Neuropsychologia 117 (2018) 513–519

Contents lists available at ScienceDirect

Neuropsychologia

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

Representational differences between line drawings and photographs of natural scenes: A dissociation between multi-voxel pattern analysis and repetition suppression

Thomas P. O'Connell[a,*], Per B. Sederberg[b], Dirk B. Walther[c]

[a] Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205, USA
[b] Department of Psychology, University of Virginia, Charlottesville, VA, USA
[c] Department of Psychology, University of Toronto, Toronto, Ontario, Canada

ARTICLE INFO

Keywords:
natural scene perception
fMRI
multi-voxel pattern analysis
repetition suppression

ABSTRACT

Distributed representations of scene categories are consistent between color photographs (CPs) and line drawings (LDs) in the parahippocampal place area (PPA) and the retrosplenial cortex (RSC), as shown using multi-voxel pattern analysis (MVPA). Here, we used repetition suppression (RS) to further investigate the degree of representational convergence between CPs and LDs of natural scenes. MVPA and RS can capture different aspects of visual representations, and RS may prove useful in elucidating important differences in the representations of CPs and LDs of natural scenes. We performed an event-related fMRI experiment, including image-repetitions either within-type (i.e., CP to CP or LD to LD) or between-types (CP to LD, LD to CP). We found significant RS for within-type repetitions in PPA, RSC and the occipital place area (OPA), but did not observe RS for between-types repetitions. By contrast, scene categories were decodable from activity patterns evoked by both CPs and LDs using SVM classification for both within-type decoding and between-types cross-decoding. We conclude that there are representational differences between CPs and LDs in scene-selective cortex despite a category-level correspondence.

1. Introduction

Over the last two decades, functional magnetic resonance imaging (fMRI) has become an indispensable tool for uncovering the neural processes underlying perception and cognition. Two analysis methods have proven particularly useful for the probing of neural representations of perception, memory, intentions, and other cognitive processes: repetition suppression (RS) and multi-voxel pattern analysis (MVPA). Both methods have helped to elucidate the representational underpinnings of perceptual and cognitive processing in the brain, but recent research has uncovered incongruencies between the two methods.

MVPA leverages spatial patterns of blood-oxygen-level-dependent (BOLD) activity using similarity measures or machine-learning to decode (predict) content from brain activity (Haxby, 2001; Cox and Savoy, 2003). RS is the reduction of BOLD signal across multiple repetitions of a stimulus (Grill-Spector and Malach, 2001; Fang et al., 2005; Grill-Spector et al., 2006; Park et al., 2007; Turk-Browne et al., 2012). Epstein and Morgan (2012) observed RS in the PPA, the RSC, and the occipital place area (OPA) for repeated landmark identity. However, they did not find RS for repeated scene category, even though scene categories were decodable using MVPA. Ward et al. (2013) showed differences in representations that track explicit and implicit memory using MVPA and RS. Park and Park (2017) found a signal for scene texture in the parahippocampal place area (PPA) using MVPA but not using RS. Recently, Hatfield et al. (2016) identified neural signatures of object orientations using MVPA but not RS in lateral occipital cortex, and proposed a framework to characterize the divergent results provided by RS and MVPA in light of the underlying neural activity thought to drive each effect. Specifically, they proposed that significant RS reflects co-activation of the same neural populations across stimulus repetitions while significant MVPA reflects activity in either identical populations of neurons or reliably clustered but not necessarily identical populations of neurons. They suggest co-use of both RS and MVPA to explore similarity in neural representations across different stimulus manipulations to better characterize the underlying neural signal driving significant MVPA effects. Here, we use this framework to explore the nature of neural populations involved in encoding natural scenes.

Natural scenes are central to humans' visual experience in tasks such as navigating through complex environments (Maguire et al.,

* Corresponding author.
E-mail address: thomas.oconnell@yale.edu (T.P. O'Connell).

https://doi.org/10.1016/j.neuropsychologia.2018.06.013
Received 3 January 2018; Received in revised form 14 June 2018; Accepted 17 June 2018
Available online 21 June 2018
0028-3932/

1997; Maguire, 1998), extracting visual context (Bar, 2004; Oliva and Torralba, 2007), and categorizing unknown scenes (Fei-Fei et al., 2007; Greene and Oliva, 2009; Walther et al., 2009). Our research has demonstrated that structural components of natural scenes, as preserved by line drawings (LDs), play a critical role in supporting scene categorization. LDs and color photographs (CPs) of natural scenes elicit analogous multi-voxel representations of scene categories in the parahippocampal place area (PPA) and the retrosplenial cortex (RSC) (Walther et al., 2011). Coarse scene structure is shared between CPs and LDs, unlike color and most texture. To further characterize which structural features preserved by LDs support scene categorization, we used computational modeling on LDs to show features including contour, curvature, and junctional type/angle discriminate between scene categories and generate error patterns correlated to human categorization errors (Walther and Shen, 2014) and neural encoding of scene categories (Choo and Walther, 2016). Thus, scene structure as preserved in LDs appears to play a pivotal role in representations of scene categories.

Previously, we have used MVPA to decode representations of scene categories from patterns of fMRI activity in high-level visual brain regions, identified regions where classifier error patterns match human categorization error patterns, and shown that CPs and LDs of natural scenes evoke equivalent and cross-decodable representations of scene categories (Walther et al., 2011; Walther and Shen, 2014; Choo and Walther, 2016). Here, we further probe the equivalence of scene category representations for CPs and LDs of natural scenes using RS. Additionally, our previous studies used category-blocked designs to maximize category-specific activity patterns. Additionally, we assess whether previously measured cross-decodability of scene categories from CPs and LDs holds for individually presented natural scenes.

We performed an event-related fMRI experiment in which participants viewed a stream of CPs and LDs depicting natural scenes with embedded repetitions of individual scenes, either within-type (CP-CP or LD-LD) or between-types (CP-LD or LD-CP) (Fig. 1**A**). Depending on the presence or absence of RS and MVPA effects for between-types conditions, we aimed to infer the degree of overlap between neural populations that represent CPs and LDs of scenes (Fig. 1**B**). We found RS only for within-type but not between-types repetitions in PPA, RSC, and OPA, suggesting that the representation of scenes accessed by RS does not transfer between CPs and LDs. By contrast, we were able to decode scene categories within image types as well as across image types in the same regions of interest, thereby confirming our previous block-based results. We conclude that there are representational differences between CPs and LDs in scene-selective cortex, and that CPs and LDs of scenes are represented by non-identical neural populations that are spatially clustered within voxels.

2. Materials and methods

2.1. Participants

Fifteen participants (nine males, mean age 22.3 years, range 19–26 years) completed the experiment, which was approved by The Ohio State University Institutional Review Board. All participants had normal or corrected-to-normal vision, were right-handed, provided written informed consent, and were paid for their participation.

2.2. Stimuli

Participants were shown 288 images (144 CPs and 144 LDs) depicting natural scenes from six categories: beaches, forests, city streets, highways, mountains, and offices. CPs had been previously rated by an independent set of human participants as highly representative of their respective category (Torralbo et al., 2013). LDs were created at the Lotus Hill Research Institute by trained artists tracing contours in the CPs (Walther et al., 2011). All images had a resolution of 800 × 600 pixels and occupied 9.5° × 7.1° of visual angle when viewed in the scanner using back projection.

2.3. Procedure

Participants completed 8 main experiment runs, in which they viewed a stream of CPs and LDs containing repetitions of individual scenes in a fast event-related design. Images were presented for 1000 ms each with a central fixation cross overlaid starting 500 ms before and ending 500 ms after stimulus presentation. The inter-stimulus interval (ISI) was jittered between 3 s and 6 s (mean = 4.5 s). Scenes were repeated within the same run as either the same (CP_1CP_2 or LD_1LD_2) or different (CP_1LD_2 or LD_1CP_2) image

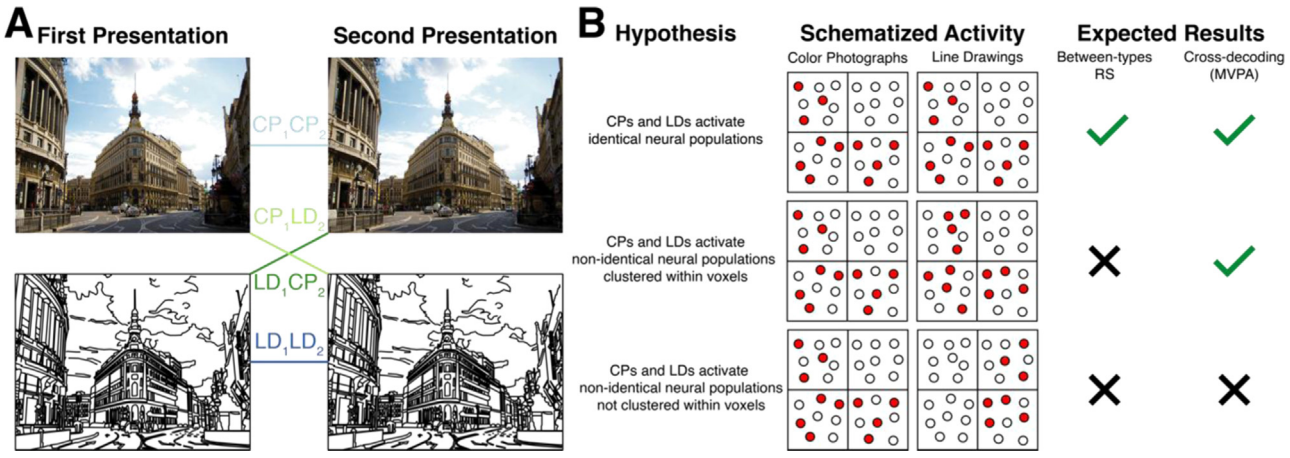

Fig. 1. A. Repetition conditions. Participants were shown color photographs (CPs) and line drawings (LDs) of natural scenes in an event-related design. Across two presentations, each scene was repeated either within image-type (CP_1CP_2 and LD_1LD_2) or between image-types (CP_1LD_2 and LD_1CP_2). **B.** Possible outcomes. The first column shows three hypotheses for how CPs and LDs may be represented relative to one another. The second column shows schematized activity for each hypothesis in a four voxel brain with eight neurons per voxel. Red indicates an active neuron. The final column shows the expected RS and MVPA results associated with each hypothesis. If the same neural populations are active for both image types, we expect to see repetition suppression (RS) for between-types repetitions (CP_1LD_2 & LD_1CP_2) and significant cross-decoding of scene categories between CPs and LDs ($CP_{Train}LD_{Test}$ & $LD_{Train}CP_{Test}$). In the second row, different neurons are active for CPs and LDs, but active neurons are clustered at a spatial scale smaller than voxels, leading to the same multi-voxel pattern of activity. In this case, we expect to see significant cross-decoding between CPs and LDs but not between-types RS. In the final row, different neurons are active for CPs and LDs, and the different neural populations do not cluster together at the sub-voxel level. In this final case, we do not expect to see RS or cross-decoding of scene category between CPs and LDs.

type (Fig. 1b). A lag of 3–7 trials (19.5–45.5 s) separated scene repetitions to avoid expectation-related RS (Summerfield et al., 2008; Turk-Browne et al., 2012; Larsson and Smith, 2012). Each run had one instance of each repetition condition for each of the six scene categories, for a total of 24 pairs (48 trials) per run. Participants performed a two-alternative forced-choice (2-AFC) natural versus manmade categorization task to engage their attention throughout the experiment.

2.4. Data acquisition

fMRI data were recorded on a 3-Tesla Siemens Tim Trio MRI with a 12-channel matrix head coil at The Ohio State University Center for Cognitive and Behavioral Brain Imaging. Functional images were obtained using a gradient echo, echo-planar imaging (EPI) sequence: TR = 2.5 s, TE = 28 ms, flip angle = 75°, matrix = 66 × 74, FOV = 220 mm, 48 transverse 3 mm slices without gap (3 × 3 × 3 mm voxels). A high-resolution (1 × 1 × 1 mm voxels) MPRAGE T1-weighted anatomical scan (TR = 1.9 s, TE = 4.68 ms, flip angle = 10°) was collected to assist in the alignment of EPI scans.

2.5. Data preprocessing

Functional data were motion corrected, spatially smoothed (4 mm FWHM), and normalized to percent signal change using AFNI (Cox, 1996). The anatomical scan and localizer scans (described below) were aligned to functional data from the main experiment.

2.6. ROIs

PPA, RSC, and OPA were identified using separate localizer sequences with blocks showing images of faces, objects, or landscapes. ROIs were defined with a standard [landscapes > objects, faces] linear contrast (Epstein and Kanwisher, 1998). A minimum threshold of $p < 0.01$ (corrected for multiple comparisons using false discovery rate) was applied to each contrast. Average bilateral ROI sizes across participants were 138 voxels (SD = 11.8) for PPA, 98 voxels (SD = 23.7) for RSC, and 99 voxels (SD = 24.9) for OPA. PPA and OPA were localized bilaterally in all 15 participants, and RSC was localized bilaterally in 12 participants, only in the right hemisphere in 2 participants, and not at all in 1 participants. Only participants with bilateral ROIs were included in the analyses for each region.

The fusiform face area (FFA) was identified in an analogous manner, using a [faces > objects, landscapes] linear contrast (Kanwisher et al., 1997). At $p < 0.01$ (corrected for multiple comparisons using false discovery rate), average bilateral FFA size was 106.47 voxels (SD = 40.46).

Primary visual cortex (V1) was included as a visually active area where we did not expect representations of scene content beyond simple visual features. We localized V1 bilaterally, using the meridian stimulation method (Sereno et al., 1995), in 14 participants. Average bilateral V1 ROIs contained 245 voxels (SD = 40.8). Technical issues prevented localization of V1 in one participant.

2.7. General linear models

We estimated single-trial activity for each participant using general linear models (GLMs). Single-trial activity within each condition and presentation combination (e.g., 1st presentation CP_1CP_2, 2nd presentation CP_1CP_2, 1st presentation LD_1LD_2) was modeled using eight separate GLMs. Regressors were defined for each individual trial in a given condition by presentation combination, and a set of nuisance regressors were defined as all other trials. We then used the stim_times_IM option in 3dDeconvolve in AFNI (Cox, 1996), which uses ordinary least squares regression to compute single-trial β-estimates for the current condition/presentation. Pooling across the eight GLMs provided single-trial β-estimates for each trial in the experiment.

2.8. Repetition suppression analysis

To control for any differences in activation magnitude for CPs and LDs, we calculated RS with respect to the average activation elicited for first presentations of the same image type as the second presentation. Average activation for novel CPs was computed by pooling across first presentations in the CP_1CP_2 and CP_1LD_2 conditions, and average activation for novel LDs was computed by pooling across first presentations in the LD_1LD_2 and LD_1CP_2 conditions. RS was then calculated as the difference between activation for a second presentation and the corresponding first presentation activation for the same image type. In the LD_1LD_2 or CP_1LD_2 conditions, RS was calculated as the activation for LD_2 minus the average activation for all first presentation LDs, whereas in the CP_1CP_2 or LD_1CP_2 conditions, RS was calculated as the activation for CP_2 minus the average activation for all first presentation CPs. For each condition, RS significance was determined using two-tailed repeated-measure *t*-tests between average baseline β-estimates and second presentation β-estimates across participants.

2.9. Whole-brain analysis

To investigate effects outside of pre-determined ROIs, whole-brain RS analysis was performed in the same manner described above on voxel-wise activity normalized to Montreal Neurological Institute (MNI) standard space. Significance of RS across subjects was assessed using paired *t*-tests ($p < 0.01$) and corrected for multiple comparisons at the cluster level.

2.10. Scene category decoding analysis

We performed six-way scene category decoding using spatial patterns of β-estimates as input for a linear support vector machine (SVM) classifier. In an eight-fold leave-one-run-out (LORO) cross validation, the SVM classifier was trained with β-estimates from seven of the eight runs for a given ROI to predict scene category and tested with β-estimates from the left-out run. In two within-type conditions, the classifier was trained and tested on evoked responses from the same image type ($CP_{Train}CP_{Test}$ and $LD_{Train}LD_{Test}$). Alternatively, it was trained on one image type and tested on the other in two between-types conditions ($CP_{Train}LD_{Test}$ and $LD_{Train}CP_{Test}$). Classification rates across participants were compared to chance accuracy (1/6) using one-tailed *t*-tests.

3. Results

3.1. Behavioral

All subjects exhibited high performance on the 2-AFC natural versus manmade task (mean accuracy = 0.94, SD = 0.077), indicating that they attended to the stimuli throughout the experiment.

3.2. Repetition suppression

We only found RS for within-type repetitions (CP_1CP_2 and LD_1LD_2, Fig. 2**A**): CP repetition evoked significant RS in the PPA, and marginal effects were seen in RSC, and OPA (see Table 1 for all RS results and statistics). Repeating the same LD produced significant RS in PPA, OPA, and FFA. We found no RS for the two between-types conditions (CP_1LD_2 and LD_1CP_2, all $p > 0.103$). However, we observed significant ($p < 0.05$) repetition enhancement in V1 for the LD_1CP_2 condition.

Within each ROI, we compared RS across conditions using a two-way, repeated-measures ANOVA, with first image type (CP or LD) and repetition type (within-type or between-types) as factors. In all scene-selective ROIs, we found a significant main effect for repetition type: PPA: $F_{(1,14)} = 23.75$, $p = 0.000246$, RSC: $F_{(1,11)} = 7.33$, $p = 0.0203$, OPA: $F_{(1,14)} = 16.23$, $p = 0.00125$. A main effect for repetition type was also found in FFA, $F_{(1,14)} = 6.07$, $p = 0.0273$. Within-type RS was

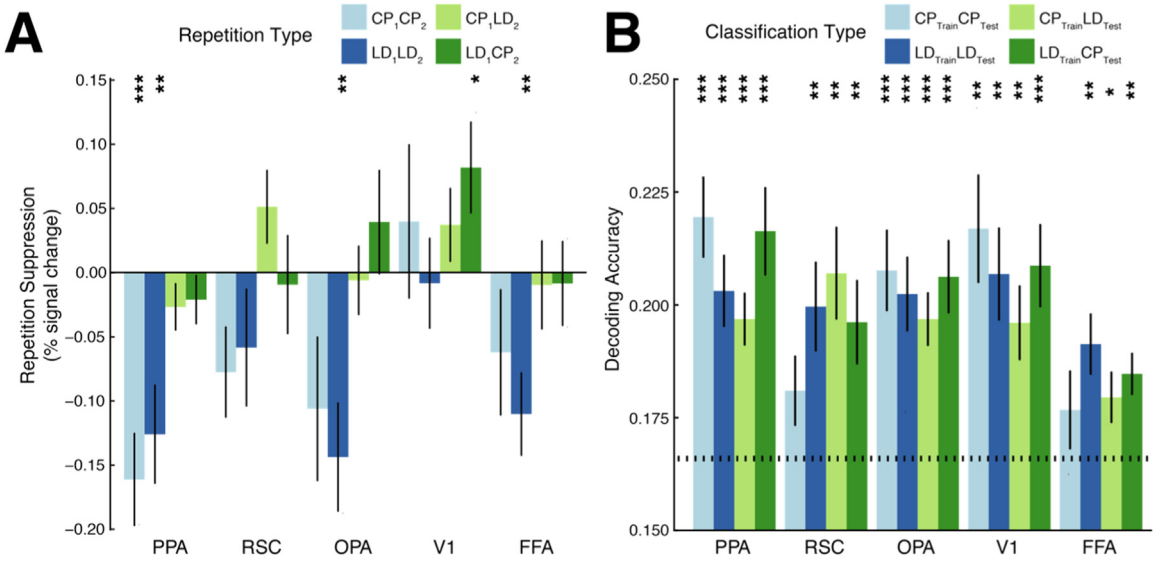

Fig. 2. A. Repetition suppression. Significant repetition suppression was seen for within-type repetitions (CP_1CP_2 & LD_1LD_2) but not between-types repetitions (CP_1LD_2 & LD_1LD_2). This is consistent with color photographs and line drawings activating identical neural populations or non-identical but spatially clustered neural populations. **B.** Decoding scene categories. Significant decoding was observed for within-image-type decoding and between-image-type cross-decoding. Along with repetition suppression just for within-type repetitions, these results are consistent with color photographs and line drawings activating non-identical but spatially clustered neural populations. ***p < 0.001, **p < 0.01, *p < 0.05.

Table 1
All results and statistics for the repetition suppression analyses.

ROI	df	Condition	% signal change	SEM	T-Stat	P-value
PPA	14	CP_1CP_2	− 0.163	0.037	− 4.44	0.0006
		LD_1LD_2	− 0.127	0.039	− 3.27	0.0056
		CP_1LD_2	− 0.027	0.019	− 1.46	0.1668
		LD_1CP_2	− 0.021	0.019	− 1.10	0.2881
RSC	11	CP_1CP_2	− 0.079	0.036	− 2.19	0.0512
		LD_1LD_2	− 0.059	0.046	− 1.28	0.2285
		CP_1LD_2	0.052	0.029	1.78	0.1031
		LD_1CP_2	− 0.010	0.039	− 0.25	0.8101
OPA	14	CP_1CP_2	− 0.107	0.057	− 1.89	0.0798
		LD_1LD_2	− 0.145	0.043	− 3.39	0.0044
		CP_1LD_2	− 0.006	0.027	− 0.23	0.8213
		LD_1CP_2	0.040	0.041	0.96	0.3510
V1	13	CP_1CP_2	0.040	0.061	0.66	0.5194
		LD_1LD_2	− 0.009	0.036	− 0.24	0.8148
		CP_1LD_2	0.037	0.029	1.29	0.2182
		LD_1CP_2	0.083	0.036	2.29	0.0394
FFA	14	CP_1CP_2	− 0.063	0.050	− 1.27	0.2260
		LD_1LD_2	− 0.112	0.033	− 3.40	0.0043
		CP_1LD_2	− 0.010	0.035	− 0.28	0.7827
		LD_1CP_2	− 0.009	0.033	− 0.26	0.7977

greater than between-types RS in each case. There were no main effects for first image type ($p > 0.362$) and no interactions between first image type and repetition type in any ROI ($p > 0.352$).

3.3. Whole brain repetition suppression

For the CP_1CP_2 condition (Fig. 3**A**), significant RS ($p < 0.01$, corrected) was present bilaterally in the parahippocampal gyri, which corresponds to our functional PPA ROIs. At a relaxed threshold ($p < 0.05$, corrected), we found RS bilaterally in the middle occipital gyrus, corresponding to our OPA ROI, and unilaterally in the left posterior cingulate cortex, corresponding to our right RSC ROI. For the LD_1LD_2 condition (Fig. 3**B**), significant RS ($p < 0.01$ corrected) was present in the right parahippocampal gyrus, which corresponds to the right PPA ROI. Additionally, the right middle occipital gyrus showed significant RS, which corresponds to our right OPA ROI. There was also a significant cluster of RS centered on the right inferior temporal gyrus and extending through the right middle temporal gyrus. Finally, at a relaxed statistical threshold ($p < 0.05$ corrected) we found bilateral RS in the inferior frontal gyrus. Neither the CP_1LD_2 nor the LD_1CP_2 contrasts yielded significant RS anywhere in the brain.

3.4. Scene category classification

In all of the ROIs (PPA, RSC, OPA, V1, and FFA), classification rates for discriminating between scene categories were significantly above chance (1/6) for all four conditions (Fig. 2**B**), with the exception of $CP_{Train}CP_{Test}$ in RSC and FFA (see Table 2 for all MVPA results and statistics), replicating previous results showing that scene category could be reliably decoded between CPs and LDs (Walther et al., 2011).

Classification rates were compared across conditions using a two-way, repeated-measures ANOVA, with training image type (CP or LD) and testing image type (within-type or between-types) as factors. No significant main effects or interactions were found in PPA or RSC: scene category was equally decodable regardless of which image type was used for training or testing in both ROIs, again replicating previous results showing scene category can be cross decoded between CPs and LDs in the PPA and the RSC (Walther et al., 2011). Additionally, no significant main effects or interactions were found in V1, OPA, or FFA.

4. Discussion

We used RS and MVPA to investigate the similarity of neural representations of CPs and LDs in scene-selective cortical areas. Using an event-related design, we replicated our previous findings (Walther et al., 2011; Choo and Walther, 2016) showing that scene categories can be decoded and cross-decoded from CPs and LDs of natural scenes across scene-selective cortex, suggesting similarities in the underlying representations. However, we found RS only for within-type repetitions in PPA, RSC, and OPA and not for between-types repetitions anywhere in the brain, suggesting representational differences between CPs and LDs of natural scenes. How can we reconcile these seemingly contradictory results provided by MVPA and RS?

Hatfield et al. (2016) presented a clear framework for interpreting

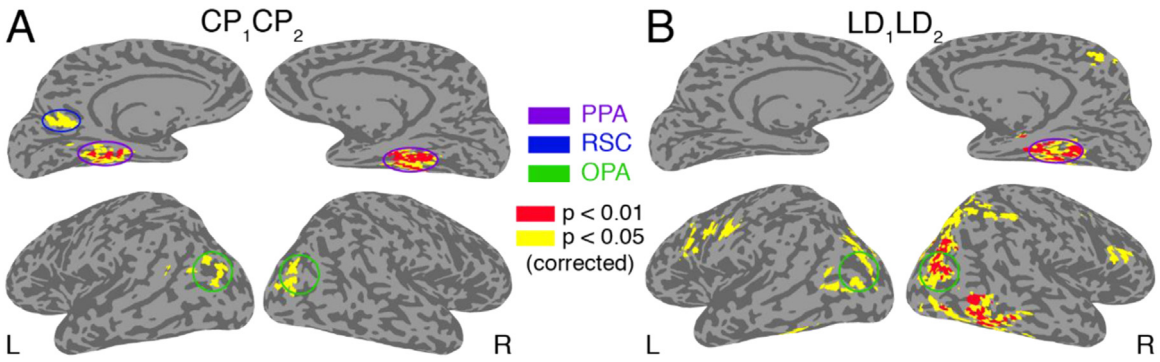

Fig. 3. Whole brain repetition suppression (RS) results. **A.)** RS for within-type repetition of color photographs. **B.)** RS for within-type repetition of line drawings. No between-types RS was observed anywhere in the brain.

Table 2
All results and statistics for the MVPA classification analyses.

ROI	df	Condition	Accuracy	SEM	T-Stat	P-value
PPA	14	$CP_{Train}CP_{Test}$	21.94%	0.89%	5.90	0.00004
		$LD_{Train}LD_{Test}$	20.31%	0.80%	4.57	0.00043
		$CP_{Train}LD_{Test}$	19.69%	0.58%	5.17	0.00014
		$LD_{Train}LD_{Test}$	21.63%	0.98%	5.09	0.00017
RSC	11	$CP_{Train}CP_{Test}$	18.10%	0.77%	1.85	0.09162
		$LD_{Train}LD_{Test}$	19.97%	0.99%	3.33	0.00676
		$CP_{Train}LD_{Test}$	20.70%	1.02%	3.94	0.00230
		$LD_{Train}LD_{Test}$	19.62%	0.93%	3.17	0.00897
OPA	14	$CP_{Train}CP_{Test}$	20.76%	0.90%	4.56	0.00044
		$LD_{Train}LD_{Test}$	20.24%	0.82%	4.34	0.00068
		$CP_{Train}LD_{Test}$	19.69%	0.59%	5.09	0.00016
		$LD_{Train}LD_{Test}$	20.63%	0.81%	4.90	0.00024
V1	13	$CP_{Train}CP_{Test}$	21.69%	1.20%	4.19	0.00106
		$LD_{Train}LD_{Test}$	20.68%	1.03%	3.91	0.00178
		$CP_{Train}LD_{Test}$	19.61%	0.83%	3.55	0.00354
		$LD_{Train}LD_{Test}$	20.87%	0.92%	4.59	0.00051
FFA	14	$CP_{Train}CP_{Test}$	17.67%	0.87%	1.15	0.26764
		$LD_{Train}LD_{Test}$	19.13%	0.68%	3.64	0.00267
		$CP_{Train}LD_{Test}$	17.95%	0.57%	2.26	0.04014
		$LD_{Train}LD_{Test}$	18.47%	0.47%	3.88	0.00167

the presence/absence of MVPA and RS effects in terms of their underlying neural signals. The presence of significant similarity in multi-voxel patterns (MVPs) could reflect repeated activation of the same population of neurons across both presentations of a stimulus or activation of different populations of neurons reliably clustered at a spatial scale smaller than a voxel (similar to hypercolumns). On the other hand, the presence of significant RS is thought to reflect repeated activation of the same population of neurons across both presentations of a stimulus (Grill-Spector and Malach, 2001; Grill-Spector et al., 2006). Even though there are several competing accounts for the physiological basis of RS (Grill-Spector et al., 2006; Malach, 2012), the conflicting theories all predict activation of the same neural populations by the initial and repeated stimulus presentations. Taken together, the presence or absence of MVPA and RS effects can be used to infer the nature of neural populations that represent different classes of stimuli (such as CPs and LDs, Fig. 1**B**). The presence of both MVPA and RS effects between stimulus classes would suggest that identical neural populations represent both classes of stimuli. Presence of MVPA but not RS would suggest that reliably clustered but distinct neural populations represent each class of stimuli.

Here, we find that scene categories can be decoded and cross-decoded from MVPs elicited by CPs and LDs of natural scenes. When interpreted using the above framework, these results suggest that CPs and LDs are either represented by the same neural populations or represented by populations of neurons reliably clustered at a sub-voxel level. However, we do not find RS between CPs and LDs of natural scenes. The presence of RS for within-type repetitions confirms that we reliably activate the same neural populations by repeating the exact same image within our experimental paradigm. The absence of RS for between-type repetitions indicates that identical neural populations are not activated by CPs and LDs of the same scenes. Taken together, the MVPA and RS results provide evidence that CPs and LDs are represented by reliably clustered, but not identical, neural populations in scene-selective cortical areas (Fig. 1**B**).

Another potential explanation for the divergent results from MVPA and RS is that the two analyses access different levels of representation in the brain. For natural scenes, Epstein and Morgan (2012) found that both landmark identity and scene category are decodable with MVPA, but RS is only present for landmark identity repetition and not scene category repetition. In our case, LDs preserve many structural components of CPs, including the distribution of curvature and contour junction features across the visual field, while eliminating color and most texture. These structural components have been shown to be computationally predictive of behavioral (Walther and Shen, 2014) and neural categorization error patterns for natural scenes (Choo and Walther, 2016). However, we do not find RS between CPs and LDs of the same scene. One explanation for this finding is that the structural components of a scene preserved by LDs are sufficient to support representation of scene category but not scene identity. Other features present in CPs but not preserved in LDs, such as color and fine-grained texture, may play a more important role in scene identification despite not being necessary for successful scene categorization. If this account is correct and RS and MVPA are differentially accessing representations supporting identification and categorization, our results would suggest that representation of scene identity requires a more complete match in image properties, such as color and fine-grained texture, than is afforded by the scene structure preserved in LDs. Additional behavioral and computational modeling experiments are necessary to confirm this hypothesis.

Between-types RS was not observed in the PPA. The PPA has been shown to represent many visual attributes of scenes (see Malcolm et al., 2016 for a review), including spatial layout (Epstein and Kanwisher, 1998; Epstein et al., 2007; Park et al., 2007, 2011; Epstein, 2008; Kravitz et al., 2011), relative distance and expanse (Kravitz et al., 2011; Harel et al., 2013), size and clutter (Park et al., 2014), contextually-relevant objects (MacEvoy and Epstein, 2011), texture (Lowe et al., 2017; Park and Park, 2017), and spatial frequency content (Rajimehr et al., 2011; Kauffmann et al., 2015; Watson et al., 2016; Berman et al., 2017). CPs and LDs of scenes differ in many of these features. The fact that we can decode and cross-decode scene categories from brain activity patterns evoked by CPs and LDs in the PPA does not indicate that no other information is being represented within the region; decoding analysis does not allow for a complete functional description of a region (Naselaris et al., 2011). In fact, several aspects of a scene including its

spatial, textural, object, or spatial frequency content are likely to be represented by populations of neurons in PPA and to contribute to the representation of scene content (such as category). Such coexistence of overlapping representations for various attributes of natural scenes within the same brain regions is consistent with neurophysiology findings in macaques (Kornblith et al., 2013) and helps explain the wide variety of scene attributes decodable from the PPA.

There are several alternative interpretations for our results. It is possible that format-specific image properties, such as color, are coded by populations of neurons confined within individual voxels and that format-general features, such as the local or global scene structure, are only encoded in populations of neurons spanning several voxels. This scenario would also explain why MVPA generalizes across CPs and LDs but RS does not. While we cannot exclude this explanation with our current data, we favor the explanation based on separate but spatially intermixed neural populations for encoding CPs and LDs of scenes, because it depends less on assumptions about a specific spatial distribution of neurons across voxels and, therefore specific choices of the spatial grid for data acquisition.

Additionally, the magnitude of RS is modulated by different psychological manipulations, including temporal context (Turk-Browne et al., 2012) and selective attention (Yi and Chun, 2005; Moore et al., 2013), and RS has been linked to explicit memory (Turk-Browne et al., 2006; but see Ward et al., 2013) and implicit memory (Maccotta and Buckner, 2004; Wig et al., 2005; Zago et al., 2005; Turk-Browne et al., 2006; Ward et al., 2013). However, the current experiment was not designed to test for these effects. In fact, we took great care to counterbalance any experimental parameters that could potentially affect attention and memory, such as the temporal separation of repeated presentations in time and the temporal position of stimuli within blocks and within the experiment session, across repetition conditions.

Another explanation could be that participants do not remember repeated instances of the same scene when the image type changes across presentations. It is possible that the natural/manmade response is bound to the memory of the image, and the magnitude of repetition suppression reflects stimulus-response learning between the scene and the response, which has been linked to RS (Dobbins et al., 2004). The observed lack of RS for between-types repetitions between CPs and LDs could reflect disrupted stimulus-response learning across image types. Future work can systematically explore this possibility by manipulating and tracking attention and memory for CPs and LDs behaviorally and linking behavioral performance to RS and MVPs.

Why do we not see RS effects in V1 for the ROI-based analysis or anywhere in early visual cortex for the whole brain analysis? Across the literature, there are two primary designs used to evoke RS: repetition within a trial (short-interval RS) and repetition between trials (long-interval RS). Epstein et al. (2008) systematically explored the difference between these two types of RS within the same set of participants. For natural scenes, they found short-interval RS in both scene selective and early visual cortex but only found long-interval RS in scene selective cortex. We here employed a long-interval RS design with a lag of 19.5–45.5 s between repetitions. Thus, the lack of RS effects in V1 and early visual areas observed in this study is consistent with previous reports.

To summarize, we have confirmed the representational equivalence of CPs and LDs of real-world scenes with respect to MVPA decoding of scene categories. However, the structural features preserved in LDs were insufficient to give rise to RS for stimulus repetitions across image type. Taken together, these results lead us to conclude that CPs and LDs of natural scenes are represented in scene-selective cortical areas by distinct neural populations that are reliably spatially clustered at a sub-voxel level. MVPA and RS are complementary analysis tools when investigating neural representations of real-world scenes. Jointly, they provide us with new insights into the organization of visual information in scene-selective visual cortex.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2015-06696 to DBW) and the Canadian Foundation for Innovation (Grant 32896 to DBW).

References

Bar, M., 2004. Visual objects in context. Nat. Rev. Neurosci. 5, 617–629.
Berman, D., Golomb, J.D., Walther, D.B., 2017. Scene content is predominantly conveyed by high spatial frequencies in scene-selective visual cortex. PLoS One 12 (Lappe M, ed., e0189828–16).
Choo, H., Walther, D.B., 2016. Contour junctions underlie neural representations of scene categories in high-level human visual cortex. NeuroImage 135, 32–44.
Cox, D.D., Savoy, R.L., 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. NeuroImage 19, 261–270.
Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.
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.
Epstein, R.A., 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. Trends Cogn. Sci. 12, 388–396.
Epstein, R.A., Kanwisher, N.G., 1998. A cortical representation of the local visual environment. Nature 392, 598–601.
Epstein, R.A., Morgan, L.K., 2012. Neuropsychol. Neuropsychol. 50, 530–543.
Epstein, R.A., Parker, W.E., Feiler, A.M., 2007. Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. J. Neurosci. 27, 6141–6149.
Epstein, R.A., Parker, W.E., Feiler, A.M., 2008. Two kinds of fMRI repetition suppression? Evid. Dissociable Neural Mech. 99, 2877–2886.
Fang, F., Murray, S.O., Kerstent, D., He, S., 2005. Orientation-tuned fMRI adaptation in human visual cortex. J. Neurophysiol. 94, 4188–4195.
Fei-Fei, L., Iyer, A., Koch, C., Perona, P., 2007. What do we perceive in a glance of a real-world scene? J. Vision. 7, 10–29.
Greene, M.R., Oliva, A., 2009. The briefest of glances the time course of natural scene understanding. Psychol. Sci. 20, 464–472.
Grill-Spector, K., Henson, R.N.A., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn. Sci. 10, 14–23(Available at: 〈http://linkinghub.elsevier.com/retrieve/pii/S1364661305003232〉).
Grill-Spector, K., Malach, R., 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol. 107, 293–321.
Harel, A., Kravitz, D.J., Baker, C.I., 2013. Deconstructing visual scenes in cortex: gradients of object and spatial layout information. Cereb. Cortex 23, 947–957.
Hatfield, M., McCloskey, M., Park, S., 2016. Neural representation of object orientation: a dissociation between MVPA and repetition suppression. NeuroImage 139, 136–148.
Haxby, J.V., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293, 2425–2430.
Kanwisher, N.G., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.
Kauffmann, L., Ramanoël, S., Guyader, N., Chauvin, A., Peyrin, C., 2015. Spatial frequency processing in scene-selective cortical regions. NeuroImage 112, 86–95.
Kornblith, S., Cheng, X., Ohayon, S., Tsao, D.Y., 2013. A network for scene processing in the Macaque temporal lobe. Neuron 79, 766–781.
Kravitz, D.J., Peng, C.S., Baker, C.I., 2011. Real-world scene representations in high-level visual cortex: it's the spaces more than the places. J. Neurosci. 31, 7322–7333.
Larsson, J., Smith, A.T., 2012. fMRI Repetition Suppression: Neuronal Adaptation or Stimulus Expectation? Cerebral Cortex. 22, 567–576.
Lowe, M.X., Rajsic, J., Gallivan, J.P., Ferber, S., Cant, J.S., 2017. Neural representation of geometry and surface properties in object and scene perception. NeuroImage 157, 586–597.
Maccotta, L., Buckner, R.L., 2004. Evidence for neural effects of repetition that directly correlate with behavioral priming. J. Cogn. Neurosci. 1–8.
MacEvoy, S.P., Epstein, R.A., 2011. Constructing scenes from objects in human occipitotemporal cortex. Nat. Neurosci. 14, 1–9.
Maguire, E.A., 1998. Knowing where and getting there: a human navigation network. Science 280, 921–924.
Maguire, E.A., Frackowiak, R.S.J., Frith, C.D., 1997. Recalling routes around London: activation of the right hippocampus in taxi drivers. J. Neurosci. 17, 7103–7110.
Malcolm, G.L., Groen, I.I.A., Baker, C.I., 2016. Making sense of real-world scenes. Trends Cogn. Sci. 20, 843–856.
Moore, K.S., Yi, D.-J., Chun, M.M., 2013. The effect of attention on repetition suppression and multivoxel pattern similarity. J. Cogn. Neurosci. 25, 1305–1314.
Naselaris, T., Kay, K.N., Nishimoto, S., Gallant, J.L., 2011. Encoding and decoding in fMRI (Available at: 〈http://dx.doi.org/10.1016/j.neuroimage.2010.07.073〉). NeuroImage 56, 400–410.
Oliva, A., Torralba, A., 2007. The role of context in object recognition. Trends Cogn. Sci. 11, 520–527.
Park, J., Park, S., 2017. Conjoint representation of texture ensemble and location in the parahippocampal place area. J. Neurophysiol. 117, 1595–1607.
Park, S., Brady, T.F., Greene, M.R., Oliva, A., 2011. Disentangling scene content from

spatial boundary: complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. J. Neurosci. 31, 1333–1340.
Park, S., Intraub, H., Yi, D.-J., Widders, D., Chun, M.M., 2007. Beyond the edges of a view: boundary extension in human scene-selective visual cortex. Neuron 54, 335–342.
Park, S., Konkle, T., Oliva, A., 2014. Parametric coding of the size and clutter of natural scenes in the human brain. Cereb. Cortex 25, bht418–bht1805.
Rajimehr, R., Devaney, K.J., Bilenko, N.Y., Young, J.C., Tootell, R.B.H., 2011. The "parahippocampal place area" responds preferentially to high spatial frequencies in humans and monkeys. PLoS Biol. 9, e1000608–e1000611 (Whitney D, ed.).
Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K., Belliveau, J., Brady, T., Rosen, B.R., Tootell, R., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268, 889–893.
Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.M., Egner, T., 2008. Neural repetition suppression reflects fulfilled perceptual expectations. Nat. Neurosci. 11, 1004–1006.
Torralbo, A., Walther, D.B., Chai, B., Caddigan, E., Fei-Fei, L., Beck, D.M., 2013. Good exemplars of natural scene categories elicit clearer patterns than bad exemplars but not greater BOLD activity. PLoS One 8 (Op de Beeck H, ed., e58594–12).
Turk-Browne , N.B., Simon, M.G., Sederberg, P.B., 2012. Scene representations in parahippocampal cortex depend on temporal context. J. Neurosci. 32, 7202–7207(Available at: 〈http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0942-12.2012〉).
Turk-Browne, N.B., Yi, D.-J., Chun, M.M., 2006. Linking implicit and explicit memory: common encoding factors and shared representations. Neuron 49, 917–927.
Walther, D.B., Caddigan, E., Fei-Fei, L., Beck, D.M., 2009. Natural scene categories revealed in distributed patterns of activity in the human brain. J. Neurosci. 29, 10573–10581.
Walther, D.B., Chai, B., Caddigan, E., Beck, D.M., Fei-Fei, L., 2011. Simple line drawings suffice for functional MRI decoding of natural scene categories. Proc. Natl. Acad. Sci. USA 108, 9661–9666.
Walther, D.B., Shen, D., 2014. Nonaccidental properties underlie human categorization of complex natural scenes. Psychol. Sci. 25, 851–860.
Ward, E.J., Chun, M.M., Kuhl, B.A., 2013. Repetition suppression and multi-voxel pattern similarity differentially track implicit and explicit visual memory. J. Neurosci. 33, 14749–14757.
Watson, D.M., Hymers, M., Hartley, T., Andrews, T.J., 2016. Patterns of neural response in scene-selective regions of the human brain are affected by low-level manipulations of spatial frequency. NeuroImage 124, 107–117.
Wig, G.S., Grafton, S.T., Demos, K.E., Kelley, W.M., 2005. Reductions in neural activity underlie behavioral components of repetition priming. Nat. Neurosci. 8, 1228–1233.
Yi, D.-J., Chun, M.M., 2005. Attentional modulation of learning-related repetition attenuation effects in human parahippocampal cortex. J. Neurosci. 25, 3593–3600.
Zago, L., Fenske, M.J., Aminoff, E., Bar, M., 2005. The rise and fall of priming: how visual exposure shapes cortical representations of objects. Cereb. Cortex 15, 1655–1665.