes did not prevent the finding of impaired object identification within the experiment.

During object identification, we measured blood-oxygen-level dependent (BOLD) signal using fMRI as a proxy for neural processing. To enable fMRI scanning with relatively high spatial resolution (2 × 2 × 2.5 mm) in visual shape processing regions, we oriented the MRI slice prescriptions on the ventral visual pathway (Fig. 3A). We first identified 10 regions of interest (ROIs) involved in visual object identification (Table 1). These were regions in which baseline, primed, and antiprimed objects elicited greater activation than blank display trials. The very frequently observed neural repetition effect (for reviews, see Grill-Spector et al., 2006; Henson, 2003) typically is measured as less activity for repeated stimuli (primed objects) than for intermixed non-repeated stimuli (antiprimed objects in our paradigm). Thus, we examined whether this difference was found in the 10 visual object identification ROIs. (Rare exceptions to the general finding of less activity for repeated stimuli have been found, but only

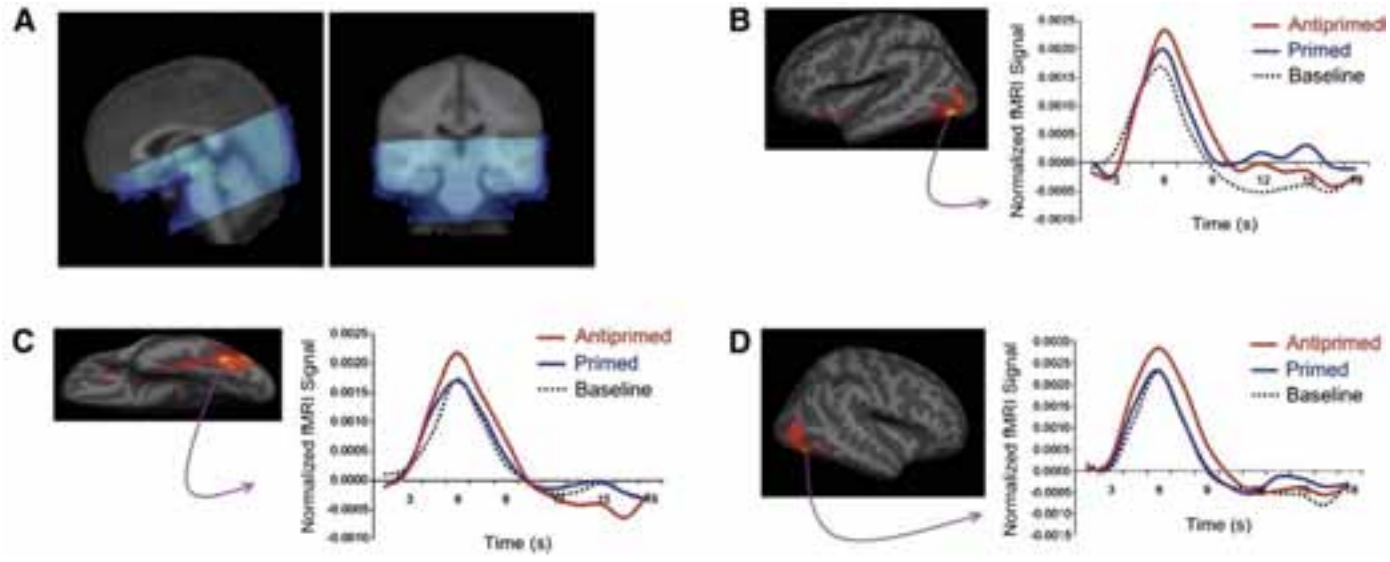

Fig. 3. fMRI slice prescriptions and fMRI results in regions that exhibited the neural repetition effect. (A) The locations of MRI slice prescriptions are shown for a representative participant. They were focused on visual object processing areas in the ventral visual stream. (B) The anatomical location and neural activity exhibited in left lateral occipital–temporal cortex, one of the three visual object identification ROIs that exhibited the typical neural repetition effect (unprimed > primed, which in our paradigm is antiprimed > primed). Neural activity was significantly greater in the antiprimed condition than in the baseline condition or in the primed condition (and the latter conditions did not differ). This pattern of results indicates that the typical neural repetition effect was due to increased activity in the antiprimed condition relative to baseline, rather than decreased activity in the primed condition relative to baseline. (C) The anatomical location and neural activity exhibited in left ventral temporal cortex, another of the three visual object identification ROIs that exhibited the typical neural repetition effect. (D) The anatomical location and neural activity exhibited in right inferior occipital cortex, the last the three visual object identification ROIs that exhibited the typical neural repetition effect.

when the stimuli are pre-experimentally novel [Henson et al., 2000; Schacter et al., 1995], not attended to during the initial presentation [Eger et al., 2004; Yi and Chun, 2005], or visually degraded [Rainer et al., 2004; Turk-Browne et al., 2007].) We observed significantly less activity for the primed objects than for the antiprimed objects in three of the 10 ROIs, consistent with previous priming research (for reviews, see Grill-Spector et al., 2006; Henson, 2003): left lateral occipital–temporal cortex (−44, −79, 6; BA37; Fig. 3B) [$t(19)=2.31$, $p<.03$], left ventral occipital–temporal cortex (−38, −56, −20; BA19; Fig. 3C) [$t(19)=2.50$, $p<.03$], and right inferior occipital cortex (48, −72, −12; BA19; Fig. 3D) [$t(19)=3.24$, $p<.005$].

Most importantly, we investigated the cause of the neural repetition effect. Extant theories ubiquitously interpret it as *decreased* neural activity for the repeated stimuli, as may occur due to adaptation or habituation (Grill-Spector and Malach, 2001), suppression or sharpening of activity (Desimone, 1996; Wiggs and Martin, 1998), or faster activations (Henson and Rugg, 2003; James and Gauthier, 2006) for the repeated stimuli. In contrast, we hypothesized that it may be due to *increased* neural activity for the non-repeated, antiprimed stimuli, as required to reestablish their previously impaired representations. Activity in the baseline condition supported our hypothesis. Antiprimed activation was greater than baseline and primed activation, which did not differ.

In particular, in left lateral occipital–temporal cortex (Fig. 3B), the typical neural repetition effect was due to increased activity for the antiprimed objects, not decreased activity for the primed objects. Activity elicited by antiprimed objects was significantly greater than activity elicited in the baseline condition [$t(19)=2.28$, $p<.04$], which in turn did not differ significantly from activity elicited by primed objects [$t<1$]. The typical neural repetition effect could not have been due to changes in representations of the primed objects, as assumed in extant theories (Desimone, 1996; Grill-Spector et al., 2006; Grill-Spector and Malach, 2001; Henson and Rugg, 2003; James and Gauthier, 2006; Wiggs and Martin, 1998). Instead, it was due to increased activity elicited by the antiprimed objects. The same pattern of results was found in the other two regions exhibiting the typical neural repetition effect (Figs. 3C and D). The increased activity from antiprimed objects relative to baseline was greater in magnitude than any difference in activity between primed objects and baseline, but the differences between antiprimed and baseline conditions did not achieve significance in those areas taken alone [p values>.05]. Although this effect is not large in every region, it is consistent across all the areas exhibiting the typical repetition effect. Thus, in order to directly assess the reliability of the antipriming effect across the different repetition effect areas, we combined the results from the three ROIs, using meta-analytic procedures. Following recommendations for such analyses (Rosenthal, 1978), we used three different methods to combine the results because the number of effects is small (three). In all three methods, the effect was reliable: for the method of adding logs (i.e., Fisher method) [$\chi^2(6)=16.1$, $p=.013$]; for the method of adding probabilities [$p=.004$]; and for the method of adding t values [$Z=2.48$, $p=.007$]. The finding of increased activation from antiprimed objects was reliable across areas. For a final confirmation, we conducted a repeated-measures analysis of variance that included data from all three regions. Collapsing across the three regions, activity from antiprimed objects was significantly greater than activity from primed objects and from baseline objects, which did not differ [$F(1,38)=4.34$, $p<.05$, for the appropriate interaction contrast].

One could suggest that the reason why activity in the antiprimed condition was greater than activity in the baseline condition was that there was a difference in overall level of fatigue or engagement in the task, or some other state difference, between the fourth and second phases of the experiment. If so, however, in the seven ROIs that did not exhibit the typical neural repetition effect (greater activity for antiprimed objects than primed objects), there should have been differences in activation between the antiprimed and primed conditions (measured in the fourth phase) and the baseline condition (measured in the second phase). Yet, there were no such differences [all p values>.05].

Texture control experiment results

A behavioral control experiment was conducted to test a similar alternative explanation for the behavioral result that antiprimed objects were identified more poorly than baseline objects. One could suggest that a state difference (e.g., fatigue) between the second and fourth phases of the fMRI experiment was responsible. If so, however, when only non-identifiable and non-nameable texture patterns are presented in place of familiar visual objects during the third phase of a control experiment, identification of (no longer antiprimed) objects should be poorer than identification of baseline objects (Fig. 4A). This was tested in the texture control experiment. Most importantly, no "antipriming" from viewing and judging texture patterns during the third phase was observed in object identification during the fourth phase (Fig. 4B). Similar to another experiment in which visual objects were not presented during the third phase (Marsolek et al., 2006), object identification during the fourth phase was not significantly less accurate (69.1%) and was not significantly slower (1321 ms) than baseline object identification during the second phase (69.5%; 1315 ms) [$F(1,39)<1$, for accuracy, and $F(1,39)<1$, for response times]. In addition, when the behavioral results from the fMRI experiment and the control experiment were combined in an analysis of variance, a significant interaction indicated that the poorer identification of antiprimed objects than baseline objects in the object-antipriming (fMRI) experiment was greater than the difference between "antiprimed" objects and baseline objects in the texture-antipriming (control) experiment [$F(1,58)=6.96$, $p<.02$]. These results are important for indicating that the significantly impaired identification of antiprimed objects in the fMRI experiment was not due to fatigue or some other state difference between baseline and antiprimed identification trials.

Neurocomputational modeling results

The finding of increased activation for antiprimed objects is counterintuitive according to extant theories; thus, to provide a clear mechanistic hypothesis for its origin, we examined biologically plausible models trained to identify our object images. These were LEABRA++ neurocomputational models (O'Reilly, 2001; O'Reilly and Munakata, 2000) that first were trained to identify gray-scaled bitmap images of the objects used in the fMRI experiment (Fig. 2). These models utilized partially overlapping representations of the different object shapes to learn to identify them; because relearning of objects takes place after every identification experience, models like these readily simulate visual object priming (Stark and McClelland, 2000) and antipriming effects (Marsolek et al., 2006). After initial object identification training, the models were used in simulated experiments following the same procedure as used in the fMRI experiment. The models successfully simulated behavioral effects of enhanced and impaired object identification. Primed objects (81.1%) were identified significantly more often than baseline objects (72.9%) [$t(19)=4.02$, $p<.001$], and antiprimed objects (52.0%) were identified significantly less often than baseline objects (72.9%) [$t(19)=7.97$, $p<.001$].

More importantly, the aspects of the models that predicted the pattern of fMRI activity (greater activation for antiprimed objects than for baseline or primed objects) were variables involved in changes of connection weights (simulating synaptic modifications), not magnitudes of unit activations (simulating neuronal spiking activity). The magnitudes of weight changes on connections between the second and third layers of units were significantly greater in the antiprimed condition than in the baseline condition [$t(19)=2.88$, $p<.01$], significantly greater in the antiprimed

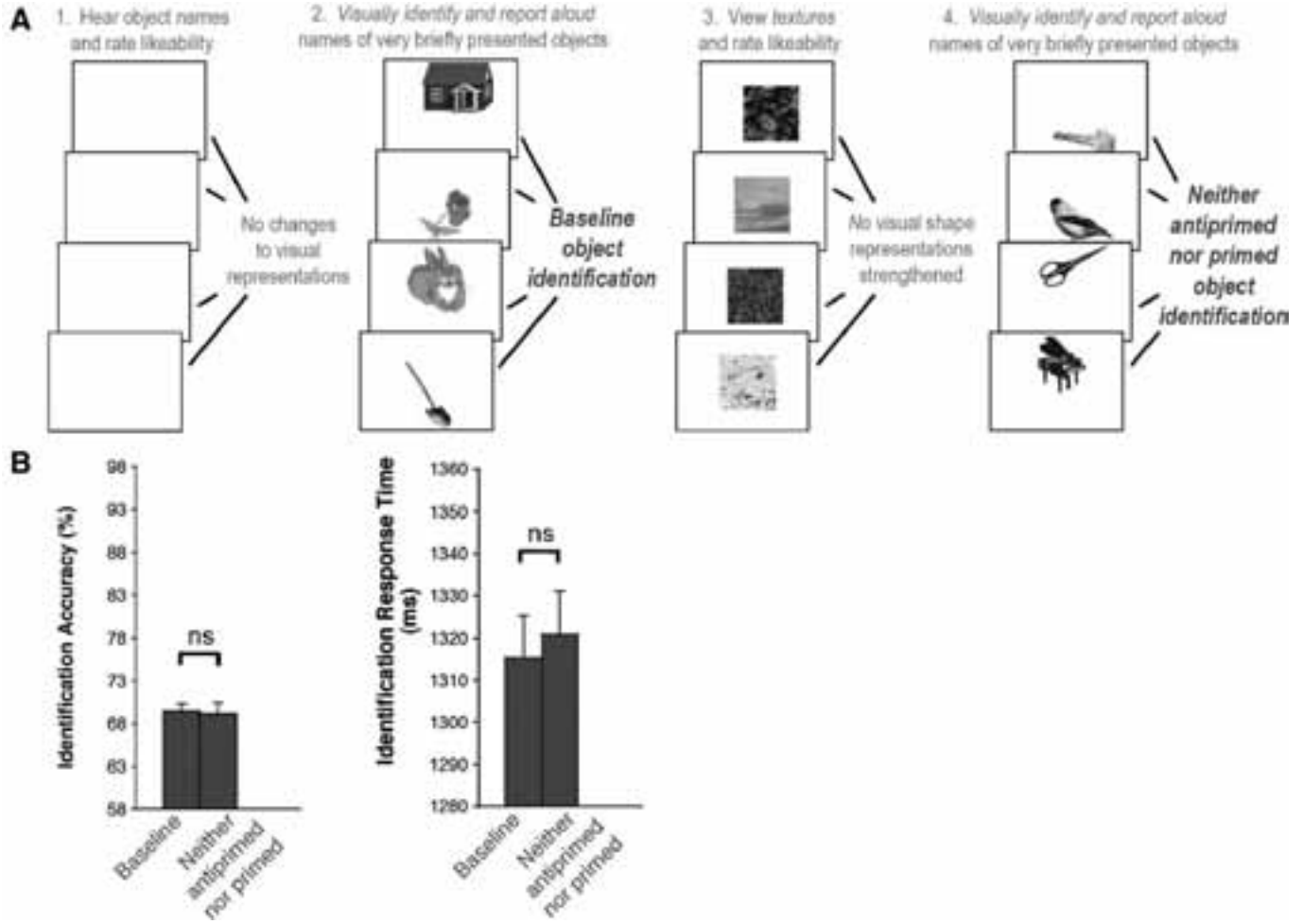

Fig. 4. Texture control experiment design and behavioral results. (A) The experiment had four phases, and example visual stimuli are shown for each phase. First, like in the fMRI experiment (Fig. 1A), participants stared at a blank visual display while they heard the names of 50 objects, each of which they rated for likeability. Second, participants visually identified and verbally reported 100 very briefly presented pictures of familiar objects that were neither primed nor antiprimed. Baseline object identification was measured during these trials, because they occurred after a period of time in which no changes in visual representations of objects should have taken place (during the first phase). Third, unlike in the fMRI experiment, participants viewed and rated the likeability of 50 texture patterns that were unfamiliar and non-nameable. Because the patterns were unfamiliar, they should not have been identifiable and they should not have instigated the relearning changes to visual object representations that are hypothesized to cause antipriming. Fourth, participants visually identified and verbally reported 50 very briefly presented objects, none of which were presented in any form previously in the experiment. Because no changes to visual object representations should have taken place in the preceding phase, the objects presented in the fourth phase were neither antiprimed nor primed. (B) The behavioral results from the texture control experiment are depicted. Mean identification accuracy (+ standard errors of the mean) and mean response times for correctly identified objects (+ standard errors of the mean) are plotted as a function of test presentation condition (baseline vs. neither antiprimed nor primed). The differences between identification of baseline objects and objects that have neither been antiprimed nor primed did not approach significance (ns = non-significant). Combined with the behavioral results from the fMRI experiment, these data indicate that strengthening of visual object representations during the third phase is required for the antipriming effect and that greater fatigue (or some other state difference) does not cause impaired identification of objects presented in the fourth phase compared with the second phase.

condition than in the primed condition [$t(19) = 2.82$, $p < .02$], and not different between the baseline and primed conditions [$t < 1$] (Table 2). The same pattern of results was observed for magnitudes of error between actual output unit activations and target output unit activations [$t(19) = 2.40$, $p < .03$, $t(19) = 3.17$, $p < .01$, and $t < 1$, respectively] (Table 2), which factor into the calculation of error-correction weight changes. These findings indicate that antiprimed objects engaged a greater degree of maintenance relearning than did primed or baseline objects, in a pattern that predicted the fMRI results. However, not all aspects of the models predicted the pattern of fMRI results. In particular, because it has been suggested that differences in neural spiking activity elicited by repeated and non-repeated stimuli may underlie the neuroimaging signal differences (Desimone, 1996; Wiggs and Martin, 1998), we examined the average magnitudes of output layer unit activations (layer 3 units were forced to be sparse-distributed at 18 units activated, so they were not examined). These results did not predict the fMRI results, as no significant differences were observed between antiprimed, baseline, and primed conditions [for antiprimed vs. baseline, $t(19) = 1.18$, $p > .25$; for antiprimed vs. primed, $t(19) = 1.19$, $p > .25$; and for primed vs. baseline, $t < 1$] (Table 2). The results from neurocomputational models support our interpretation that the differences in neural activity elicited by antiprimed and primed items reflect increased neural processing required for maintenance relearning of the antiprimed objects.

Therefore, the models provide a concrete specification of a system that (a) successfully simulates the priming and antipriming effects in object identification performance and (b) provides an explanation for the pattern of fMRI activation results that we observed. Indeed, the models predicted both behavioral and fMRI results, despite the fact that these were *different patterns* of results. The different patterns of behavioral and fMRI results may not be surprising given that behavioral repetition effects and fMRI repetition effects in posterior cortex (but not in prefrontal regions) usually are not correlated (Schacter et al., 2007).

ERP experiment results

Finally, we used electroencephalography, a technique with high temporal resolution, in a parallel experiment to test whether the pattern of fMRI results—greater signal elicited by antiprimed objects than by primed or baseline objects—may reflect maintenance relearning of the antiprimed objects. If so, because synaptic modification processes take place after object identification, the pattern of fMRI results should be observable via scalp-recorded brain electrical activity relatively late in the object identification trials.

In the ERP experiment (Fig. 5A), significant behavioral effects of priming and antipriming were observed in object identification performance (Fig. 5B). Primed objects were identified more accurately (91.2%) and faster (832 ms) than baseline (78.9%; 858 ms) [$F(1,60) = 105.74$, $p < .001$, for accuracy, and $F(1,60) = 9.52$, $p < .01$, for response times]. In addition, antiprimed objects were identified less accurately (76.0%) and slower (875 ms) than baseline (78.9%; 858 ms) [$F(1,60) = 5.67$, $p < .03$, for accuracy, and $F(1,60) = 4.05$, $p < .05$, for response times].

Moreover, there were enough participants in this experiment to enable a test of whether the visual similarities between different object images were important for predicting the magnitudes of antipriming effects (the magnitudes of less accurate identification of antiprimed objects than baseline objects). According to our theory, the greater the visual similarity between two object images, the more likely their visual shape representations are superimposed, and hence the greater the likelihood that strengthening one representation will weaken the other representation. Indeed, over the 250 objects used in the experiment, the degrees of visual shape similarity between each of them and the 50 other visual objects that could have antiprimed them in the preceding phase (when they were used in the antiprimed condition) significantly predicted the magnitudes of antipriming observed for the objects (poorer identification when they were used in the antiprimed condition than when they were used in the baseline condition) [$r(248) = .22$; $p < .001$].

Supporting the prediction that the fMRI results reflect maintenance relearning of antiprimed objects, the ERPs elicited by antiprimed objects showed greater positivity than ERPs elicited by baseline and primed objects (which were equivalent), and this occurred only in late time windows, after the overall mean identification response time of 855 ms (Fig. 6). The results in early time windows were different, in a pattern presumably reflecting facilitated access to perceptual representations for primed items (Grill-Spector et al., 2006; Rugg, 1995; Schendan and Kutas, 2003) or declarative memory for the primed items (Grill-Spector et al., 2006; Voss and Paller, 2008).

In particular, in the analysis of mean ERP amplitudes measured from the five electrodes closest to the left occipital–temporal fMRI areas that exhibited the neural repetition effect (Fig. 6), the interaction between test presentation condition (baseline vs. primed

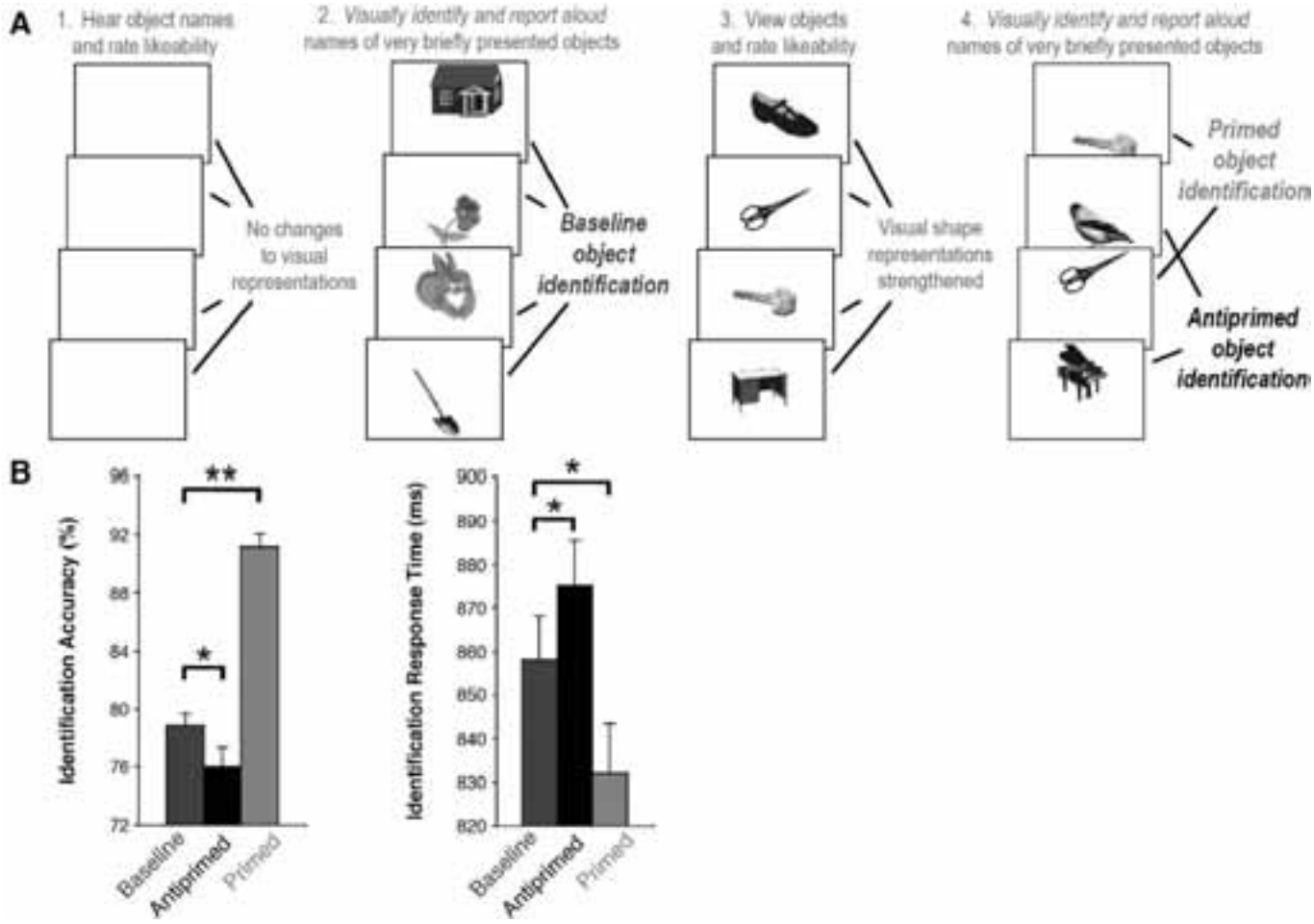

Fig. 5. ERP experiment design and behavioral results. (A) The only difference in procedure between the ERP experiment and the fMRI experiment (Fig. 1A) was that participants visually identified and spoke aloud an acceptable name for each briefly presented object presented in the second and fourth phases (speaking aloud was prohibited in the fMRI experiment to avoid motion artifacts). (B) The behavioral results from the ERP experiment are depicted. Mean identification accuracy (+ standard errors of the mean) and mean response times for correctly identified objects (+ standard errors of the mean) are plotted as a function of test presentation condition (baseline, antiprimed, and primed). Primed objects were identified significantly more accurately and significantly faster than baseline objects, and antiprimed objects were identified significantly less accurately and significantly slower than baseline objects, thus both priming and antipriming effects were observed. *$p < .05$. **$p < .001$.

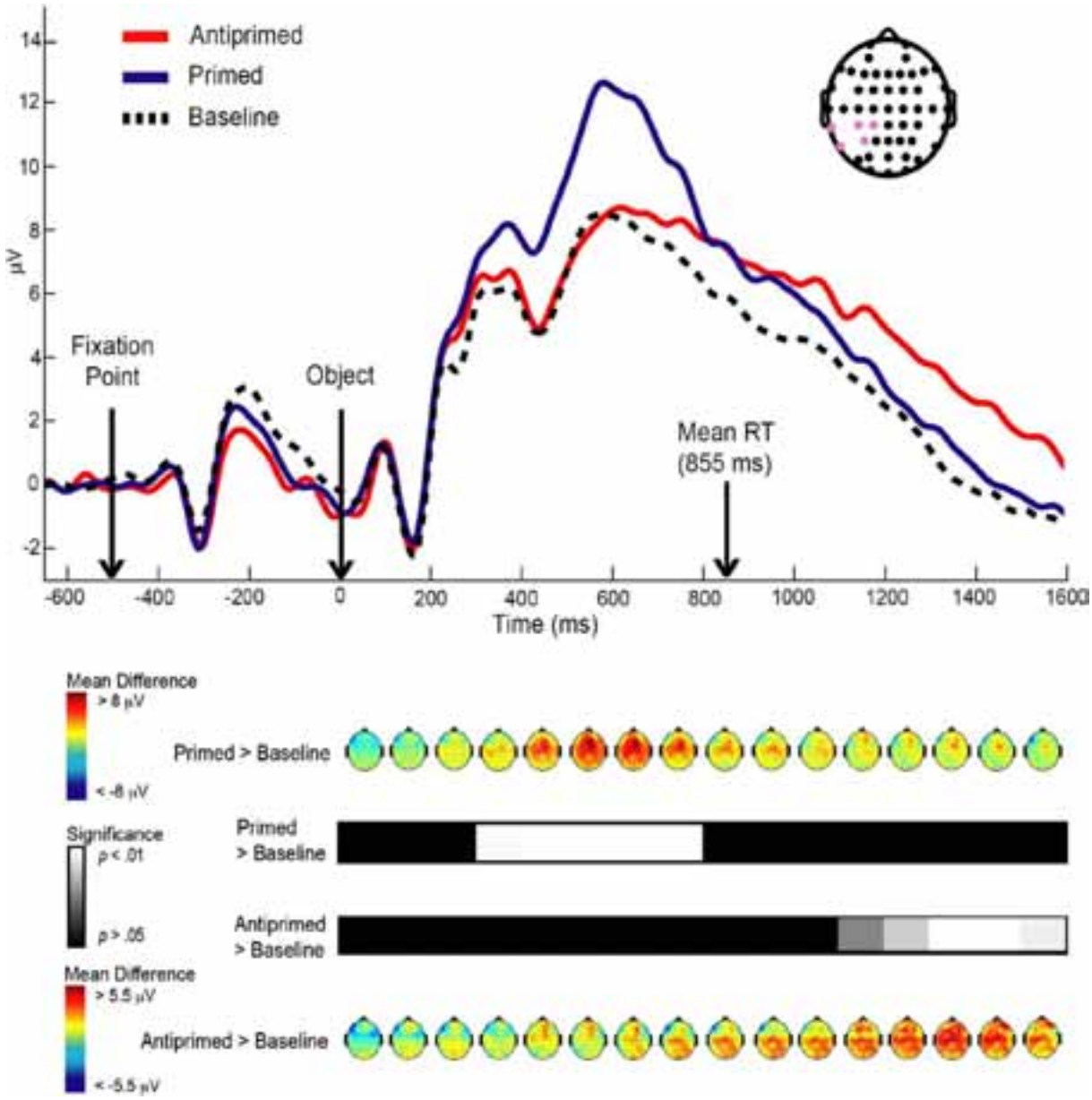

Fig. 6. ERP results. The ERP waveforms reflect activity from the five electrodes closest to the left occipital–temporal fMRI areas that exhibited the neural repetition effect (locations depicted in color in the inset), for the baseline, primed, and antiprimed conditions. Levels of gray reflect the significance of primed > baseline and antiprimed > baseline tests of those waveforms for every 100-ms time window post-object. The scalp topographies reflect primed > baseline and antiprimed > baseline condition differences for all electrodes for each time window. These results indicate that the pattern of fMRI results (greater activation for antiprimed objects than for baseline or primed objects) occurred only relatively late after object presentations (after the overall mean identification response time of 855 ms). This supports the theory that the fMRI results reflect relearning of previously weakened representations of (antiprimed) objects.

vs. antiprimed) and time window was significant [$F(30,1170) = 6.81$, $p < .001$]. The *t*-tests used to examine this interaction are reported in Fig. 6). Most importantly, the pattern of results observed in fMRI (greater signal elicited by antiprimed objects than by primed or baseline objects) was observed with ERPs only relatively late following object presentations (only in post-1100 ms time windows). This supports the hypothesis that enhanced fMRI activation for antiprimed objects reflects maintenance relearning of those objects.

Discussion

We have shown that identification of visual objects can impair subsequent identification of other visual objects, as predictable from important aspects of the neural implementation of vision. The neural representations of whole object shapes are partially overlapping in the ventral visual stream. Neurons in different cortical columns are maximally sensitive to different moderately complex features of whole objects. These features are simple enough that they appear in multiple whole objects but also complex enough that they do not appear in nearly all whole objects (Tanaka, 1993, 2003). Similarly, neuroimaging studies indicate that patterns of activation elicited by different objects (e.g., houses, bottles, shoes, faces, chairs, etc.) are overlapping in human ventral temporal cortex (Haxby et al., 2001; Ishai et al., 1999). This representational scheme likely evolved for increased efficiency of storage of many items and for enhanced generalization abilities (McClelland et al., 1995; McClelland and Rumelhart, 1985). Such neural representations undergo learning-related changes after identification experiences (Kobatake et al., 1998; Sigala and Logothetis, 2002) on a continuous basis (i.e., learning/relearning is never turned

off), likely via synaptic modifications as in long-term potentiation and long-term depression (Chen et al., 1996).

A previously unappreciated implication of these properties of the visual system is that the continual tweaking of overlapping representations can lead to both subsequent impairments and subsequent enhancements in identifying objects. The results reported here provide the first evidence in support of our theory that continual adjustments of overlapping object representations lead to impaired object identification (antipriming) in addition to enhanced object identification (priming), and because of the impairments, increases in neural activity are elicited by antiprimed objects in order to reestablish their weakened representations.

The present results are crucial for reassessing one of the most frequently observed effects in neuroimaging, namely that neural activity is lower for repeated than for non-repeated stimuli (for reviews, see Grill-Spector et al., 2006; Henson, 2003). Extant explanations are that it reflects a decrease in neural activity elicited by the repeated stimuli, due to adaptation or habituation (Grill-Spector and Malach, 2001), suppression or sharpening of activity (Desimone, 1996; Wiggs and Martin, 1998), or faster activations (Henson and Rugg, 2003; James and Gauthier, 2006) for the repeated stimuli. Our results indicate that it reflects *increased* activity (relative to an unbiased baseline) due to the increased need to strengthen or reestablish the recently weakened representations of objects in the non-repeated, antiprimed condition. Moreover, the results from neurally plausible models indicate that successful predictors of the neural repetition effect involve greater magnitudes of synaptic changes for the recently weakened representations, not changes in neural firing for the strengthened representations. Finally, only late time window ERP results reflect the pattern of fMRI results, in support of the hypothesis that post-identification, maintenance relearning processes underlie the fMRI results.

These conclusions are provocative because they run counter to the extant interpretations of the neural repetition effect in perceptual cortical areas. The field may need to make such a fundamental shift in interpretation, according to accumulating neural evidence. The typical interpretations of fMRI repetition effects cannot account for why behavioral repetition effects and fMRI repetition effects in posterior (but not prefrontal) cortex are not correlated (Schacter et al., 2007), nor for why behavioral repetition effects and firing-rate repetition suppressions in inferior temporal cortex are not correlated (McMahon and Olson, 2007).

Our conclusions also run counter to a common assumption that BOLD fMRI signals are related to neuronal firing rates. An intriguing aspect of our theory is that it is more in line with findings that BOLD fMRI signals have been associated with local field potentials from dendritic activity, and not with spiking activity, at sites exhibiting transient responses (Logothetis et al., 2001). Because one could argue whether dendritic activity (local field potentials) reflects synaptic changes or spiking activity in upstream regions, it is crucial to note that more recent work indicates a strong coupling between local field potentials and changes in tissue oxygen concentrations even in the absence of spikes (Viswanathan and Freeman, 2007). These findings give us sure enough footing to hypothesize that our BOLD results are due to synaptic changes. This footing is at least as sure as that associated with the hypothesis that BOLD is due to spiking activity. Thus, although our conclusion runs counter to the ubiquitous assumption that spiking activity underlies magnitudes of BOLD signal, our explanation is that the present fMRI results reflect synaptic activity used for maintenance relearning of object shape information.

The unbiased baseline condition is critical in our procedure for teasing apart the enhanced ability to identify repeated stimuli (relative to baseline) and the impaired ability to identify non-repeated stimuli (relative to baseline). Other studies of repetition effects do not utilize such a baseline condition in which performance is unaffected by recent changes to visual representations. Typically in other studies, new or "unprimed" stimuli are presented intermixed with repeated stimuli, and performance is compared between those two conditions. By our theory, the new/"unprimed" stimuli are actually antiprimed, and without a baseline measure, differences between primed performance and new/"unprimed" performance conflate priming benefits and antipriming costs. In addition, another way in which repetition effects can be investigated behaviorally (Biederman and Cooper, 1991; Buckner et al., 1998; Sayres and Grill-Spector, 2006) and neurally (Buckner et al., 1998; Grill-Spector and Malach, 2001; Sayres and Grill-Spector, 2006) is to measure differences between the initial presentation and the subsequent presentation(s) of the same items (or slightly transformed items). However, such comparisons require using the same task between initial and subsequent presentations. This enables rapid learning of stimulus–response associations to underlie repetition effects, a form of learning that has different properties from the more purely perceptual repetition effects of interest here (Dobbins et al., 2004; Horner and Henson, 2008; Schnyer et al., 2006).

One could suggest an alternative explanation for why identification was poorer in the antiprimed condition than in the baseline condition. The suggestion is that baseline objects could have been *primed* by recent experiences with them in the real world, before the experiment. It is important to note that the only way this suggestion could work is by accepting completely implausible assumptions about the participants' real world experiences during a very circumscribed time, namely about 15–30 minutes before the start of the baseline phase. Any real world experiences that could have primed the baseline objects would have had to occur about 15 minutes before the onset of the baseline trials (participants stared at the blank visual display during the first phase of the experiment and during anatomical scan acquisition for about 15 minutes). However, if priming of baseline objects can take place from real world experiences, then priming of antiprimed objects can similarly take place. Any real world object experiences that could have primed the antiprimed objects would have had to occur about 30 minutes before the onset of the antiprimed trials (the time between the onset of the experiment and the onset of phase 4 was about 30 minutes). Thus, any greater priming of baseline objects than priming of antiprimed objects from real world experiences would have had to be due to real world experiences with baseline objects that took place only between 15 and 30 minutes before the baseline trials. This is implausible. Stimulus objects in the experiment were pseudo-randomly assigned to the baseline, antiprimed, and primed conditions, in a counterbalanced manner across subjects. This renders it highly implausible that a participant's personal experiences in the real world occurring only 15 to 30 minutes before the baseline phase would correspond to specific objects that were presented in the baseline condition for that particular subject. That an explanation of this sort could account for consistent effects across the current fMRI experiment, the current ERP experiment, and in multiple previous behavioral experiments, is extremely unlikely.

Even more importantly, two of the key findings in the present study rule out the alternative suggestion. First, in line with our theory, the degrees of visual shape similarity between antiprimed test objects and the visual objects that could have antiprimed them in the preceding phase significantly predicted the magnitudes of antipriming for those test objects. This result should not have been observed if the difference between baseline and antiprimed identification was due to priming of baseline objects from real world experiences. Second, in the texture control experiment, no antipriming occurred when non-nameable texture patterns were presented instead of objects in the third phase of the experiment. These results should not have been observed if priming of the baseline objects was the reason why baseline and antiprimed identification differed in the other experiments.

One also could suggest an alternative explanation for why neural activity was greater in the antiprimed condition than in the baseline condition. The suggestion is that scanner drift (differential heating of the MRI receiver coils that changes the observed signal over the course of many trials) could account for the increased signal between the antiprimed trials (in phase 4) and the baseline trials (in phase 2). It is important to note why scanner drift could not account for our fMRI antipriming effect. First, baseline trials during the second phase and primed/antiprimed trials during the fourth phase took place in separate fMRI runs that were initiated at the beginning of phase 2 and at the beginning of phase 4. More importantly, the method we used for deconvolution of the fMRI signal "mean normalized" the signal, such that it was expressed as a percent signal change from the zero point of not doing a task (fixating on blank displays) *within* a run. This was done after high-pass filtering that removed low-frequency signal drift during a run (as might be observed from scanner heating). Comparing signals characterized in this way between runs is not a problem, with the sole implausible exception that participants did something very different when they were not doing a task (fixating on blank displays) between phases 2 and 4, as that is what was associated with the zero points for computing percent signal change. Finally, our overall pattern of results rules out the scanner drift hypothesis. If scanner drift was responsible for greater BOLD signal elicited by antiprimed items than by baseline items, then this should have been observed in all brain regions. In fact, seven of the ten object identification ROIs revealed no differences between baseline, antiprimed, and primed trials.

Overall, the present results exemplify the utility of integrating behavioral, neuroimaging, and neurocomputational modeling results to mutually constrain hypotheses and more effectively understand the neural implementation of perception and memory. Future work should be aimed at more directly testing the theory that neural representations commonly are weakened by the same changes that cause other representations to be strengthened. Relearning of weakened representations may be a crucial aspect of perception and memory that is not well appreciated but that is needed to maintain sets of superimposed representations in neocortical long-term memory.

Acknowledgments

We thank Chris Azorson, Steve Engel, and Yuhong Jiang for valuable comments on a previous version of the manuscript; Susan Park Anderson, Sarah Davis, Sarah Fitch, Luke Kane, Alvina Kittur, Rachel LaNasa, Mithra Sathishkumar, and Benjamin Spear for help with data collection and analysis; and Randy O'Reilly for help with the simulation models. The research was supported by the National Institute of Mental Health (D.M.S., M.V., and C.J.M.). It was also supported by the Medical Research Service of the Department of Veterans Affairs and the Center for Cognitive Sciences in conjunction with the National Institute of Health and Human Development and the Office of the Vice President for Research and Dean of the Graduate School of the University of Minnesota.

References

Biederman, I., Cooper, E.E., 1991. Evidence for complete translational and reflectional invariance in visual object priming. Perception 20, 585–593.

Buckner, R.L., Goodman, J., Burock, M., Rotte, M., Koutstaal, W., Schacter, D.L., Rosen, B., Dale, A.M., 1998. Functional–anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. Neuron 20, 285–296.

Chao, L.L., Weisberg, J., Martin, A., 2002. Experience-dependent modulation of category-related cortical activity. Cereb. Cortex 12, 545–551.

Chen, W.R., Lee, S., Kato, K., Spencer, D.D., Shepherd, G.M., Williamson, A., 1996. Long-term modifications of synaptic efficacy in the human inferior and middle temporal cortex. Proc. Natl. Acad. Sci. U. S. A. 93, 8011–8015.

Dehaene, S, Jobert, A., Naccache, L., Ciuciu, P., Poline, J.-B., Le Bihan, D., Cohen, L., 2004. Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence. Psychol. Sci. 15, 307–313.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., Riviére, D., 2001. Cerebral mechanisms of word masking and unconscious repetition priming. Nat. Neurosci. 4, 752–758.

Desimone, R., 1996. Neural mechanisms for visual memory and their role in attention. Proc. Natl. Acad. Sci. U. S. A. 93, 13494–13499.

Dobbins, I.G., Schnyer, D.M., Verfaellie, M., Schacter, D.L., 2004. Cortical activity reductions during repetition priming can result from rapid response learning. Nature 428, 316–319.

Eddy, M.D., Schnyer, D., Schmid, A., Holcomb, P.J., 2007. Spatial dynamics of masked picture repetition effects. NeuroImage 34, 1723–1732.

Eger, E., Henson, R.N.A., Driver, J., Dolan, R.J., 2004. BOLD repetition decreases in object-responsive ventral visual areas depend on spatial attention. J. Neurophysiol. 92, 1241–1247.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb. Cortex 14, 11–22.

Grill-Spector, K., Malach, R., 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol. 107, 293–321.

Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn. Sci. 10, 14–23.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293, 2425–2430.

Henson, R., Shallice, T., Dolan, R., 2000. Neuroimaging evidence for dissociable forms of repetition priming. Science 287, 1269–1272.

Henson, R.N.A., 2003. Neuroimaging studies of priming. Prog. Neurobiol. 70, 53–81.

Henson, R.N.A., Rugg, M.D., 2003. Neural response suppression, haemodynamic repetition effects, and behavioural priming. Neuropsychologia 41, 263–270.

Horner, A.J., Henson, R.N., 2008. Priming, response learning and repetition suppression. Neuropsychologia 46, 1979–1991.

Ishai, A., Ungerleider, L.G., Martin, A., Schouten, J.L., Haxby, J.V., 1999. Distributed representation of objects in the human ventral visual pathway. Proc. Natl. Acad. Sci. U. S. A. 96, 9379–9384.

James, T.W., Gauthier, I., 2006. Repetition-induced changes in BOLD response reflect. accumulation of neural activity. Hum. Brain Mapp. 27, 37–46.

Jasper, H.H., 1958. The ten-twenty electrode system of the International Federation. Electroen. Clin. Neuro. 10, 371–375.

Jolicœur, P., Gluck, M.A., Kosslyn, S.M., 1984. Pictures and names: making the connection. Cogn. Psychol. 16, 243–275.

Kobatake, E., Wang, G., Tanaka, K., 1998. Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. J. Neurophysiol. 80, 324–330.

Lades, M., Vorbruggen, J.C., Buhmann, J., Lange, J., von der Malsburg, C., Wurtz, R.P., Konen, W, 1993. Distortion invariant object recognition in the dynamic link architecture. IEEE T. Comput. 42, 300–311.

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neurophysiological investigation of the basis of the fMRI signal. Nature 412, 150–157.

Marsolek, C.J., 2008. What antipriming reveals about priming. Trends Cogn. Sci. 12, 176–181.

Marsolek, C.J., Schnyer, D.M., Deason, R.G., Ritchey, M., Verfaellie, M., 2006. Visual antipriming: evidence for ongoing adjustments of superimposed visual object representations. Cogn. Affect. Behav. Ne. 6, 163–174.

McClelland, J.L., Rumelhart, D.E., 1985. Distributed memory and the representation of general and specific information. J. Exp. Psychol. Gen. 114, 159–188.

McClelland, J.L., McNaughton, B.L., O'Reilly, R.C., 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. 102, 419–457.

McMahon, D.B.T., Olson, C.R., 2007. Repetition suppression in monkey inferotemporal cortex: Relation to behavioral priming. J. Neurophysiol. 97, 3532–3542.

Moses, Y., Adini, Y., Ullman, S., 1994. Face recognition: the problem of compensating for illumination changes. In: Eklundh, J.O. (Ed.), Proceedings of the European Conference on Computer Vision. Springer-Verlag, Stockholm, Sweden, pp. 286–296.

O'Reilly, R.C., 2001. Generalization in interactive networks: the benefits of inhibitory competition and Hebbian learning. Neural Comput. 13, 1199–1241.

O'Reilly, R.C., Munakata, Y., 2000. Computational Explorations in Cognitive Neuroscience. MIT Press, Cambridge, MA.

Rainer, G., Lee, H., Logothetis, N.K., 2004. The effect of learning on the function of monkey extrastriate visual cortex. PLoS Biol. 2, 275–283.

Rosenthal, R., 1978. Combining results of independent studies. Psychol. Bull. 85, 185–193.

Rugg, M.D., 1995. ERP studies of memory. In: Rugg, M.D., Coles, M.G.H. (Eds.), Electrophysiology of Mind: Event-Related Brain Potentials and Cognition. Oxford University Press, New York, pp. 132–170.

Sayres, R., Grill-Spector, K., 2006. Object-selective cortex exhibits performance-independent repetition suppression. J. Neurophysiol. 95, 995–1007.

Schacter, D.L., 1987. Implicit memory: history and current status. J. Exp. Psychol. Learn. 13, 501–518.

Schacter, D.L., Reiman, E., Uecker, A., Polster, M.R., Yun, L.S., Cooper, L.A., 1995. Brain regions associated with retrieval of structurally coherent visual information. Nature 376, 587–590.

Schacter, D.L., Wig, G.S., Stevens, W.D., 2007. Reductions in cortical activity during priming. Curr. Opin. Neurobiol. 17, 171–176.

Schendan, H.E., Kutas, M., 2003. Time course of processes and representations supporting visual object identification and memory. J. Cogn. Neurosci. 15, 111–135.

Schnyer, D.M., Dobbins, I.G., Nicholls, L., Schacter, D.L., Verfaellie, M., 2006. Rapid response learning in amnesia: delineating associative learning components in repetition priming. Neuropsychologia 44, 140–149.
Semlitsch, H.V., Anderer, P., Schuster, P., Presslich, O., 1986. A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. Psychophysiology 23, 695–703.
Sigala, N., Logothetis, N.K., 2002. Visual categorization shapes feature selectivity in the primate temporal cortex. Nature 415, 318–320.
Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23, 208–219.
Stark, C.E.L., McClelland, J.L., 2000. Repetition priming of words, pseudowords, and nonwords. J. Exp. Psychol. Learn. 26, 945–972.
Tanaka, K., 1993. Neuronal mechanisms of object recognition. Science 262, 685–688.
Tanaka, K., 2003. Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. Cereb. Cortex 13, 90–99.
Tulving, E., Schacter, D.L., 1990. Priming and human memory systems. Science 247, 301–305.
Turk-Browne, N.B., Yi, D.-J., Leber, A.B., Chun, M.M., 2007. Visual quality determines the direction of neural repetition effects. Cereb. Cortex 17, 425–433.
van Turennout, M., Ellmore, T., Martin, A., 2000. Long-lasting cortical plasticity in the object naming system. Nat. Neurosci. 3, 1329–1334.
van Turennout, M., Bielamowicz, L., Martin, A., 2003. Modulation of neural activity during object naming: effects of time and practice. Cereb. Cortex 13, 381–391.
Viswanathan, A., Freeman, R.D., 2007. Neurometabolic coupling in cerebral cortex reflects synaptic more than spiking activity. Nat. Neurosci. 10, 1308–1312.
Voss, J.L., Paller, K.A., 2008. Brain substrates of implicit and explicit memory: the importance of concurrently acquired neural signals of both memory types. Neuropsychologia 46, 3021–3029.
Wiggs, C.L., Martin, A., 1998. Properties and mechanisms of perceptual priming. Curr. Opin. Neurobiol. 8, 227–233.
Yi, D.-J., Chun, M.M., 2005. Attentional modulation of learning-related repetition attenuation effects in human parahipppocampal cortex. J. Neurosci. 25, 3593–3600.