Concordance between perceptual and categorical repetition effects in the ventral visual stream

Scott L. Fairhall, Stefano Anzellotti, Petra E. Pajtas, and Alfonso Caramazza. Concordance between perceptual and categorical repetition effects in the ventral visual stream. J Neurophysiol 106: 398–408, 2011. First published May 4, 2011; doi:10.1152/jn.01138.2010.—The process of object categorization is an integral part of human cognition. In the present study, we have used a repetition suppression paradigm to determine the degree to which the ventral visual cortex is sensitive to categorical relationships. By using images of animals and tools, suppression across perceptual (stimulus level) and categorical repetitions (basic level and domain level) was compared and contrasted across the domain-selective and hierarchical organization of the ventral visual stream. Both perceptual and categorical repetition effects were insensitive to domain-selective tuning, with suppression most prominent in regions responding maximally to images, irrespective of stimulus domain. Likewise, both perceptual and categorical repetition produced overlapping suppression across multiple regions of the visual hierarchy. Some divergent patterns were observed. The right superior temporal sulcus demonstrated repetition suppression only at the basic level (different examples of the same basic object), and the right anterior fusiform gyrus was sensitive to direct stimulus repetition but not basic-level categorical repetition. Because of the high concordance between the response profiles of perceptual and categorical repetition effects, we conclude they arise from a common cognitive mechanism.

Functional magnetic resonance imaging-adaptation; repetition suppression; functional magnetic resonance imaging; visual cortex; domain

HOW OBJECTS ARE REPRESENTED within the visual system has been investigated extensively. Studies have identified brain regions responsive to particular classes of objects such as tools, houses, faces, and animals (Chao and Martin 2000; Downing et al. 2001, 2006; Ishai et al. 2000; Kanwisher et al. 1997; Martin 2007). However, it remains unclear to what extent processing within these regions is influenced by the perception of visual form or by categorical information. A promising method to begin addressing this issue is repetition suppression (RS) or functional magnetic resonance imaging (fMRI)-adaptation (e.g., Grill-Spector and Malach 2001; Henson et al. 2000; Kourtzi and Kanwisher 2001). RS arises when the neural response to the second presentation of a stimulus is decreased compared with the response to the first presentation and may indicate neuronal selectivity to specific stimulus attributes (Grill-Spector 2006) or top-down influences on the perceptual process (Mur et al. 2010; Summerfield et al. 2008). Suppression effects over two stimulus presentations identify neural populations sensitive to repeated attributes, whereas recovery from RS will detect neural populations insensitive to those attributes. This differential sensitivity of RS across stimulus attributes makes it amenable to investigate the manner in which varying relationships between objects influence the function of the visual cortex.

It is possible to establish a loose taxonomy of the varying relationships that can exist between objects. At a broad level, relationships can be of either a perceptual or categorical nature. Relationships of a perceptual nature can be further subdivided into the direct perceptual event, the same object when viewed under the same conditions, with all associated sensory properties (stimulus level), as well as the same object when viewed under different perceptual conditions, varying, for instance, size or point of view (identity level). In contrast, object relationships can be determined by categorical membership (Rosch 1976). Two examples of categorical relationships are the basic-object level (2 different exemplars of the object “hammer,” hereafter referred to as basic level) and the domain level (2 different exemplars drawn from the domain, e.g., “animal”).

Perceptual repetition effects are relatively well characterized. Even identity-level repetition, where size or point-of-view changes produce dramatic alterations in visual input, induce RS across higher level regions of the visual system, demonstrating that the visual cortex is sensitive to more than merely consistent retinal input. These regions include the inferior occipital gyrus (IOG), posterior and anterior fusiform gyrus of the ventral stream, and regions of the intraparietal sulcus in the dorsal stream (Andresen et al. 2009; Grill-Spector et al. 1999; James et al. 2002; Konen and Kastner 2008; Mur et al. 2010; Pourtois et al. 2009; Vuilleumier et al. 2002). In addition, there is some evidence for posterior-anterior hierarchical organization in identity-level RS, akin to that observed in the response properties of neurons in the cortex of nonhuman primates. Neurons in the monkey cortex represent simple stimulus features in early visual cortex (i.e., V4) but represent complex conjunctions of features in anterior inferior temporal cortex (Kobatake and Tanaka 1994; Kobatake et al. 1998). A similar pattern is present in humans: moving anteriorly from V4 along IOG to the posterior fusiform gyrus, object representations progress from low-level perceptual information to more abstract representations in more anterior regions (Lerner et al. 2001). Identity-level perceptual repetition effects have been reported to follow this pattern with suppression arising across variations in stimulus size as early as the IOG (Grill-Spector et al. 1999), whereas persistent RS across higher level manipulations of viewpoint does not manifest before the left anterior fusiform gyrus (Andresen et al. 2009; Vuilleumier et al. 2002).
However, recent studies suggest that viewpoint-generalized RS (Mur et al. 2010; Pourtois et al. 2009) may also be present earlier in the visual processing hierarchy, challenging the notion of a rigid posterior-anterior organization (see also Eger et al. 2008a, 2008b).

Findings of categorical repetition effects within the visual system have been less consistent. Thus far, only basic-level (e.g., 2 separate hammers) repetition effects have been investigated. Although some studies have shown basic-level RS (Koutstaal et al. 2001; Simons et al. 2003), others have not (Chouinard et al. 2008; Vuilleumier et al. 2002). Recently, Chouinard et al. (2008) proposed that the reason for previous findings of basic-level RS was the conflation of perceptual with categorical factors, due to the increased image-based visual similarity between same-object pairs compared with different-object control pairs. Notwithstanding this potential confound, studies that did report persistent RS across basic-level repetitions provide tentative evidence for a posterior-anterior hierarchical organization. As was the case with repetition across viewpoint (Vuilleumier et al. 2002), basic-level RS effects were present in the left anterior fusiform gyrus and not in earlier visual regions. These results have been taken as evidence that the visual cortex, specifically the left fusiform gyrus, is sensitive to semantic/conceptual relationships, whereas other visual regions are more involved in processes specific to visual form (Koutstaal et al. 2001; Simons et al. 2003).

Embedded within the posterior-anterior hierarchical structure of the visual system is a second prominent aspect of visual organization, domain selectivity. At the same time that all regions of inferior occipital and fusiform gyri respond to all objects (Haxby et al. 2001; Ishai et al. 1999), they exhibit anatomically distinct patterns of preference for a variety of object domains (Chao and Martin 2000; Downing et al. 2001, 2006; Ishai et al. 2000; Kanwisher et al. 1997; Peelen and Downing 2005). Although some studies have reported a relationship between these domain-selective visual regions and domain-selective repetition effects (James et al. 2002; Mahon et al. 2007), the majority of studies have found cross-domain RS. In these studies, perceptual repetition affected both preferred and nonpreferred regions of the visual cortex (Andersen et al. 2009; Avidan et al. 2002; Chao et al. 2002; Konen and Kastner 2008; Mur et al. 2010; Pourtois et al. 2009; Weiner et al. 2010).

In the present study, we aimed to determine the extent to which repetition across perceptual and categorical dimensions represents the same process. We investigated this by comparing the response profile of perceptual and categorical repetition in terms of sensitivity to both the hierarchical and domain-selective organization of the visual cortex. Using stimulus-level, basic-level, and domain-level RS, we reexamined the influence of basic-level repetition (now controlling closely for image similarity), and, for the first time, we addressed the influence of categorical repetition at the domain level.

**METHODS**

**Participants**

Sixteen subjects (9 females, age: 25.3 ± 4.8 yr, mean ± SD) participated in this study. All subjects gave written informed consent, and all procedures were approved by the Harvard University Institutional Review Board.

**Stimuli**

Stimuli consisted of grayscale images of 90 tool and 90 animal stimuli subtending ~9° of visual angle, presented against a white background. Three exemplars were chosen for each object, resulting in 30 objects for each domain. Care was taken in the selection of basic-level and control repeats. First, the varying exemplars of each object were selected to be subjectively as distinct and visually dissimilar as possible. In a minority of cases (~6.1%), it was not possible to unambiguously convey that exemplars depicted separate identities (c.f. the panda in Fig. 1). Thus it is not possible to exclude the possibility that in this small number of cases, basic-level repetitions were confused with identity-level repetitions. After selecting subjectively dissimilar basic-level exemplars, we employed multiple metrics to match image-based visual similarity between basic-level and control repeats for pixelwise similarity (Chouinard et al. 2008; Grill-Spector et al. 1999), two-dimensional spectral similarity, and the degree of shape deformation needed to map one image onto the space of the other (Belongie et al. 2002; Fig. 1A).

The pixelwise technique was selected to permit comparisons with previous RS studies (Chouinard et al. 2008; Grill-Spector et al. 1999). To more extensively control for image similarity, we augmented this measure with indexes reflecting similarity of shape (Belongie et al. 2002) as well as the complexity and angular properties of images (spectral). The shape base technique we employed is widely applied in the field of computer vision and has previously been used to control visual similarity during the investigation of RS (Mahon et al. 2007). This technique is robust to simple object transformation and captures an important element of object form. The 2D Fourier measure is sensitive to a wide range of image properties, capturing elegant fine and coarse detail in addition to the dominance of angles within the image. We observed that the combined application of these three metrics produced robust measures of similarity with high prima facie validity (see Supplemental Fig. S1 for examples of image pairs). (Supplemental material for this article is available online at the Journal of Neurophysiology website.) To validate these measures, we compared the objective ratings obtained for 60 image control pairs with the subjective ratings of 6 observers. To minimize semantic bias due to categorical membership, subjective ratings were limited to comparisons between different basic-level objects. This revealed strong (r = 0.69) and highly significant correlations (P < 0.0001) between our objective metric and subjective measures of visual similarity (see Supplemental Materials).

Control pairs were iteratively selected by a supervised algorithm, separately for each subject. Within runs, pairs of randomly occurring same-domain S1 presentations were selected that matched the lag-distribution of the S1–S2 pairs in the experimental conditions. For each domain, if the randomly selected control pairs were less visually similar than the basic-level pairs, alternative pairs replaced the most dissimilar subset of pairs (still matched for S1–S2 lag). This process was continued until basic-level and control pairs were matched on pixelwise, shape, and spectral measures. If visual similarity could not be matched for a stimulus sequence, the entire stimulus sequence was regenerated and the process repeated.

**Procedure**

The experiment consisted of four sessions, ~5.5 min each. In each session, 120 images and 35 null (fixation) events were presented. Images were presented for 300 ms, followed by a fixation period of 1,700 ms. Within each session, objects were presented twice, with either a repetition of the same-object exemplar (stimulus level) or two different exemplars of the same object (basic level; see Fig. 1B). In even sessions, stimuli presented in the preceding session as a stimulus-level repeat were now presented as a basic-level repeat and vice versa. To ensure that participants attended to the task, subjects indicated whether the viewed object was smaller or larger than a 12-in. box.
In an event-related design, repetition from the first stimulus (S1) to the second (S2) occurred with a lag of between zero and seven intervening stimuli. Control stimuli were taken from the S1 presentations of two different objects from the same domain (see Fig. 1B). For each subject, presentation order was iteratively generated to ensure that control and basic-level repeats were matched for image-based visual similarity and for the mean and distribution of S1-S2 lag.

Repetition effects across conditions were controlled through counterbalancing. Because of the randomized distribution of the four conditions, the probability of a given S1 or S2 presentation from one of the four conditions being preceded by an S2 presentation was equal across all conditions. In this way, any potential influence of across-condition repetition was present equally across conditions and across S1/S2 presentations.

Data Acquisition

Functional neuroimaging data were acquired using a 3T MAGNETOM Trio scanner (Siemens Medical Systems, Erlangen, Germany) with a 12-channel head coil used for radiofrequency transmission and reception. Visual stimuli were projected into the scanner through a mirror system mounted on the head coil. Head movement was minimized by mild restraint and cushioning. Thirty-one interleaved axial slices of functional MR images were acquired using echo planar blood oxygen level-dependent (BOLD) imaging ($3 \times 3 \times 3$ mm, TR = 2 s, TE = 30 ms), covering the entire cortex.

Fig. 1. Stimulus matching and functional MRI paradigm. A: to ensure that low-level visual similarity could not account for repetition suppression (RS) difference between object-level and control repeats, stimuli were matched across 3 separate metrics. First, differences at each pixel were accrued across the image; second, spectral dissimilarities were quantified using a 2D Fourier transform; and third, the degree of shape deformation needed to map one image onto another was determined. Differences were matched over the specific stimuli sequence used for each subject box, and whisker plots represent mean dissimilarity values for each subject. B: event-related design. Participants viewed a sequence of images. Within the sequence, stimulus-level, object-level, and control repeats (always selected from the same domain) were present. The lag between the first and second stimulus ranged from 0 to 7 intervening stimuli. C: examples of control stimuli selected as visually similar using the image-based matching process explained in A. These stimuli are drawn from the lower end of the dissimilarity distribution ($Z < -1$; see Supplemental Fig. S1 for details).
Data Analysis

Preprocessing. Data were analyzed using SPM8 software (www.fil.ion.ucl.ac.uk/spm/). After the first four volumes were discarded, all volumes were slice-time corrected, realigned to the first volume, corrected for motion artifacts, normalized into standard stereotactic space using the Montreal Neurological Institute (MNI) template, and smoothed using an 8-mm full width at half-maximum Gaussian kernel. The time series were high-pass filtered to eliminate low-frequency components (filter width 128 s) and prewhitened by means of an autoregressive model, AR(1).

Stimulus- and basic-level repetition effects. First-level analysis was based on a conventional general linear model (GLM; Friston et al. 1995a, 1995b) using eight regressors (S1 and S2 for both repetition types and both domains). In a separate GLM, matched control pairs (in terms of image similarity and S1-S2 lag) were taken to the second level, and nonmatched and S2 events were modeled as effects of no interest. Because S1-S2 lag is not explicitly modeled in these GLMs, the results reflect RS pooled across lag-latencies. A repetition suppression index (RS index) was calculated to capture the suppression of the response to the repeated stimulus (S2) as a proportion of the original (S1) response: (S1 - S2)/(S1 + S2), where if any value was below zero, the absolute value of the minimum was added to both S1 and S2 to remove any confounding effects of negative values (c.f., Simmons et al. 2007). Voxelwise RS index maps were calculated for each subject for each of the four RS conditions, and these were included in a second-level random-effects GLM. All reported voxel-wise results are corrected for multiple comparisons at P < 0.05, corrected for family-wise error (FWE) at the cluster level (Poline et al. 1997). For the region of interest (ROI) analyses, data from five peak locations for the animal and tool domain-selective regions (as determined by an independent contrast, all animals vs. all tools) were extracted. These values were then subjected to separate analyses of variance (ANOVA) for each group of domain-selective regions with the factors region (5 levels), RS type (stimulus level, basic level), and domain (animals, tools). Greenhouse-Geisser correction was applied in cases where the assumption of sphericity was violated.

Domain-level repetition effects. In addition to stimulus- and basic-level effects, the event-related design allowed us to investigate repetition effects at the domain level by determining the degree to which the response to a stimulus is suppressed by the number of preceding stimuli from the same domain. It is not possible to directly assess this using the BOLD response, since successive repetition of within-domain stimuli preferentially stimulates domain-selective cortex. However, by using a regressor that determines the parametric alterations in the canonical hemodynamic response on each trial, contingent on the number of preceding within-domain items, it is possible to model how repetition modulates the evoked BOLD response (Henson et al. 2000). Specifically, for the first instance of an exemplar within each run (to exclude direct repetition effects), the proportion of within/ across-domain exemplars of the preceding five trials was calculated to create a proportional rating (from 0 to 1 in steps of 0.2). This rating was used to create a parametric regressor for each subject and each domain, and the resulting beta values were included in a second-level random-effects GLM.

RESULTS

Behavioral Responses

Subjects performed a size-judgment task (“Is the item bigger or smaller than a 12-in. box?”) within the scanner (all subjects responded on >96% of trials). Reaction time data are presented in Table 1. Median reaction times were faster for tools (mean = 656 ms, SD = 61 ms) than for animals (mean = 680 ms, SD = 67 ms; P < 0.01), despite the fact that similar proportions of tools (50.3%) and animals (52.6%) were judged to be larger than a 12-in. box (t < 1). Although interesting, this behavioral difference does not influence the results or conclusions of this study and is thus outside the scope of this investigation. Paired t-tests confirmed a significant effect of priming for both stimulus domains at both the basic level and the stimulus level (all P < 0.001).

RS Due to Image Similarity

As noted earlier, control pairs were matched on pixelwise, spectral, and shape-based criteria to assess an alternate hypothesis that basic-level RS effects are attributable to image-based visual similarity (Chouinard et al. 2008). Across all analyses (both ROI and whole brain), image-based visual similarity between control pair images produced no significant RS. This result demonstrates that mere visual similarity, as captured by image-based measures, is not a plausible rival hypothesis for the presence of basic-level RS reported in earlier studies.

Consistent Patterns of Stimulus-Level and Basic-level Repetition Effects

Domain-selective regions were identified by contrasting all-tool presentations with all-animal presentations (Fig. 2A and Table 1). This revealed the typical network of tool- and categorical repetition effects (Makin et al. 2000).

Table 1. Location, Z-score, and activity vs. fixation baseline of peak value for the contrast tools vs. animals

<table>
<thead>
<tr>
<th>Hemisphere</th>
<th>Region</th>
<th>Coordinates</th>
<th>Z Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>x  y  z</td>
<td></td>
</tr>
<tr>
<td>Animal</td>
<td></td>
<td>Preferred v.</td>
<td>Animals vs. null</td>
</tr>
<tr>
<td>Left</td>
<td>pIFG</td>
<td>−39 −64 −17</td>
<td>3.66</td>
</tr>
<tr>
<td>Right</td>
<td>pIFG</td>
<td>42 −64 −17</td>
<td>5.28</td>
</tr>
<tr>
<td>Right</td>
<td>pIFT</td>
<td>51 −76 1</td>
<td>8.38</td>
</tr>
<tr>
<td>Right</td>
<td>aIFG</td>
<td>42 −46 −23</td>
<td>7.22</td>
</tr>
<tr>
<td>Right</td>
<td>STS</td>
<td>51 −43 16</td>
<td>3.98</td>
</tr>
<tr>
<td>Tool</td>
<td></td>
<td>Preferred v.</td>
<td>Animals vs. null</td>
</tr>
<tr>
<td>Left</td>
<td>SOG</td>
<td>−36 −91 19</td>
<td>4.85</td>
</tr>
<tr>
<td>Left</td>
<td>MTG</td>
<td>−51 −58 −11</td>
<td>4.07</td>
</tr>
<tr>
<td>Left</td>
<td>IPS</td>
<td>−18 −67 49</td>
<td>4.05</td>
</tr>
<tr>
<td>Left</td>
<td>aIFG</td>
<td>−27 −58 −11</td>
<td>5.71</td>
</tr>
<tr>
<td>Right</td>
<td>aIFG</td>
<td>30 −49 −11</td>
<td>4.61</td>
</tr>
</tbody>
</table>

Note that for both stimulus classes, the right superior temporal sulcus (STS) is less active during image presentation than fixation. Null, fixation event; pIFG, posterior lateral fusiform gyrus; pIT, posterior inferotemporal cortex; aIFG, anterior lateral fusiform gyrus; SOG, superior occipital gyrus; MTG, middle temporal gyrus; IPS, intraparietal sulcus; aIFG, anterior medial fusiform gyrus.
animal-selective regions, exhibiting left hemispheric dominance for tools and right hemispheric dominance for animals (c.f. Chao and Martin 2000; Downing et al. 2006; Mahon et al. 2007). Tool-selective regions included bilateral anterior medial fusiform gyrus (amFG), left middle temporal gyrus (MTG), intraparietal sulcus (IPS), and a section of the superior occipital gyrus (SOG). Regions more responsive to animals than tools consisted of bilateral posterior fusiform gyri (pFG) as well as the right anterior lateral fusiform gyrus (alFG), posterior inferotemporal cortex (pIT), and superior temporal sulcus (STS).

To characterize perceptual and categorical effects, RS indexes were extracted from each ROI (see Fig. 2B and Table 1). Both animal- and tool-selective regions were subjected to separate ANOVAs with the factors region (5 regions), RS type (control, object-level, and stimulus-level), and domain. Testing the alternate hypothesis that the RS index > 0 revealed that RS occurred in both tool and animal regions.
demonstrated a greater proportional change than other regions. Together, these results indicate that there is evidence for an overall effect of RS across animal- and tool-selective regions but no evidence that RS differed for either of the two RS types or within vs. across domains. Although there appears to be greater RS for stimulus level than basic level, this trend did not approach significance. In a supplemental voxelwise analysis, we documented the high spatial extent of domain-insensitive RS effects across the ventral stream (see Fig. S3) and also verified that these results were not an artifact of the proportional RS index used (see Supplemental Materials). Of note, image-based visual similarity was not matched between the animal and tool domains. Consequently, the lack of differences in RS between domains is further support that image-based visual similarity is not the driving force behind ventral stream RS.

In contrast, regional differences were seen to manifest within animal-selective cortex in terms of interactions between region for both RS type \([F(1,2.18) = 8.2, P < 0.01]\) and stimulus category \([F(1,4.21) = 5.3, P < 0.05]\). These effects reflected the highly significant interaction between region and RS type when only the influence of animal stimuli in animal regions was considered \([F(3,45) = 10.6, P < .00001; \text{c.f. Fig. 2B, bottom left}]\). This interaction was driven by the inverted pattern of basic-level and stimulus-level RS observed in the STS \([\text{RS}_{\text{basic level}} = 0.44, \text{RS}_{\text{stimulus level}} = -0.06, T_{(15)} = 3.5, P < 0.005; \text{Fig. 2B, bottom left}]\). This demonstrates that RS only occurs in this region during basic-level repetition of animal stimuli. No other main effect of, or interactions with, RS type were observed, and no similar effects were observed when the complimentary analysis was conducted in tool regions \((P > 0.13\) across all comparisons).

The magnitude of RS effects was not homogenous across regions. Collapsing across RS type, differing strengths of RS effects were seen across tool-selective regions of the visual system \([F(3,7,55.6) = 4.7, P < 0.005; \text{c.f. Fig. 2B, left}]\). MTG demonstrated a greater proportional change than other regions (right amFG and left SOG, \(P < 0.05\), Bonferroni corrected; in addition, the left amFG and IPS were significantly different at an uncorrected level, \(P < 0.05\)). This interaction (in conjunction with the RS index used) formally demonstrates that RS effects are not a simple proportional scaling of the initial response and that sensitivity to repetition differs across cortical regions.

A voxelwise analysis was performed to determine the overlap in stimulus-level and basic-level repetition effects across the entire ventral stream. Specifically, we sought to address the theoretical question, do regions showing strong stimulus-level RS also show significant evidence of basic-level RS? To this end, ventral visual stream regions demonstrating stimulus-level RS were identified at a high threshold \((P < 0.001)\), and within this localizer mask, basic-level RS effects were examined at a relaxed threshold \((P < 0.05)\). These results, presented in Fig. 3, demonstrate the large degree of overlap in the cortical regions contributing to stimulus-level and basic-level RS. Within occipitotemporal regions, 75% of voxels demonstrating a strong stimulus-level RS also showed significant evidence for basic-level RS. Even though we employed a relaxed initial threshold, because of the large extent of these clusters, they remain significant after correction \((P < 0.05, \text{FWE corrected})\). For reference, clusters showing RS at a relaxed threshold \((P < 0.05)\) for both stimulus-level and basic-level RS are also represented in light blue in Fig. 3. At this lenient threshold, a region of the left prefrontal cortex is also apparent. This region is consistent with findings of previous studies (Chouinard et al. 2008; Vuilleumier et al. 2002); however, caution is appropriate in interpreting activation at this unconventional threshold. Because of the uncertain nature of this effect, this finding is not discussed further. This analysis allows us to avoid the erroneous conclusion that these brain regions are not influenced by basic-level repetition. This allows us to make the qualitative statement that those regions we identified as sensitive to stimulus-level RS are also sensitive to basic-level RS but does not demonstrate that these effects are quantitatively similar. Quantitative effects are discussed below.

Fig. 3. Overlapping neural substrates for stimulus-level and object-level RS. To determine whether object-level RS occurs in the same cortical regions as stimulus-level RS in an attenuated form, those regions where stimulus-level RS was observed (threshold, \(P < 0.001\)) were examined for the presence of item-level RS at a lenient threshold \((P < 0.05)\); due to the expanse of these clusters, both stimulus-level and object-level RS effects are significant at the cluster level, \(P < 0.001, \text{family-wise error (FWE) cluster corrected}\). Yellow outlines enclose regions where RS was significantly larger at the item level than the stimulus level (threshold, \(P < 0.001; \text{cluster-level correction}, P < 0.05, \text{FWE cluster corrected}\)). Only the anterior FG in the right hemisphere met the double criteria of significantly greater stimulus-level RS and no significant basic-level RS at 0.05 (cluster-level significance, \(P < 0.05, \text{FWE cluster corrected}\)). For reference, sections colored in light blue indicate those clusters of activation (>200 voxels) outside the a priori focus of this analysis that show moderate \((P < 0.05, \text{uncorrected})\) RS for both stimulus-level and basic-level repetition.
Although there was extensive overlap between the regions showing stimulus-level and basic-level RS, some regions did exhibit dissociations. Superior sections of the lateral occipital cortex and IPS showed significant stimulus-level, but not basic-level, RS. However, there were no significant differences between these two conditions, limiting the ability to make formal statements about the role of these regions. In contrast, bilateral MTG (MNIxyz = −45, 64, 5; 48, −58, −11) and the left medial anterior fusiform (−30, −34, −23) showed significantly greater RS effects at the stimulus level (yellow outlines in Fig. 3), indicating that the physical properties of the stimulus had a greater influence on RS. However, these regions still showed some, albeit weaker, basic-level RS. Only a subsection of the right amFG (36, −46, −23) met the dual criteria of no evidence for basic-level RS effects and significantly greater RS at the stimulus level. This result indicates that this region is specifically sensitive to visual form and not to basic-level categorical membership.

### Domain-Level Repetition Effects

Finally, using a parametric regressor, we investigated whether the largely overlapping effects observed between basic-level and stimulus-level repetition extend to repetition at the domain level. In this analysis, image-based visual similarity was not matched between stimulus domains. Since we observed no influence of image-based similarity on RS and no differential pattern of RS effects between stimulus domains in our analysis of stimulus-level and object-level repetition effects, it is unlikely that image-based similarity can account for domain-level effects. However, it is not possible to exclude the possibility that domain-level RS is in some way influenced by image-based visual similarity. Because these parametric effects were anticipated to be moderate, a more sensitive analytic approach was employed whereby ventral stream regions showing a main effect of RS were used as a mask for voxelwise analysis (thereby reducing the severity of corrections required for multiple comparisons), and an initial threshold of $P < 0.01$ (instead of 0.001) was used. The use of the mask increased statistical power but limited the ability to make inference to those regions showing maximal stimulus-level RS (which were of greatest a priori interest). The main effect of domain-level repetition was significant in two clusters ($P < 0.05$, FWE corrected) in the left and right anterior FG (see Fig. 4). Within these regions, post hoc testing revealed no significant difference between animal and tool domain-level RS. This finding demonstrates that repetition, even within broad categories such as domain, influences neural representation within the visual cortex.

To further determine the degree of concordance between perceptual RS (stimulus level) and both levels of categorical RS (basic level and domain level), we identified six ventral stream regions where stimulus-level RS was strongest. These were bilateral IOG, MTG, and aFG (Table 2). Interestingly, these peaks correspond to the three most strongly activated regions of visual cortex when viewing either tool or animal stimuli (compared with fixation). Within these peak locations, the presence of categorical RS effects was tested independently. As expected from previous analyses, basic-level RS was indeed present in all of these locations. In addition, domain-level RS was observed in five of the six regions. Together, these results indicate that, at multiple levels of the visual processing hierarchy, both basic-level and domain-level repetition effects rely on brain regions previously only thought to exhibit stimulus-level repetition effects.

### Table 2. Effects of RS within ventral stream regions

<table>
<thead>
<tr>
<th>Hemisphere</th>
<th>Region</th>
<th>Coordinates</th>
<th>Stimulus Level</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$x$</td>
<td>$y$</td>
<td>$z$</td>
<td>$P$ (FWE)</td>
<td>Extent</td>
</tr>
<tr>
<td>Left</td>
<td>IOG</td>
<td>−45</td>
<td>−73</td>
<td>−8</td>
<td>$&lt;0.001$</td>
<td>1,565</td>
</tr>
<tr>
<td>Left</td>
<td>MTG</td>
<td>−42</td>
<td>−58</td>
<td>−14</td>
<td></td>
<td>7.07</td>
</tr>
<tr>
<td>Left</td>
<td>aFG</td>
<td>−33</td>
<td>−40</td>
<td>−26</td>
<td></td>
<td>6.65</td>
</tr>
<tr>
<td>Right</td>
<td>aFG</td>
<td>36</td>
<td>−43</td>
<td>−23</td>
<td>$&lt;0.001$</td>
<td>1,301</td>
</tr>
<tr>
<td>Right</td>
<td>MTG</td>
<td>45</td>
<td>−58</td>
<td>−14</td>
<td></td>
<td>6.78</td>
</tr>
<tr>
<td>Right</td>
<td>IOG</td>
<td>36</td>
<td>−85</td>
<td>7</td>
<td></td>
<td>6.46</td>
</tr>
</tbody>
</table>

Peak coordinates, cluster significance, and cluster size (voxels) are presented for the simple effect of stimulus-level repetition. In addition, at the peak of each region, the effects of basic-level and domain-level repetition suppression (RS) are independently calculated. FWE, familywise error corrected.
DISCUSSION

In this study we sought to determine the similarities and differences between repetition effects over perceptual and categorical dimensions. To accomplish this, we determined the degree to which these forms of repetition have similar response properties in terms of the hierarchical and domain-selective organization of the visual cortex. Surprisingly, a highly similar pattern of perceptual and categorical repetition effects was observed. That is, both forms of RS were independent of domain selectivity (i.e., tool stimuli produce RS in animal-selective visual cortex and vice versa), and both occurred across multiple levels of the visual hierarchy. In contrast to this overall pattern, dissociations were also observed. The right STS demonstrated robust basic-level, but not stimulus-level, RS for animal stimuli. Conversely, several regions of the ventral stream demonstrated stronger stimulus-level RS (bilateral MTG and aFG), with a section of the right aFG showing stimulus-level RS in the absence of detectable basic-level RS.

Consistent Patterns of Perceptual and Categorical Repetition Effects

Overall patterns across domain-selective regions revealed that stimulus-level and basic-level RS were independent of domain selectivity. This finding is in accord with the majority of studies that have examined RS across perceptual dimensions (Andresen et al. 2009; Avidan et al. 2002; Chao et al. 2002; Konen and Kastner 2008; Mur et al. 2010; Pourtois et al. 2009). Pourtois et al. (2009) found RS effects for houses, chairs, and faces generalized across multiple ventral visual regions. In that study, only in a specific region of the aFG (outside the region identified in an independent face localizer) was RS observed for faces that did not reach significance for either chairs or buildings. However, the results of the present study, where RS was observed at the same location for both animals and tools (see Supplemental Materials), suggest that this previous finding may be attributable to a lack of statistical power. In another study, Mur et al. (2010) investigated identity-level RS for faces in a fixed-effects analysis and likewise found RS in both face-selective and -nonselective regions. Similarly, Weiner et al. (2010) investigated RS in house-, face-, and body-selective regions and found cross-domain RS for a range of object categories.

The present study supports these previous findings and extends them to show that the pattern of insensitivity to domain selectivity generalizes to basic-level categorical repetitions. Although domain selectivity is a clear organizational principle of the ventral stream (e.g., Kanwisher et al. 1997), neurons in the visual processing stream outside of domain-selective regions contain sufficient information to encode both distinct objects and different domains (Eger et al. 2008a, 2008b; Haxby et al. 2001). Although the comparison of one domain against another emphasizes the domain-selective nature of the visual cortex, it can also obscure the fact that the majority of neural activity elicited by a stimulus (compared with null events) is shared between domains while a small proportion of the overall neural activity demonstrates domain selectivity. In the present study, the strongest proportional repetition effects were in those regions that also demonstrated the strongest overall response, bilateral IOG, MTG, and aFG. With the exception of the left MTG, these regions were not identified as domain selective (the right aFG was not overlapping with that identified in the domain selectivity analysis). The ROI analysis and the proportional RS index employed in this study reveal that the susceptibility to RS in these regions extends beyond a mere scaling of the initial response and suggests that maximally recruited regions are functionally biased toward repetition sensitivity. Together, these results indicate that it is the magnitude of the neural response rather than the domain-selective profile that plays the greatest role in determining susceptibility to RS.

In addition to determining the generality of RS effects across domain-selective cortex, we investigated RS effects across the hierarchical architecture of the visual system. Within the ventral stream, 75% of voxels sensitive to stimulus-level repetition were also sensitive to basic-level repetition. This result demonstrates that basic-level RS is widespread throughout the visual system and is not limited to anterior regions in the left FG, as had been suggested by previous work (Koutstaal et al. 2001; Simons et al. 2003). This finding mirrors recent developments in the identity-level RS literature, where the apparent localization of identity-level RS within the left aFG (Vuilleumier et al. 2002) has been shown to extend across greater sections of the visual cortex (Mur et al. 2010; Pourtois et al. 2009). Furthermore, we used a novel analytic technique to extend the study of categorical repetition effects beyond basic-level relationships to locate the subtle influence of domain-level repetition. A voxelwise analysis revealed domain-level RS bilaterally within the aFG. We extended this analysis by investigating domain-level RS in the three bilateral regions that showed the strongest stimulus-level RS. This analysis demonstrated that in these three bilateral regions, domain-level RS (as well as basic-level RS) was observed across hierarchical levels, in the left IOG and bilaterally in the MTG and anterior FG.

The high degree of similarity in the pattern of stimulus-level and categorical repetition effects indicates that the processing of perceptual and categorical relationships occurs across a similar neural substrate and may rely on shared cognitive characteristics. However, some differences were also observed between these two types of RS, and these are discussed below.

Dissociation between Perceptual and Categorical RS Effects

Within the ROI analysis of domain-selective regions, the right STS showed RS only during basic-level repetition of animal stimuli. This result is of particular interest because stimulus-level repeats also contain identical basic-level representations, in addition to the repeated perceptual factors. Therefore, the right STS is sensitive to basic-level repetition only when properties of the stimulus vary, suggesting that this region may be involved in linking occurrences of dissimilar representations of the same biological objects. This would be in accord with the diverse role of this structure in varied forms of cognition, ranging from the perception of biological stimuli to attention and even to integration across the modalities (c.f. Hein and Knight 2008) and its role in default mode, internalized processes (Greicius et al. 2003) (note the negative values compared with baseline in Table 1). However, future studies are needed to understand the exact nature of this effect within the right STS.

When stimulus-level and basic-level RS effects were analyzed across voxels in the ventral visual stream, stimulus-level
RS was greater in the MTG and aFG. This indicates that the reoccurrence of specific stimulus events is reflected in the activity of multiple regions of the visual system. Moreover, a medial subsection of the right aFG demonstrated only stimulus-level RS with no evidence for basic-level RS. This is of interest because it suggests that a particular perceptual event, a particular stimulus viewed from a particular angle, influences activity independently of the categorical aspects of that stimulus in anterior regions of the visual system. Our results are consistent with recent findings of viewpoint-dependent RS in anterior face-selective fusiform regions and viewpoint-invariant RS in the lateral occipital sulcus (Pourtois et al. 2009). However, our RS effects challenge theories that suggest representations at higher levels of the visual processing hierarchy are less stimulus dependent. Together, these results indicate that a simple, uniform posterior-anterior axis running from stimulus-bound to abstracted representations is an insufficient account of the neural organization of object processing. Instead, these results imply that some anterior fusiform regions form representations that remain tightly bound to the physical properties of a particular environmental event. The location of this representation within anterior regions of the FG suggests that such stimulus-bound representations may be beneficial to higher level processing.

Role of Visual Similarity in Ventral Stream RS

In the present study, we sought to discount the hypothesis that basic-level RS in the ventral stream is solely the repercussion of image similarity. To this end, we augmented the pixelwise measures used in previous studies with spectral and shape-based indexes of dissimilarity. The use of these multiple metrics allowed the tight control of image-based properties, but these measures may not have been sensitive to all aspects of perceptual similarity. Validation of our scale with subjective ratings of visual similarity indicated that the objective measure had high construct validity while avoiding any potential semantic bias arising due to categorical membership. However, semantic bias is an intrinsic part of visual similarity processing. Targeted electrophysiological studies using words have shown shape-based visual semantic features are sufficient to produce RS (e.g., coin-button; Kellenbach et al. 2000). This suggests that, due to stored visual knowledge, visual features may be inextricably related to categorical groups.

The results of the present study are supported by the recent finding that prior exposure to equivalent word stimuli produces suppression in lateral occipital cortex during picture presentation (Horner and Henson, in press). Together with the present study, this suggests that RS is not dependent on low-level perceptual similarity. However, even in the absence of common perceptual input, intrinsic relationships between categorical knowledge and stored knowledge about visual features may play a role in ventral stream RS.

Implications of Repetition Suppression for Cognitive Function

fMRI repetition suppression has been widely interpreted to be an index of neuronal selectivity and a potential means to probe neuronal selectivity at the subvoxel level (Grill-Spector et al. 1999; Grill-Spector and Malach 2001; Koutstaal et al. 2001; Simons et al. 2003). This interpretation is based largely on findings of neuronal adaptation at the single-unit level (Dragoi et al. 2000; Kohn and Movshon 2003) and assumes that repeated, bottom-up stimulation of a neural population causes either the scaling or sharpening of neurons encoding the relevant stimulus feature (Grill-Spector et al. 2006). However, the correspondence between fMRI-indexed adaptation and single-unit neuronal adaptation has been contested (Sawamura et al. 2006; Tolias et al. 2005). Moreover, fMRI-indexed RS does not appear to arise solely due to passive bottom-up processes. RS is reduced when repetition events are rare and therefore unexpected, demonstrating the influence of top-down processes (Summerfield et al. 2008). These top-down effects are extensive. Viewpoint-independent RS for individual face identity extends downstream as far as early visual cortex (anatomically consistent with V1; Mur et al. 2010). These findings illustrate that RS does not solely arise from computations within the suppressed neural populations but also as a repercussion of operations computed elsewhere. This places mechanistic models of RS that only reflect local neuronal properties in doubt. Furthermore, the results of the present experiment are also inconsistent with the predictions of these models.

A definitive prediction of the scaling model is that the degree of RS is scaled by a constant fraction of the initial $S_1$ response. This means that absolute RS will be greatest in optimally responsive neural populations. This is due to the fact that an equivalent proportional change in activity will translate into the maximal absolute change in optimally responsive neurons (Grill-Spector et al. 2006). Because of the proportional RS index used in the present study, this prediction would manifest as an equivalent RS index value across brain regions. As noted earlier, this is not the case. Those regions that produced the greatest response to the $S_1$ stimulus (MTG, IOG, and aFG) are those that show the greatest proportional RS. In other words, we observed a hyperscaling: maximally responsive brain regions demonstrate a greater scaling factor than nonoptimally responsive neurons. Interestingly, the factor that determines optimality is the absolute responsiveness to the visually presented stimulus and not domain selectivity (see Fig. 2B and also Fig. S3). This hyperscaling indicates that a simple scaling model cannot explain RS but that different neuronal populations are disproportionately sensitive to the processes that cause RS.

The observed independence between RS and domain selectivity is inconsistent with the main prediction of the sharpening model. This model predicts that the tuning of the neural response to a feature becomes more selective after repetition (Grill-Spector et al. 2006). Thus activity outside domain-selective regions will show the most attenuation following repetition of stimuli from that domain. Since our results show equivalent RS for both preferred and nonpreferred stimuli, sharpening does not appear to be a reliable model for RS effects in the ventral stream observed in this study (however, see Weiner et al. 2010, who suggest that sharpening may be present in some ventral stream regions over longer repetition time scales).

Both the previous literature and the results of the present study are inconsistent with sharpening or scaling models. Consequently, any interpretation of RS or fMRI-adaptation as a direct index of neuronal tuning at a subvoxel level cannot be assumed. Rather, when considering RS, one must consider the influence of other brain regions and other higher level cognitive processes that may influence the suppressed region. Accordingly, the appropriate
mechanistic model of RS must also account for these properties. The facilitation model of RS emphasizes interactions occurring over multiple levels of the processing stream and posits that increases in neural efficiency, as indexed by behavioral priming, will result in a faster, briefer neural response (Grill-Spector et al. 2006). A model by which repetition produces increased efficiency and neural facilitation is offered within the predictive coding framework. This framework assumes that incoming sensory information is assessed against prior (implicit) expectation and knowledge rather than solely bottom-up sensory input (Friston 2005). Under this model, RS results from informed expectation, effectively disambiguating incoming sensory information and reducing the neural activity required for successful perceptual or cognitive categorization (Friston 2005; Henson 2003; Summervield et al. 2008). The hyperscaling observed in this study is consistent with this model and would suggest that the maximum increase in efficiency occurs in regions carrying the weight of the processing load.

Within this framework, the largely similar response profiles of perceptual and categorical repetition observed in the present study demonstrate that predictions based on either prior perceptual or categorical information have qualitatively comparable influences on the processing of incoming information. That is, disambiguation based on prior experience of a perceptual or categorical nature has largely similar qualitative effects on the processing of the current stimuli. In contrast, the absence of basic-level RS in the right alFG despite the presence of strong stimulus-level RS shows that categorical information has neither a bottom-up nor a top-down influence on the processing of this region. This suggests that categorical information is largely irrelevant to computations within this region and that exact sensory information (see also Vuilleumier et al. 2002) largely determines processing in the right alFG. As noted earlier, this challenges rigid hierarchical notions of visual processing, suggesting that some concrete and not abstracted visual representations are processed at higher levels of the visual system.

Conclusion

In this study we employed perceptual and categorical repetition effects to further understand the nature of object representation in the visual cortex. We found that stimulus-level and basic-level repetition display highly similar patterns of RS, both in their relationship to domain preference and in their distribution across the visual hierarchy. Furthermore, repetition effects at the domain level were also observed across multiple hierarchical levels of the visual cortex. In contrast, a region of the right amFG responded to stimulus-level, but not object-level, repetition. These findings challenge views of a strict unidirectional, hierarchical anterior-to-posterior organization of the ventral visual stream. Instead, they suggest that the physical properties of particular stimulus events are represented in higher visual regions and that categorical relationships influence processing across multiple levels of the visual cortex.

ACKNOWLEDGMENTS

We thank Marius Peelen for helpful comments on an earlier draft of this manuscript.

GRANTS

This research was supported by the Provincia Autonoma di Trento and the Fondazione Cassa di Risparmio di Trento e Rovereto.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES


