J Neurophysiol 123: 224–233, 2020.
First published November 27, 2019; doi:10.1152/jn.00582.2019.

RESEARCH ARTICLE | *Sensory Processing*

Temporal tuning of repetition suppression across the visual cortex

Matthias Fritsche, Samuel J. D. Lawrence, and Floris P. de Lange
Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

Submitted 10 September 2019; accepted in final form 21 November 2019

Fritsche M, Lawrence SJ, de Lange FP. Temporal tuning of repetition suppression across the visual cortex. *J Neurophysiol* 123: 224–233, 2020. First published November 27, 2019; doi:10.1152/jn.00582.2019.—The visual system adapts to its recent history. A phenomenon related to this is repetition suppression (RS), a reduction in neural responses to repeated compared with nonrepeated visual input. An intriguing hypothesis is that the timescale over which RS occurs across the visual hierarchy is tuned to the temporal statistics of visual input features, which change rapidly in low-level areas but are more stable in higher level areas. Here, we tested this hypothesis by studying the influence of the temporal lag between successive visual stimuli on RS throughout the visual system using functional (f)MRI. Twelve human volunteers engaged in four fMRI sessions in which we characterized the blood oxygen level-dependent response to pairs of repeated and nonrepeated natural images with interstimulus intervals (ISI) ranging from 50 to 1,000 ms to quantify the temporal tuning of RS along the posterior-anterior axis of the visual system. As expected, RS was maximal for short ISIs and decayed with increasing ISI. Crucially, however, and against our hypothesis, RS decayed at a similar rate in early and late visual areas. This finding challenges the prevailing view that the timescale of RS increases along the posterior-anterior axis of the visual system and suggests that RS is not tuned to temporal input regularities.

NEW & NOTEWORTHY Visual areas show reduced neural responses to repeated compared with nonrepeated visual input, a phenomenon termed repetition suppression (RS). Here we show that RS decays at a similar rate in low- and high-level visual areas, suggesting that the short-term decay of RS across the visual hierarchy is not tuned to temporal input regularities. This may limit the specificity with which the mechanisms underlying RS could optimize the processing of input features across the visual hierarchy.

adaptation; fMRI; repetition suppression; temporal tuning; visual processing

INTRODUCTION

How do our brains efficiently process the vast amount of information impinging on our senses? Importantly, sensory input is not random but exhibits spatial and temporal regularities (Dong and Atick 1995; Simoncelli and Olshausen 2001). For instance, our world is relatively stable over short timescales and our brains could exploit this temporal stability to optimize neural processing.

One phenomenon that has been linked to optimized processing of visual information over time is repetition suppression (RS) (Dragoi et al. 2000; Gotts et al. 2012; Grill-Spector et al. 2006; Kohn and Movshon 2004; Krekelberg et al. 2006; Müller et al. 1999). RS refers to the reduction of the neural response to repeated compared with novel sensory input and has been extensively studied and utilized in functional (f)MRI research (Grill-Spector et al. 1999; Henson et al. 2000; Kourtzi and Kanwisher 2001; Tootell et al. 1998; Vuilleumier et al. 2002). Crucially, if RS is a consequence of optimized neural processing due to the exploitation of temporal regularities of input features, it should be adapted to these specific regularities. That is, RS should be limited to timescales for which input features are likely repeated. Importantly however, visual features of different complexity exhibit different temporal regularities. Simple low-level features, such as contrast edges, represented in primary visual cortex (Hubel and Wiesel 1959), change rapidly within hundreds of milliseconds, for instance, due to eye movements. More complex features, such as faces, represented in downstream ventral areas (Kanwisher et al. 1997), are much more stably present in the input over short timescales. In line with this, it has been proposed that the temporal integration window of neural populations representing increasingly complex features progressively increases along the posterior-anterior axis of the brain (Hasson et al. 2008; Honey et al. 2012; Lerner et al. 2011). The current study was aimed at investigating whether a similar gradient exists for the temporal tuning of RS along the visual hierarchy as measured with fMRI. In particular, we hypothesized that RS should decay over shorter timescales in posterior compared anterior visual areas.

Some indirect evidence for a temporal tuning gradient in RS exists. For instance, previous fMRI studies using brief adaptor stimuli (≤1,000 ms) have observed RS in primary visual cortex (V1) only for short time lags of 100–200 ms between successively repeated input (Kourtzi and Huberle 2005) but not time lags of 400 ms or longer (Boynton and Finney 2003; Fang et al. 2005; Kourtzi and Huberle 2005; Murray et al. 2006; but see Weigelt et al. 2012). Conversely, RS in higher level object and face selective areas was observed over timescales of several seconds, while these studies generally failed to find RS in early visual areas over similar timescales (Sayres and Grill-Spector 2006; Weiner et al. 2010). Finally, a recent study reported increasing timescales for the recovery of subadditive blood oxygen level-dependent (BOLD) responses due to repeated visual input along the posterior-anterior axis of the visual system (Zhou et al. 2018).

While these previous findings point toward a gradient in temporal tuning, the evidence is synthesized from studies using

Address for reprint requests and other correspondence: M. Fritsche, Donders Institute for Brain, Cognition and Behaviour, Radboud Univ. Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, Netherlands (e-mail: m.fritsche@donders.ru.nl).

different stimuli, timings, and tasks and potentially conflating stimulus-specific with unspecific effects (Huettel and McCarthy 2001; Zhou et al. 2018). Therefore, in the current study we sought to systematically investigate the short-term temporal decay of RS across the visual hierarchy with one stimulus set consisting of natural images. To preview, we found strong RS to natural images in all visual areas under investigation. Surprisingly, even in V1 RS persisted for temporal lags of 1 s. Furthermore, when expressing RS in terms of a relative reduction of BOLD responses to repeat versus nonrepeat trials, RS gradually increased along the visual hierarchy, in line with previous findings (Soon et al. 2003; Weiner et al. 2010). Crucially, however, against our hypothesis, RS decayed at a similar rate in anterior and posterior visual areas. Therefore, the current results do not provide evidence for the hypothesis that RS reflects a mechanism of optimized neural processing due to the exploitation of temporal input regularities but are consistent with a more constant mechanism of short-term adaptation throughout the visual system.

MATERIALS AND METHODS

Data Availability

All data and code used for stimulus presentation and analysis are available from the Donders Institute for Brain, Cognition, and Behaviour repository at https://data.donders.ru.nl/collections/di/dccn/DSC_3018029.05_518.

Participants

Fourteen healthy participants (2 authors; 9 women, aged 23.4 ± 3.3 yr, means ± SD) took part in the study. All participants reported normal or corrected-to-normal vision and gave written, informed consent before the start of the study. The study was approved by the local ethical review board (CMO region, Arnhem-Nijmegen, The Netherlands) and was in accordance with the Declaration of Helsinki. Our a priori defined target sample size comprised 12 participants, for which we estimated a power of around 80 to 95% to detect RS effects in the fusiform face area (FFA) and lateral occipital complex (LOC) in response to repeated natural images with an interstimulus interval of 500 ms (Kovács et al. 2013). Participants completed three main task sessions and one session consisting of a retinotopic mapping and functional localizer task. All sessions were conducted on different days. Of the 14 participants, 1 participant became unavailable after the first 2 experiment sessions and 1 participant was excluded due to excessive head movements. Consequently, data of 12 participants entered data analysis.

Experimental Design

Stimuli and experimental paradigm. DISPLAY. Stimuli were displayed via an LCD projector (EIKI LC-XL100; resolution: 1,024 × 768 pixels; refresh rate: 60 Hz) onto a back-projection screen in the bore of the magnet. Participants viewed the screen (distance: 96.6 cm; field of view: horizontal 24°; vertical: 18°) through an angled mirror. Stimulus presentation was controlled using the Psychophysics toolbox (Brainard 1997; Pelli 1997) for MATLAB (The MathWorks, Natick, MA). For the three main task sessions, eye gaze was monitored and recorded with a high precision Eyelink 1000 eye tracker (SR Research).

MAIN TASK. An example trial of the main task is depicted in Fig. 1. Trials comprised two presentations of natural images (500 ms each), separated by one of five interstimulus intervals (50, 100, 200, 500, or 1,000 ms). On a given trial, either the same image could be presented twice (repeat trial) or two different images could be presented (nonrepeat trial). The experiment therefore implemented a 2 (repeat/nonrepeat) × 5 (ISIs) factorial design. Images were presented in color on a mid-gray background and were windowed with a circular aperture (8° radius) with a linearly ramping boundary (3.2° from 0 to 100%). Each trial was followed by a variable intertrial interval drawn from a truncated exponential distribution (minimum: 2,000 ms; average: 3,720 ms). To stabilize attention and to avoid potential confounds of attentional factors interacting with ISIs, participants were instructed to perform a demanding one-back digit task at fixation throughout each run (Kay et al. 2013a, 2013b; Zhou et al. 2018). Digits ranging from 0 to 9 (0.25 × 0.25° visual angle) were presented on top of a central, dark gray fixation disk (0.45° diameter). Each digit was on for 500 ms and off for 167 ms before the next digit appeared. Digit presentations were not synchronized to the presentation of the natural images. Participants were asked to press a button every time a digit repeated. On average there were 28 digit repetitions per run (~1 repetition per 30 s), and no more than one digit was repeated successively. To reduce visual adaptation, the color of digits alternated between black and white for successive presentations. Participants completed four runs (846 s each; 1,151 volumes, including 10 initial dummy volumes). Each run consisted of 160 trials and contained all ISI levels in pseudorandomized order. This resulted in 64 trials per repetition × ISI combination per session and thus 192 trials across all three main task sessions. To drive visual responses across multiple levels of the visual hierarchy, visual stimuli comprised a set of natural images. Images were taken from the ImageNet database (http://www.image-net.org/). We selected 32 images for each of the categories "Animal," "Body," "Face," "Fruit," "House," "Indoor," "Landscape," "Object," "Outdoor," and "Vehicle," resulting in a total of 320 images. For each participant and session, we created 160 randomly assigned pairs of images A and B from the full set of 320 images, under the constraint that A and B were of different categories. Image pairs were then randomly assigned to one of the five ISI conditions. Each image pair was presented once per run in one of the four possible orders (A-A, B-B, A-B, or B-A), and every order was presented once per session. Consequently, the overall visual stimulation across repeat and nonrepeat trials within each ISI condition was equal.

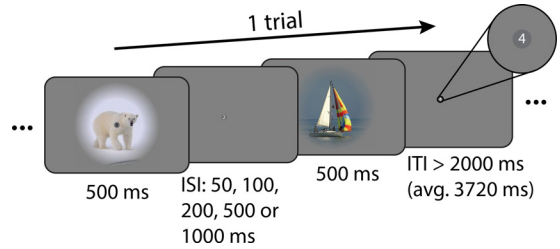

Fig. 1. Example trial of the main task. Each trial comprised 2 stimulus presentations of 500 ms each. Either the same image was presented twice (repeat trial) or 2 different images were presented (nonrepeat trial). Image presentations were separated by 1 of 5 interstimulus intervals (ISIs; 50, 100, 200, 500, or 1,000 ms). Participants were instructed to fixate on a small gray disk in the center of the screen (see *inset*) and performed a 1-back task on digits (0 to 9), which were shown within the fixation disk at a presentation rate of 1.5 Hz. Trials were separated by a variable intertrial interval (ITI), during which only the fixation disk and digits were shown.

FUNCTIONAL LOCALIZER. For mapping functional regions of interest (ROIs), participants viewed centrally presented grayscale images (16 × 16° visual angle) from the fLoc functional localizer package (Stigliani et al. 2015), while fixating on a central fixation dot (0.2° visual angle diameter). Stimuli were presented in a blocked design. Each block lasted for 16 s and comprised 32 individual images of the same category (2-Hz presentation rate). Each run (423 s; 575 volumes, including 10 initial dummy volumes) consisted of 26 blocks, of which the first and last blocks were always baseline blocks in which no

images were presented. The 24 interjacent blocks consisted of object, scene, face, body part, phase-scrambled, and baseline blocks with the same frequency (4 per each run). The order of blocks was pseudo-randomized such that adjacent blocks were always of a different category. Participants performed a one-back task on the images. Each participant completed three runs in total.

RETINOTOPY. For retinotopic mapping, participants viewed high contrast, contrast-reversing (6 Hz) 90° rotating wedge and expanding ring stimuli that mapped responses to the central region of the visual field (radius 9° of visual angle). Stimulus runs lasted 297 s (198 volumes), including eight full stimulus cycles and six starting dummy volumes that were discarded before analyses. During stimulus runs, participants were instructed to fixate on a small central fixation cross (0.24° visual angle across) and press a button every time it changed color from red to green or green to red (30 switches per run). Participants completed six runs of wedge stimulus presentation and one run of ring stimulus presentation.

fMRI data acquisition. Functional and anatomical images were collected on a 3T Prisma Fit MRI system (Siemens), using a 32-channel headcoil. Functional images in the main task and functional localizer were acquired using a whole-brain T2*-weighted multiband-6 sequence [repetition time (TR)/echo time (TE) = 735/39 ms, 48 slices, voxel-size 2.4 mm isotropic, 52° flip angle, A/P phase encoding direction]. Functional images in the retinotopy task were acquired using a whole-brain T2*-weighted multiband-4 sequence (TR/TE = 1,500/39.6 ms, 68 slices, voxel-size 2-mm isotropic, 75° flip angle, A/P phase encoding direction). One high-resolution anatomical image was acquired with a T1-weighted MP-RAGE sequence (TR/TE = 2300/3.03 ms, voxel-size 1-mm isotropic, 8° flip angle, GRAPPA acceleration factor 2).

Data Analysis

fMRI data preprocessing. fMRI data of the main and functional localizer tasks were preprocessed using FSL 6.00 (FMRIB Software Library; Oxford, UK; https://www.fmrib.ox.ac.uk/fsl; Smith et al. 2004). The preprocessing pipeline encompassed brain extraction (BET), motion correction (MCFLIRT), temporal high-pass filtering (100 s cutoff), spatial smoothing with a Gaussian kernel (full width at half maximum of 5 mm), and low-pass filtering using a Savitzky-Golay filter with window length $m = 7$ TRs and polynomial order $n = 3$ (Savitzky and Golay 1964). The first 10 volumes of each run were discarded to allow for signal stabilization.

Main task analysis. To estimate BOLD activations to repeat and nonrepeat trials at different ISIs, we employed finite impulse response (FIR) models (Ollinger et al. 2001). Briefly, we defined FIR basis functions of unit impulse from 0 to 16 s after trial onset, in steps of 1 s for each of the 10 trial types (5 ISIs × repeat/nonrepeat). This resulted in a total of 170 regressors (17 time points × 10 trial types), which were used to fit voxelwise generalized linear models (GLMs) to each participant's run data using FSL FEAT. Subsequently, the estimated BOLD responses for each trial type were averaged across runs and across all voxels within a given ROI (see *Functional ROI definition* and *Retinotopic ROI definition*) and fit with a double gamma function of the form

$$h(t) = a\left(\frac{t^{\alpha_1-1}\beta_1{}^{\alpha_1}e^{-\beta_1 t}}{\Gamma(\alpha_1)} - c\frac{t^{\alpha_2-1}\beta_2{}^{\alpha_2}e^{-\beta_2 t}}{\Gamma(\alpha_2)}\right),$$

where t designates time from trial onset; a controls the amplitude; α and β control the shape and scale, respectively; and c determines the ratio of the response to undershoot (Lindquist et al. 2009). The peaks of the positive response and late undershoot were constrained to occur between 3 and 8 s, and 8 and 18 s, respectively ($3 \le \frac{\alpha_1}{\beta_1} \le 8$ and $8 \le \frac{\alpha_2}{\beta_2} \le 18$). Given the current experimental design, this approach of estimating the shape of the hemodynamic response function (HRF) from the data appears more applicable than a standard GLM analysis assuming a single canonical HRF for all trial types and brain areas. First, due to the different ISIs across trials, the onset and peak latencies of the BOLD response are expected to differ slightly across ISI conditions (Supplemental Fig. S1; see https://figshare.com/articles/Supplemental_Figure_S1/10043114). Second, it is likely that HRFs will differ in shape across regions. Both aspects could lead to biased activation estimates across ISIs and ROIs in a standard GLM. From the estimated models, we derived the peak amplitudes A for each trial type (Fig. 2). Subsequently, we quantified RS within each ROI at every ISI as

$$RS = \frac{A_{nonrep} - A_{rep}}{A_{nonrep}}$$

where A_{rep} and A_{nonrep} reflect the peak BOLD amplitudes to repeat and nonrepeat trials, respectively. In other words, the *RS* score denotes the percentage reduction of BOLD amplitude for repeat relative to nonrepeat trials. Of note, the *RS* score captures the BOLD amplitude reduction to the compound BOLD response of the first and second image of each trial and therefore does not directly reflect the response reduction to the second repeated or nonrepeated image alone. Since the BOLD response to the first image will be similar for repeat and nonrepeat trials, the actual response reduction to the second image will be larger than indicated by the *RS* score, which factors in responses to both images. In fact, for the ISIs employed in this experiment, the actual response reduction to the second image will be approximately double the size of the estimated *RS* score. That is, a *RS* score of 25% corresponds to a response reduction of ~50% to the second repeated versus nonrepeated image. Importantly, simulations reveal that the relationship between *RS* scores and the corresponding response reduction to the second image are preserved across the different ISIs employed in this experiment, allowing an unbiased comparison RS across the different ISIs (Supplemental Fig. S1: see https://figshare.com/articles/Supplemental_Figure_S1/10043114). To estimate the decay of RS with increasing ISI, we fit linear models to each participant's *RS* scores as a function of ISI, such that the slope of the model described the change of *RS* scores over time. Intercepts and slopes of these linear models were statistically evaluated with

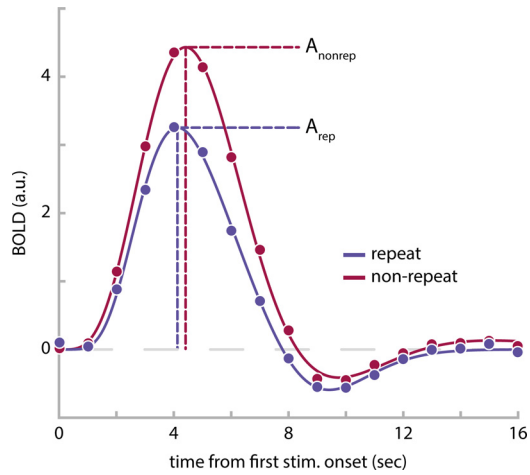

Fig. 2. Estimated blood oxygen level-dependent (BOLD) responses in the parahippocampal place area (PPA) of 1 example observer for trials with an interstimulus interval of 50 ms. The BOLD response for repeat trials (blue data points) is reduced in comparison to nonrepeat trials (red data points). BOLD amplitudes, A_{rep} and A_{nonrep}, were estimated by fitting double-gamma functions (solid lines) to the BOLD time courses and extracting the peak height of the double gamma function, respectively; a.u., arbitrary units. See Supplemental Fig. S2 (https://figshare.com/articles/Supplemental_Figure_S2/10043135) for group averaged BOLD responses and their respective double-gamma model fits for each region of interest and interstimulus interval combination.

separate repeated measures ANOVAs. We applied Greenhouse-Geisser correction, when a violation of sphericity was indicated. We hypothesized to find reduced BOLD amplitudes for repeat versus nonrepeat trials, leading to positive model intercepts. Furthermore, we expected RS to decay with increasing temporal lag between stimulus repetitions, which should be reflected in overall negative slopes. To assess the hypothesis of a faster decay of RS for posterior compared with anterior visual areas we sought to compare the decay of RS across areas independently of the initial magnitude of RS in each area. To this end, we first normalized each participant's slope estimate of a given ROI by its corresponding group mean intercept estimate. Consequently, the normalized slope parameter expressed the decay of RS relative to the initial magnitude of RS within a visual area and thus accounts for differences in slope due to differences in the initial magnitude of RS across visual areas. The normalized slope estimates were statistically evaluated with a repeated measures ANOVA. We expected to find a faster decay of RS for posterior compared with anterior visual areas, which would be captured by steeper, more negative slopes for more posterior compared with anterior ROIs. To test the hypothesis of gradually steeper slopes for posterior compared with anterior ROIs more directly, we further computed Spearman rank correlations between the normalized RS slope estimates and the position of the ROIs in the visual hierarchy. While areas V1 to V3 were assigned to successively increasing levels in the visual hierarchy (Felleman and Van Essen 1991), we assigned areas V4, LO1, LO2, and LOC to the same hierarchical level, as there is no clearly established distinction in hierarchical level between ventral V4 and dorsal LO1/2 and as the retinotopically defined areas LO1/2 and functionally defined area LOC are partly overlapping (Larsson and Heeger 2006). Finally, areas FFA and parahippocampal place area (PPA) were assigned the highest level among the ROIs of the current experiment. The rank-order correlation assessed a monotonic relationship between an area's position in the visual hierarchy and its rate of decay of the RS effect. Rank-order correlations were computed for each individual participant, and the resulting correlation coefficients were tested against zero on the group-level using a two-sided Wilcoxon signed-rank test, with a significance threshold of $\alpha = 0.05$. We expected to find a positive correlation between the ROIs' position in the visual hierarchy and their normalized RS slope estimates, indicating a gradually slower decay of RS in higher level visual areas.

Furthermore, to assess whether expressing RS as a percentage reduction of BOLD amplitude for repeat relative to nonrepeat trials (i.e. normalizing to nonrepeat trials) impacted our conclusions, we repeated the above analysis but expressed RS as the absolute BOLD amplitude difference between repeat and nonrepeat trials

$$RS = A_{nonrep} - A_{rep}$$

This has the advantage that any stimulus-unspecific adaptation effects, such as subadditive BOLD responses to any stimuli presented in close succession will largely cancel out, as they will be similarly present in both repeat and nonrepeat trials. It comes at the disadvantage that overall differences in BOLD responses across visual areas, for instance, due to the particular stimulus set employed in the experiment, are not taken into account in this analysis. This may complicate the comparison of the overall magnitude of RS across areas but should not influence the comparison of the RS decay.

Functional ROI definition. To map face, place, and object selective brain areas, voxelwise GLMs were fit to each participant's run data using FSL FEAT. Separate regressors for object, scene, face, body part, phase-scrambled, and baseline blocks were added to the model. Blocks were modeled with a 16 s duration boxcar and convolved with a gamma hemodynamic response function. Moreover, nuisance regressors in the form of first-order temporal derivatives of all event types and 24 motion regressors (6 motion parameters, the derivatives of these motion parameters, the squares of the motion parameters, and the squares of the derivatives; comprising FSL's standard extended set of motion parameters) were added to the model. Data were combined across runs using FSL's fixed effect analysis. For identifying object-selective area LOC, we contrasted average activations to objects versus scrambled blocks. For identifying face-selective area FFA and place-selective area PPA, we contrasted activations of faces versus scene blocks and scene versus face blocks, respectively. Differential activations were visualized on an inflated cortical surface and initially thresholded at $P < 1e\text{-}4$ (uncorrected). When necessary, the thresholds were raised on a per participant and per ROI basis until clusters of activations were limited to a reasonable extent, comparable in size and location to previous studies (Epstein and Kanwisher 1998; Grill-Spector et al. 2001; Kanwisher and Yovel 2006; Sayres and Grill-Spector 2006).

Retinotopic ROI definition. Anatomical data were automatically segmented into white matter, gray matter, and cerebral spinal fluid using FreeSurfer (http://surfer.nmr.mgh.harvard.edu/) and manually corrected. Functional data were analyzed using the phase encoding approach in MrVista (http://web.stanford.edu/group/vista/cgi-bin/wiki/index.php/Software). Polar angle and eccentricity data were visualized on an inflated cortical surface and the boundaries of V1, V2, V3, V4, LO1, and LO2 were drawn manually using established criteria (Amano et al. 2009; Engel et al. 1994; Larsson and Heeger 2006; Sereno et al. 1995; Wandell et al. 2007; Winawer and Witthoft 2017).

RESULTS

Repeating visual stimuli led to a clear reduction of BOLD amplitude across all visual areas (Fig. 3*A*). Accordingly, linear model fits to *RS* scores of each individual *ROI* exhibited significantly positive intercepts, indicating robust RS [$F(1,11) = 92.5$, $P = 1e\text{-}6$]. Furthermore, RS gradually increased from posterior to anterior visual areas (Fig. 3*A*), supported by a significant effect of ROI on the intercept term of the linear models [Fig. 3*B*, $F(2.5,27.3) = 6.2$, $P = 0.004$]. This was further corroborated by a significant positive correlation between the ROI position in the visual hierarchy and the intercept of RS (mean rho = 0.63 ± 0.10 SE; Wilcoxon $W = 76$, $P = 0.004$). In fact, on the individual participant level, 8 out of 12 participants showed a significant positive correlation ($P < 0.05$). Moreover, in line with prior research (Henson et al. 2000, 2004), the strength of RS decreased with increasing temporal lag between repeated stimuli, as revealed by consistently negative model slopes [Fig. 3*C*, $F(1,11) = 11.6$, $P = 0.006$]. Crucially, however, our initial hypothesis that posterior visual areas would show a steeper decay of RS compared with anterior areas was not borne out in the data. First, a repeated-measures ANOVA revealed no significant effect of ROI on normalized slopes [Fig. 4, $F(2.7,29.3) = 0.33$, $P = 0.78$]. Second, a more sensitive rank-order correlation analysis revealed no significant correlation between the ROI position in the visual hierarchy and the normalized slope of the RS decay (mean rho = −0.21 ± 0.18 SE; Wilcoxon $W = 15$, $P = 0.23$). Furthermore, a Bayesian statistical analysis revealed moderate evidence for the null hypothesis of no correlation between ROI position in the visual hierarchy and the normalized slope of the RS decay, compared with our initial hypothesis of a positive correlation ($BF_{0+} = 6.02$). Instead, RS decay was relatively similar across regions, such that RS was reduced to approximately half of its initial magnitude in all areas after 1,000 ms (Fig. 4). Interestingly, this meant that RS was observed for a prolonged period of time even in the earliest regions of visual cortex. Against our expectations, V1 showed robust RS throughout all temporal lags, even for the longest lag of 1,000

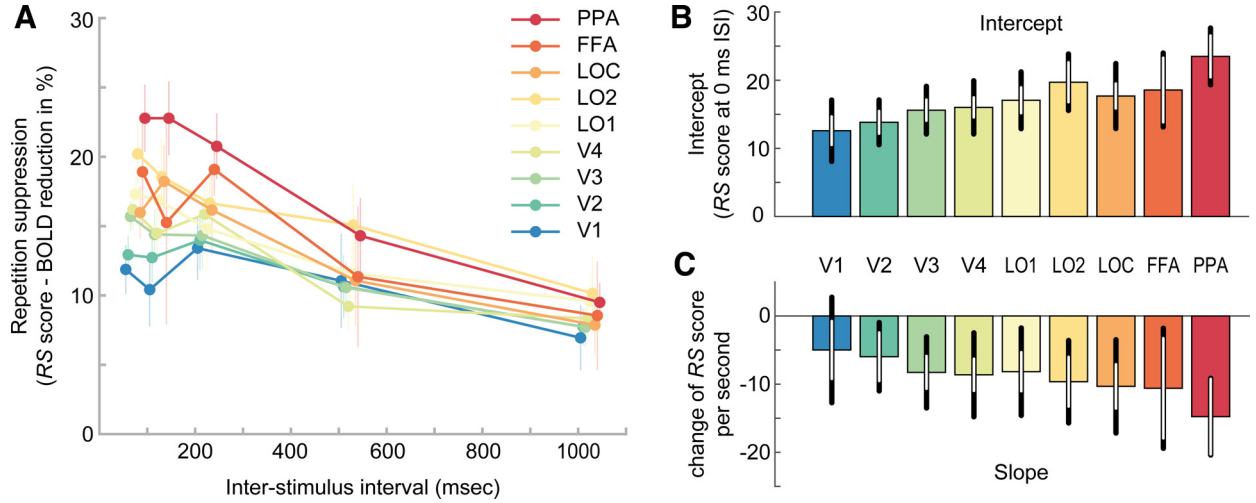

Fig. 3. Relative repetition suppression across visual areas and interstimulus intervals. *A*: repetition suppression (*RS*) score of each region of interest (ROI; color coded) as a function of interstimulus interval (ISI). The *RS* score indicates the percentage reduction of blood oxygen level-dependent (BOLD) amplitude for repeat relative to nonrepeat trials. Positive *RS* scores mark RS. While all ROIs show RS, the strength of RS increases from posterior to anterior regions. Furthermore, RS decreases for increasing temporal lags between image presentations. Error bars denote SEs. *B*: mean intercept estimates of linear models fitted to single subject data in *A*. Intercept estimates increase from posterior to anterior regions. Black error bars denote 95% confidence intervals. White error bars denote 95% within-subject confidence intervals. *C*: mean slope estimates of linear models, indicating the decay of *RS* scores over time. FFA, fusiform face area; PPA, parahippocampal place area; V1–V4, visual areas 1–4; LOC, lateral occipital complex; LO1 and LO2, lateral occipital areas 1 and 2. See Supplemental Fig. S3 (https://figshare.com/articles/Supplemental_Figure_S3/10043183) for single subject estimates of intercepts and slopes, Supplemental Fig. S4 (https://figshare.com/articles/Supplemental_Figure_S4/10043192) for model fits, and Supplemental Fig. S5 (https://figshare.com/articles/Supplemental_Figure_S5/10043198) for residual plots.

ms [$t(11) = 2.9$, $P = 0.016$]. This stands in apparent contrast to several previous findings on short-term adaptation, which do not find stimulus-specific RS in early visual cortex for temporal lags exceeding a couple hundred of milliseconds (Boynton and Finney 2003; Kourtzi and Huberle 2005; Murray et al. 2006; Weiner et al. 2010).

When repeating the above analyses on the absolute BOLD amplitude difference between repeat and nonrepeat trials, not normalized to the BOLD amplitude of nonrepeat trials, we obtained largely similar results (Fig. 5). For these absolute RS effects, linear model intercepts were significantly positive [$F(1,11) = 131.85$, $P = 2e\text{-}7$], while slopes were significantly negative [$F(1,11) = 10.92$, $P = 0.007$], indicating a decay of RS with increasing temporal lag between repeated stimuli. Furthermore, there was a significant effect of ROI on the intercept of RS [$F(2.88, 31,63) = 10.47$, $P = 2e\text{-}5$]. However, in contrast to the analysis of relative *RS* scores, there was no

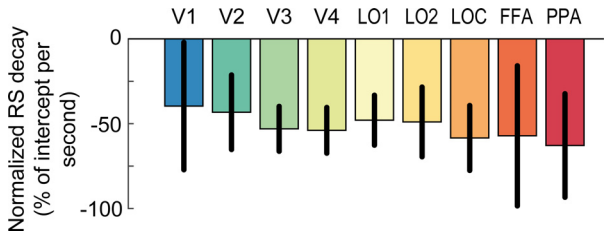

Fig. 4. Decay of repetition suppression normalized to its initial magnitude. The slope parameter estimates of the linear model fits to repetition suppression (*RS*) scores (see Fig. 3*C*) are normalized by the group mean intercept estimates for each region (see Fig. 3*B*), respectively. The normalized slope parameter estimate accounts for differences in slope due to differences in the initial magnitude of RS across visual areas. RS decays similarly across posterior and anterior visual areas and is reduced to approximately half of its initial magnitude after 1 s. Error bars denote 95% within-subject confidence intervals. FFA, fusiform face area; PPA, parahippocampal place area; V1–V4, visual areas 1–4; LOC, lateral occipital complex; LO1 and LO2, lateral occipital areas 1 and 2.

significant correlation between the ROI position in the visual hierarchy and the intercept of RS (mean rho = 0.20 ± 0.11 SE; Wilcoxon $W = 59$, $P = 0.13$). While RS monotonically increased from V1 to LO1, and was maximal for PPA, areas LO2, LOC, and FFA showed reduced RS effects (Fig. 5*B*). Those regions also showed the lowest overall BOLD activations of all ROIs (Supplemental Fig. S2; see https://figshare.com/articles/Supplemental_Figure_S2/10043135), potentially pointing toward a scaling of RS by the overall activation of the ROI. Importantly, in line with our analysis of relative RS effects, there was no evidence of a steeper decay of RS for posterior compared with anterior areas. In particular, there was no significant effect of ROI on the slope of RS decay [normalized to each ROI's intercept estimate, $F(2.37,26.04) = 0.695$, $P = 0.53$, Fig. 5*C*] and no significant correlation between the ROI position in the visual hierarchy and slope (normalized to each ROI's intercept estimate; mean rho = −0.20 ± 0.17 SE; Wilcoxon $W = 22$, $P = 0.20$). A Bayesian analysis revealed moderate evidence for the null hypothesis of no correlation between ROI position in the visual hierarchy and the normalized slope of the RS decay ($BF_{0+} = 6.64$). These results corroborate our finding that RS appears to decay at a similar rate across regions.

DISCUSSION

The aim of the current study was to investigate the temporal tuning of short-term RS to natural images across the visual system. Overall, we found strong reductions in BOLD responses for repeated compared with nonrepeated images across all visual areas under investigation. This RS effect gradually decayed with increasing temporal lag between stimuli. Crucially, however, against our initial hypothesis, there was no evidence for a faster decay of RS for low-level compared with high-level visual areas. Instead, RS in low- and high-level

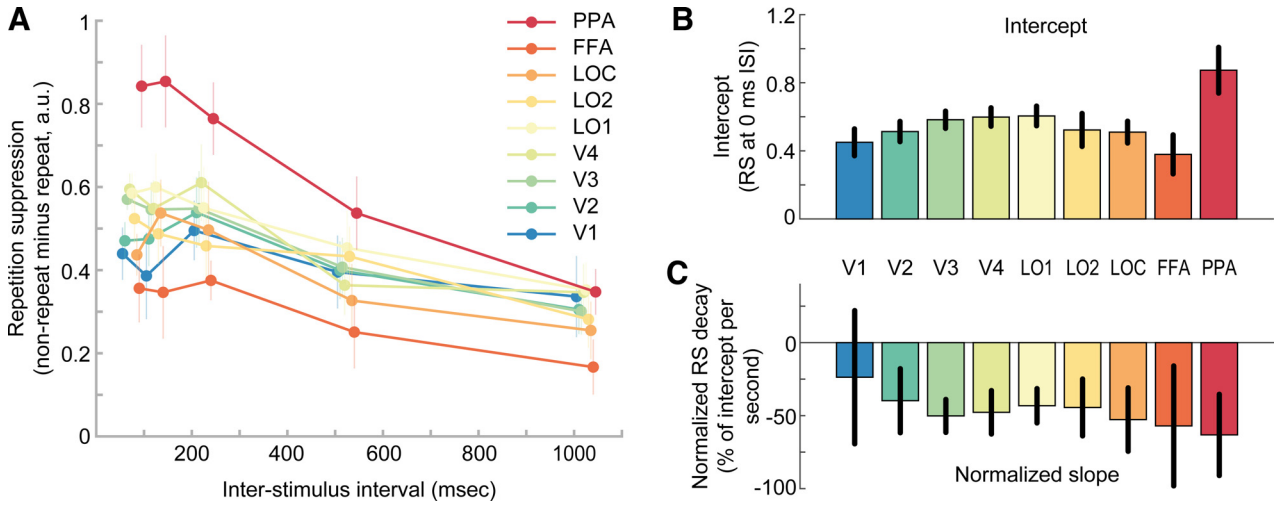

Fig. 5. Absolute repetition suppression across visual areas and interstimulus intervals. *A*: absolute repetition suppression [blood oxygen level-dependent (BOLD) amplitude difference between repeat and nonrepeat trials] of each region of interest (ROI) (color-coded) as a function of interstimulus interval (ISI). Positive values mark repetition suppression (RS); a.u., arbitrary units. Error bars denote SEs. See also Fig. 3*A* for relative RS effect. *B*: mean intercept estimates of linear models fitted to single subject data in *A*. Intercepts increase from visual area 1 (V1) to lateral occipital area 1 (LO1) and are maximal for the parahippocampal place area (PPA) but are reduced for LO2, lateral occipital complex (LOC), and fusiform face area (FFA). The latter 3 areas also show the lowest overall activation across all trials (Supplemental Fig. S2; see https://figshare.com/articles/Supplemental_Figure_S2/10043135). Error bars denote 95% within-subject confidence intervals. See also Fig. 3*B* for intercepts of relative RS effect. *C*: mean slope estimates of linear models normalized to the group mean intercept of each ROI, indicating the decay of RS over time. See also Fig. 4 for normalized slopes of relative RS effect. Error bars denote 95% within-subject confidence intervals.

visual areas appeared to decay at a similar rate. Therefore, the current results do not provide evidence for the hypothesis that short-term RS in the visual system is tuned to the temporal statistics of visual features of different complexity and challenge a prevailing view that RS generally persists over longer timescales in high-level cortical areas (Barron et al. 2016).

The finding that RS does not decay faster in posterior compared with anterior visual areas stands in apparent contrast to several previous findings, which collectively suggest such a differential decay. For instance, previous studies demonstrated RS in higher level object selective cortex over time-periods of several seconds, in the absence of similar effects in early visual areas (Sayres and Grill-Spector 2006; Weiner et al. 2010). Notably, RS in higher level occipitotemporal regions has been observed over time lags of minutes (Henson et al. 2000, 2004) and days (van Turennout et al. 2000, 2003). In contrast, RS in primary visual cortex is typically only observed for very short temporal lags of around 200 ms (Kourtzi and Huberle 2005; Kourtzi et al. 2010) but absent for longer lags (Boynton and Finney 2003; Fang et al. 2005; Kourtzi and Huberle 2005; Murray et al. 2006), although RS for lags of 1 s has been found in long-term adaptation designs with initial adaptors of 100-s duration and 4-s long top-up adaptors (Larsson et al. 2006; Montaser-Kouhsari et al. 2007). Our current findings differ from these previous reports in two aspects. First, in the current experiment RS in anterior object-selective cortex decreases rapidly over the tested time window, at a rate of about 50% of its initial magnitude per second. Extrapolating this to longer time windows, one would expect that RS should be completely absent for lags longer than ~2 s. Second, we find sustained RS to brief adaptors in early visual areas, such as V1, for temporal lags far longer than 200 ms. In fact, there is robust RS in V1, even at a temporal lag of 1 s.

What could account for the substantially faster decay of RS in anterior regions, compared with previous studies? In contrast to most previous studies on RS, which employed tasks that required participants to actively attend and process each stimulus, the current experiment involved a demanding task at fixation, which diverted attention away from the repeated or nonrepeated visual stimulation. Attention has been shown to modulate RS effects (Henson and Mouchlianitis 2007; Larsson and Smith 2012; Murray and Wojciulik 2004), and it is conceivable that it may change the temporal dynamics of RS across cortical areas. That is, attending and actively processing visual stimuli might lead to long-term synaptic changes in higher level cortical areas, which could support reductions of neural responses over longer timescales of minutes to days. Related to this, long-term RS involving several intervening stimuli between repetitions appears to be influenced by the precise task performed on the stimuli (Henson et al. 2002) and stimulus familiarity (Henson et al. 2000). Generally, it seems likely that long-term RS effects in category-selective areas reflect a combination of memory-dependent and task-dependent processes (see Henson 2016), which are very different from the process of short-term sensory adaptation studied in the current experiment. In turn, the current results indicate that, while short-term RS can be observed in category-selective cortex in the absence of attention toward the stimuli, it is characterized by a much more rapid temporal decay than suggested by previous findings.

Intriguingly, in contrast to previous studies, we also observed robust RS to brief adaptor stimuli in V1 for temporal lags of up to 1,000 ms. One clear difference between the current and previous studies is the stimulus set used to probe RS. While previous studies examining short-term adaptation in early visual cortex typically employed simple grating stimuli (Boynton and Finney 2003; Fang et al. 2005; Murray et al. 2006) or two-dimensionl shape stimuli consisting of locally oriented Gabors (Kourtzi and Huberle 2005), the current study used natural images to drive visual responses across multiple

levels of the visual hierarchy. Importantly, previous electrophysiological studies have shown differences in the oscillatory neural activity elicited by artificial grating stimuli versus naturalistic visual input in early visual areas, likely due to differences in both the temporal and spatial input statistics (Hermes et al. 2015; Kayser et al. 2003). In particular, in contrast to many natural images, grating stimuli elicit increased narrowband gamma power (Hermes et al. 2015), which has been linked to gain control or normalization (Hermes et al. 2019; Ray et al. 2013). We speculate that if these inhibitory normalization processes, which are strongly engaged by grating stimuli, are themselves subject to adaptation, this may explain the generally weaker and less sustained RS effects for grating stimuli in comparison with natural images. Interestingly, a recent study has described modulations of narrowband gamma responses by natural image structure, with more structured images eliciting stronger gamma responses (Brunet and Fries 2019). Following the above hypothesis, one may thus expect to find weaker adaptation to structured versus more random natural images in primary visual cortex. This could be an interesting topic for future research. While the above explanation would mainly rely on adaptation of inhibitory lateral connections within V1, there is an alternative explanation involving recurrent connections across cortical areas. Importantly, since BOLD responses reflect an integral of neural activity over several seconds, it is possible that the RS effect observed in V1 at a lag of 1,000 ms is due to feedback signals from higher level areas, which maintain RS over longer timescales. If natural images elicit more recurrent activity across the visual hierarchy, then grating stimuli, such an indirect RS effect due to feedback signals from high-level visual areas could explain the different lag dependence in V1. In this context, a recent study has investigated layer-dependent microcircuit dynamics during visual adaptation in V1 in alert monkeys (Westerberg et al. 2019). The authors reported robust adaptation to 200 and 500 ms grating stimuli with interstimulus intervals of 200 ms. Interestingly, stimulus repetition reduced synaptic processing most strongly in the supragranular layer of V1, which contains the bulk of cortico-cortical connections. Although this points toward intracortical processing underlying adaptation in V1, the findings are inconsistent with an explanation solely based on active inhibition via increased feedback from higher level areas, which would predict increased supragranular synaptic activity. While this study focused on grating stimuli, future studies may directly compare short-term RS effects for gratings and natural stimuli and assess potential effects of intracortical processing with electrophysiological methods and laminar fMRI.

In apparent contradiction to the current results, a recent study has found that early visual areas show a faster recovery from subadditive BOLD responses due to repeated visual input in comparison to downstream areas (Zhou et al. 2018). While BOLD responses to two stimuli presented in rapid succession were smaller than predicted by the sum of BOLD responses to an individual stimulus, the BOLD response in early visual cortex quickly approached the linear prediction for increasing temporal lags between stimulus presentations. Conversely, BOLD responses in more anterior visual areas remained below their linear prediction over progressively longer timescales. This adaptation effect and its temporal tuning across visual areas were adequately captured by a temporal compressive summation model. However, it should be noted that Zhou et al. (2018) studied adaptation using repeated presentations of the same visual stimulus, and therefore cannot distinguish between stimulus-specific and unspecific adaptation. Crucially, stimulus-unspecific adaptation effects have been widely observed in fMRI studies investigating the effects of stimulus repetition (Boynton and Finney 2003; Kourtzi and Huberle 2005; Murray et al. 2006; Soon et al. 2003) and are likely also present in the data by Zhou et al. (2018). Crucially, such unspecific adaptation effects may be driven by vascular mechanisms, unrelated to neural adaptation, and therefore, it is important to separate them from stimulus-specific adaptation. By contrasting BOLD responses of repeat and nonrepeat trials, for which the overall visual input was equated, we account for stimulus-unspecific adaptation effects in our analysis.

Furthermore, we have found that RS gradually increased in magnitude along the posterior-anterior axis of the visual system. It should be noted that this effect was only statistically significant, when expressing RS as a *relative* reduction of BOLD amplitude between repeat and nonrepeat trials. Conversely, there was no significant correlation between the ROI position in the visual hierarchy and the *absolute* difference of BOLD responses in repeat versus nonrepeat trials. While absolute RS monotonically increased from V1 to LO1, and was maximal for PPA, areas LO2, LOC, and FFA showed reduced absolute RS effects. Importantly, the latter three regions also showed the lowest overall BOLD activations of all ROIs throughout all trials (Supplemental Fig. S2; see https://figshare.com/articles/Supplemental_Figure_S2/10043135). This might suggest that RS scales with the overall activation of an ROI in response to the presented stimuli (Alink et al. 2018) and could explain why the gradual increase of RS was not robustly present when quantifying RS in terms of absolute BOLD amplitude differences, regardless of the areas' overall activation.

There exist several explanations for the gradual increase of relative RS magnitude along the posterior-anterior axis that we observed in the current study. First, it is possible that anterior regions exhibit intrinsically stronger RS than posterior regions. Second, RS effects in downstream cortical regions could, to some degree, be inherited from upstream regions (Kohn and Movshon 2003; Larsson and Harrison 2015). Such an inheritance effect would naturally produce an incremental increase of RS along the visual hierarchy. Third, the estimate of increasing RS along the visual hierarchy could be partly driven by cross-adaptation effects in nonrepeat trials. The activation in nonrepeat trials was used as a baseline for quantifying stimulus-specific response reductions when stimuli were repeated. However, since different natural images may exhibit some overlap in visual features, there likely occurs some feature-specific adaptation between the different stimuli in nonrepeat trials. This component of RS to shared features between different stimuli is not visible in our RS estimate. The magnitude of this cross-adaptation component in each visual area will depend on the area's feature selectivity and the amount of shared features between the employed stimuli. While this appears difficult to quantify, we estimated an approximate overlap of features with increasing complexity using a deep neural network, which has been shown to capture the gradient in the complexity of neural representations across the ventral visual stream (Güçlü and van Gerven 2015; Seeliger et al.

2018). We found that the stimulus pairs of nonrepeat trials were indeed more similar in terms of low-level features compared with high-level features in this artificial neural network (see Supplemental Fig. S6; see https://figshare.com/articles/Supplemental_Figure_S6/10339313). This hints at the possibility that cross adaptation may be more pronounced in early visual areas, selective for low-level features, compared with higher level visual areas. As a consequence, the overall feature-selective RS effect might be somewhat underestimated in low-level compared with high-level areas, potentially contributing to the gradual increase in RS along the visual hierarchy. Lastly, increased RS in area MT compared with V1 has been attributed to an interaction of receptive field and stimulus size, leading to different amounts of surround suppression across areas (Patterson et al. 2014; Wissig and Kohn 2012). However, these findings were based on drifting grating stimuli and only occurred for prolonged adaptation of 40 s but not short adaptation used in the current study.

Finally, we would like to point out several important limitations of the current study. First, our current results may depend on the stimulus set and experiment parameters such as adaptor and test stimulus timing employed in this study. As mentioned above, using grating stimuli instead of natural images may change the strength and temporal dynamics of RS across visual areas. For instance, different stimulus families, such as gratings versus natural images, may drive the neural circuitry, which ultimately gives rise to adaptation effects, in different ways. Furthermore, the duration of the adaptor stimulus will likely influence the magnitude and decay of RS and these effects could potentially differ across visual areas. For instance, while RS in V1 to brief adaptor gratings ($\leq$1 s) appears to be absent for temporal lags longer than 400 ms (Boynton and Finney 2003; Fang et al. 2005; Kourtzi and Huberle 2005; Murray et al. 2006), RS effects for a lag of 1 s have been found in long-term adaptation designs, where the adaptor grating is presented for 100 s (Larsson et al. 2006; Montaser-Kouhsari et al. 2007). In the current study, we were interested whether the timescale over which RS occurs across the visual hierarchy is tuned to the temporal statistics of visual input features. Therefore, we chose to focus on short-term adaptation to natural images, containing a wide range of input features of differing complexity and approximating natural viewing conditions. Extending the current results to other stimulus sets, stimulus timings, and tasks is an interesting topic for future research. A second limitation derives from the observation that adaptation effects in downstream cortical regions can be inherited from upstream regions (Kohn and Movshon 2003; Larsson and Harrison 2015). The potential presence of such inheritance effects may complicate the assessment of intrinsic temporal dynamics of RS within a particular visual area, independently of the dynamics in upstream areas. While feedforward inheritance effects can to some extent be circumvented by changing the spatial location of adaptor and test stimuli, thus exploiting the increase in receptive field size from low- to high-level visual areas (Larsson and Harrison 2015), this would preclude a direct comparison of adaptation in low- and high-level visual areas. Furthermore, inheritance effects might not be limited to feedforward processing, as adaptation effects in higher level areas could be passed down to lower level regions via recurrent connections. When interpreting the results of the current study, it should be noted, however, that if the temporal decay of RS in higher level regions would be mainly driven by the decay of RS in lower level regions, and higher level regions would add little to no RS themselves, such a pattern RS across the visual cortex would be incompatible with a view of optimized visual processing throughout the visual hierarchy, in which each area would exploit the temporal regularities of the features which it is tuned to. Third, caution must be taken when quantitatively inferring neural repetition suppression from the measured repetition suppression in the BOLD signal. While there is evidence for an approximately linear relationship between neural activity and the BOLD signal for brief stimuli (Arthurs et al. 2000; Boynton et al. 1996; Logothetis et al. 2001, 2003), this evidence is mainly based on studies investigating primary sensory areas and it is not entirely clear whether this similarly holds for higher level areas. A nonlinear relationship between neural activity and BOLD would complicate the quantitative inference from fMRI repetition suppression to neural repetition suppression across cortical areas. Furthermore, it is well known that stimuli presented in rapid succession will lead to a nonlinear summation of BOLD responses (Boynton and Finney 2003; Kourtzi and Huberle 2005; Murray et al. 2006; Soon et al. 2003), a stimulus-unspecific effect that might have a vascular, rather than purely neural, origin (de Zwart et al. 2009; Zhang et al. 2008). However, such a nonlinearity due to vascular mechanisms is likely stimulus-unspecific and therefore similarly present in repeat and nonrepeat trials. Therefore, by contrasting repeat with nonrepeat trials, nonlinear components of vascular origin should practically cancel out, thus accounting for stimulus-unspecific effects. Finally, our conclusion that RS in low- and high-level visual areas decays at a similar rate is based on a null result in a relatively small sample size ($n = 12$). We chose to collect an extensive amount of data per participant (4 sessions each), as previous studies indicated relatively consistent RS effects across individuals. We find that our data provide moderate evidence for the null hypothesis of no difference in RS decay across regions, compared with our initial hypothesis that RS should decay at a faster rate in low- compared with higher level visual areas. Nevertheless, one has to treat such a null result with caution and further work will be required to corroborate the current finding. To conclude, the current study revealed strong RS to natural images across the visual hierarchy, even in primary visual cortex, up to delays of 1,000 ms. Crucially, however, RS does not appear to decay more slowly in anterior compared with posterior visual areas. Thus the current results do not provide evidence for the hypothesis that short-term RS is tuned to the temporal statistics of visual features with different complexities. This suggests that RS is not a consequence of optimized neural processing due to the exploitation of temporal regularities of input features but appears to reflect an adaptation process that has a fixed time constant across the visual system. This may limit the specificity with which the mechanisms underlying RS could optimize the processing of input features at any level of the visual hierarchy.

GRANTS

This work was supported by European Union Horizon 2020 Program European Research Council Starting Grant 678286, "Contextvision."

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

M.F., S.J.D.L., and F.P.D.L. conceived and designed research; M.F. and S.J.D.L. performed experiments; M.F. analyzed data; M.F. and F.P.D.L. interpreted results of experiments; M.F. prepared figures; M.F. drafted manuscript; M.F., S.J.D.L., and F.P.D.L. edited and revised manuscript; M.F., S.J.D.L., and F.P.D.L. approved final version of manuscript.

ENDNOTE

At the request of the authors, readers are herein alerted to the fact that additional materials related to this manuscript may be found at the institutional website of one of the authors, which at the time of publication they indicate is: https://data.donders.ru.nl/collections/di/dccn/DSC_3018029.05_518. These materials are not a part of this manuscript, and have not undergone peer review by the American Physiological Society (APS). APS and the journal editors take no responsibility for these materials, for the website address, or for any links to or from it.

REFERENCES

Alink A, Abdulrahman H, Henson RN. Forward models demonstrate that repetition suppression is best modelled by local neural scaling. *Nat Commun* 9: 3854, 2018. doi:10.1038/s41467-018-05957-0.

Amano K, Wandell BA, Dumoulin SO. Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. *J Neurophysiol* 102: 2704–2718, 2009. doi:10.1152/jn.00102.2009.

Arthurs OJ, Williams EJ, Carpenter TA, Pickard JD, Boniface SJ. Linear coupling between functional magnetic resonance imaging and evoked potential amplitude in human somatosensory cortex. *Neuroscience* 101: 803–806, 2000. doi:10.1016/S0306-4522(00)00511-X.

Barron HC, Garvert MM, Behrens TEJ. Repetition suppression: a means to index neural representations using BOLD? *Philos Trans R Soc Lond B Biol Sci* 371: 20150355, 2016. doi:10.1098/rstb.2015.0355.

Boynton GM, Engel SA, Glover GH, Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16: 4207–4221, 1996. doi:10.1523/JNEUROSCI.16-13-04207.1996.

Boynton GM, Finney EM. Orientation-specific adaptation in human visual cortex. *J Neurosci* 23: 8781–8787, 2003. doi:10.1523/JNEUROSCI.23-25-08781.2003.

Brainard DH. The Psychophysics Toolbox. *Spat Vis* 10: 433–436, 1997. doi:10.1163/156856897X00357.

Brunet NM, Fries P. Human visual cortical gamma reflects natural image structure. *Neuroimage* 200: 635–643, 2019. doi:10.1016/j.neuroimage.2019.06.051.

de Zwart JA, van Gelderen P, Jansma JM, Fukunaga M, Bianciardi M, Duyn JH. Hemodynamic nonlinearities affect BOLD fMRI response timing and amplitude. *Neuroimage* 47: 1649–1658, 2009. doi:10.1016/j.neuroimage.2009.06.001.

Dong DW, Atick JJ. Statistics of natural time-varying images. *Network* 6: 345–358, 1995. doi:10.1088/0954-898X_6_3_003.

Dragoi V, Sharma J, Sur M. Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron* 28: 287–298, 2000. doi:10.1016/S0896-6273(00)00103-3.

Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN. fMRI of human visual cortex. *Nature* 369: 525, 1994. doi:10.1038/369525a0.

Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature* 392: 598–601, 1998. doi:10.1038/33402.

Fang F, Murray SO, Kersten D, He S. Orientation-tuned FMRI adaptation in human visual cortex. *J Neurophysiol* 94: 4188–4195, 2005. doi:10.1152/jn.00378.2005.

Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1: 1–47, 1991. doi:10.1093/cercor/1.1.1.

Gotts SJ, Chow CC, Martin A. Repetition priming and repetition suppression: A case for enhanced efficiency through neural synchronization. *Cogn Neurosci* 3: 227–237, 2012. doi:10.1080/17588928.2012.670617.

Grill-Spector K, Henson R, Martin A. Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10: 14–23, 2006. doi:10.1016/j.tics.2005.11.006.

Grill-Spector K, Kourtzi Z, Kanwisher N. The lateral occipital complex and its role in object recognition. *Vision Res* 41: 1409–1422, 2001. doi:10.1016/S0042-6989(01)00073-6.

Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, Malach R. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24: 187–203, 1999. doi:10.1016/S0896-6273(00)80832-6.

Güçlü U, van Gerven MA. Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *J Neurosci* 35: 10005–10014, 2015. doi:10.1523/JNEUROSCI.5023-14.2015.

Hasson U, Yang E, Vallines I, Heeger DJ, Rubin N. A hierarchy of temporal receptive windows in human cortex. *J Neurosci* 28: 2539–2550, 2008. doi:10.1523/JNEUROSCI.5487-07.2008.

Henson R, Shallice T, Dolan R. Neuroimaging evidence for dissociable forms of repetition priming. *Science* 287: 1269–1272, 2000. doi:10.1126/science.287.5456.1269.

Henson RN. Repetition suppression to faces in the fusiform face area: a personal and dynamic journey. *Cortex* 80: 174–184, 2016. doi:10.1016/j.cortex.2015.09.012.

Henson RN, Mouchlianitis E. Effect of spatial attention on stimulus-specific haemodynamic repetition effects. *Neuroimage* 35: 1317–1329, 2007. doi:10.1016/j.neuroimage.2007.01.019.

Henson RN, Rylands A, Ross E, Vuilleumeir P, Rugg MD. The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *Neuroimage* 21: 1674–1689, 2004. doi:10.1016/j.neuroimage.2003.12.020.

Henson RN, Shallice T, Gorno-Tempini ML, Dolan RJ. Face repetition effects in implicit and explicit memory tests as measured by fMRI. *Cereb Cortex* 12: 178–186, 2002. doi:10.1093/cercor/12.2.178.

Hermes D, Miller KJ, Wandell BA, Winawer J. Stimulus dependence of gamma oscillations in human visual cortex. *Cereb Cortex* 25: 2951–2959, 2015. doi:10.1093/cercor/bhu091.

Hermes D, Petridou N, Kay K, Winawer J. An image-computable model for the stimulus selectivity of gamma oscillations. *eLife* 8: e47035: 2019. doi:10.7554/eLife.47035.

Honey CJ, Thesen T, Donner TH, Silbert LJ, Carlson CE, Devinsky O, Doyle WK, Rubin N, Heeger DJ, Hasson U. Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* 76: 423–434, 2012. doi:10.1016/j.neuron.2012.08.011.

Hubel DH, Wiesel TN. Receptive fields of single neurones in the cat's striate cortex. *J Physiol* 148: 574–591, 1959. doi:10.1113/jphysiol.1959.sp006308.

Huettel SA, McCarthy G. Regional differences in the refractory period of the hemodynamic response: an event-related fMRI study. *Neuroimage* 14: 967–976, 2001. doi:10.1006/nimg.2001.0900.

Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17: 4302–4311, 1997. doi:10.1523/JNEUROSCI.17-11-04302.1997.

Kanwisher N, Yovel G. The fusiform face area: a cortical region specialized for the perception of faces. *Philos Trans R Soc Lond B Biol Sci* 361: 2109–2128, 2006. doi:10.1098/rstb.2006.1934.

Kay KN, Winawer J, Mezer A, Wandell BA. Compressive spatial summation in human visual cortex. *J Neurophysiol* 110: 481–494, 2013a. doi:10.1152/jn.00105.2013.

Kay KN, Winawer J, Rokem A, Mezer A, Wandell BA. A two-stage cascade model of bold responses in human visual cortex. *PLoS Comput Biol* 9: e1003079, 2013B. doi:10.1371/journal.pcbi.1003079.

Kayser C, Salazar RF, König P. Responses to natural scenes in cat V1. *J Neurophysiol* 90: 1910–1920, 2003. doi:10.1152/jn.00195.2003.

Kohn A, Movshon JA. Neuronal adaptation to visual motion in area MT of the macaque. *Neuron* 39: 681–691, 2003. doi:10.1016/S0896-6273(03)00438-0.

Kohn A, Movshon JA. Adaptation changes the direction tuning of macaque MT neurons. *Nat Neurosci* 7: 764–772, 2004. doi:10.1038/nn1267.

Kourtzi Z, Huberle E. Spatiotemporal characteristics of form analysis in the human visual cortex revealed by rapid event-related fMRI adaptation. *Neuroimage* 28: 440–452, 2005. doi:10.1016/j.neuroimage.2005.06.017.

Kourtzi Z, Kanwisher N. Representation of perceived object shape by the human lateral occipital complex. *Science* 293: 1506–1509, 2001. doi:10.1126/science.1061133.

Kourtzi Z, Tolias AS, Altmann CF, Augath M, Logothetis NK. Integration of local features into global shapes: monkey and human fMRI studies. *J Vis* 3: 191, 2010. doi:10.1167/3.9.191.

Kovács G, Kaiser D, Kaliukhovich DA, Vidnyánszky Z, Vogels R. Repetition probability does not affect fMRI repetition suppression for objects. *J Neurosci* 33: 9805–9812, 2013. doi:10.1523/JNEUROSCI.3423-12.2013.

Krekelberg B, van Wezel RJA, Albright TD. Adaptation in macaque MT reduces perceived speed and improves speed discrimination. *J Neurophysiol* 95: 255–270, 2006. doi:10.1152/jn.00750.2005.

Larsson J, Harrison SJ. Spatial specificity and inheritance of adaptation in human visual cortex. *J Neurophysiol* 114: 1211–1226, 2015. doi:10.1152/jn.00167.2015.

Larsson J, Heeger DJ. Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci* 26: 13128–13142, 2006. doi:10.1523/JNEUROSCI.1657-06.2006.

Larsson J, Landy MS, Heeger DJ. Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *J Neurophysiol* 95: 862–881, 2006. doi:10.1152/jn.00668.2005.

Larsson J, Smith AT. fMRI repetition suppression: neuronal adaptation or stimulus expectation? *Cereb Cortex* 22: 567–576, 2012. doi:10.1093/cercor/bhr119.

Lerner Y, Honey CJ, Silbert LJ, Hasson U. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J Neurosci* 31: 2906–2915, 2011. doi:10.1523/JNEUROSCI.3684-10.2011.

Lindquist MA, Meng Loh J, Atlas LY, Wager TD. Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling. *Neuroimage* 45, *Suppl*: S187–S198, 2009. doi:10.1016/j.neuroimage.2008.10.065.

Logothetis NK. The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci* 23: 3963–3971, 2003. doi:10.1523/JNEUROSCI.23-10-03963.2003.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412: 150–157, 2001. doi:10.1038/35084005.

Montaser-Kouhsari L, Landy MS, Heeger DJ, Larsson J. Orientation-selective adaptation to illusory contours in human visual cortex. *J Neurosci* 27: 2186–2195, 2007. doi:10.1523/JNEUROSCI.4173-06.2007.

Müller JR, Metha AB, Krauskopf J, Lennie P. Rapid adaptation in visual cortex to the structure of images. *Science* 285: 1405–1408, 1999. doi:10.1126/science.285.5432.1405.

Murray SO, Olman CA, Kersten D. Spatially specific FMRI repetition effects in human visual cortex. *J Neurophysiol* 95: 2439–2445, 2006. doi:10.1152/jn.01236.2005.

Murray SO, Wojciulik E. Attention increases neural selectivity in the human lateral occipital complex. *Nat Neurosci* 7: 70–74, 2004. doi:10.1038/nn1161.

Ollinger JM, Shulman GL, Corbetta M. Separating processes within a trial in event-related functional MRI I. The Method. *Neuroimage* 13: 210–217, 2001. doi:10.1006/nimg.2000.0710.

Patterson CA, Duijnhouwer J, Wissig SC, Krekelberg B, Kohn A. Similar adaptation effects in primary visual cortex and area MT of the macaque monkey under matched stimulus conditions. *J Neurophysiol* 111: 1203–1213, 2014. doi:10.1152/jn.00030.2013.

Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10: 437–442, 1997. doi:10.1163/156856897X00366.

Ray S, Ni AM, Maunsell JH. Strength of gamma rhythm depends on normalization. *PLoS Biol* 11: e1001477, 2013. doi:10.1371/journal.pbio.1001477.

Savitzky A, Golay MJ. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem* 36: 1627–1639, 1964. doi:10.1021/ac60214a047.

Sayres R, Grill-Spector K. Object-selective cortex exhibits performance-independent repetition suppression. *J Neurophysiol* 95: 995–1007, 2006. doi:10.1152/jn.00500.2005.

Seeliger K, Fritsche M, Güçlü U, Schoenmakers S, Schoffelen JM, Bosch SE, van Gerven MAJ. Convolutional neural network-based encoding and decoding of visual object recognition in space and time. *Neuroimage* 180: 253–266, 2018. doi:10.1016/j.neuroimage.2017.07.018.

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268: 889–893, 1995. doi:10.1126/science.7754376.

Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. *Annu Rev Neurosci* 24: 1193–1216, 2001. doi:10.1146/annurev.neuro.24.1.1193.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, *Suppl* 1: S208–S219, 2004. doi:10.1016/j.neuroimage.2004.07.051.

Soon CS, Venkatraman V, Chee MWL. Stimulus repetition and hemodynamic response refractoriness in event-related fMRI. *Hum Brain Mapp* 20: 1–12, 2003. doi:10.1002/hbm.10122.

Stigliani A, Weiner KS, Grill-Spector K. Temporal processing capacity in high-level visual cortex is domain specific. *J Neurosci* 35: 12412–12424, 2015. doi:10.1523/JNEUROSCI.4822-14.2015.

Tootell RB, Hadjikhani NK, Vanduffel W, Liu AK, Mendola JD, Sereno MI, Dale AM. Functional analysis of primary visual cortex (V1) in humans. *Proc Natl Acad Sci USA* 95: 811–817, 1998. doi:10.1073/pnas.95.3.811.

van Turennout M, Bielamowicz L, Martin A. Modulation of neural activity during object naming: effects of time and practice. *Cereb Cortex* 13: 381–391, 2003. doi:10.1093/cercor/13.4.381.

van Turennout M, Ellmore T, Martin A. Long-lasting cortical plasticity in the object naming system. *Nat Neurosci* 3: 1329–1334, 2000. doi:10.1038/81873.

Vuilleumier P, Henson RN, Driver J, Dolan RJ. Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nat Neurosci* 5: 491–499, 2002. doi:10.1038/nn839.

Wandell BA, Dumoulin SO, Brewer AA. Visual field maps in human cortex. *Neuron* 56: 366–383, 2007. doi:10.1016/j.neuron.2007.10.012.

Weigelt S, Limbach K, Singer W, Kohler A. Orientation-selective functional magnetic resonance imaging adaptation in primary visual cortex revisited. *Hum Brain Mapp* 33: 707–714, 2012. doi:10.1002/hbm.21244.

Weiner KS, Sayres R, Vinberg J, Grill-Spector K. fMRI-adaptation and category selectivity in human ventral temporal cortex: regional differences across time scales. *J Neurophysiol* 103: 3349–3365, 2010. doi:10.1152/jn.01108.2009.

Westerberg JA, Cox MA, Dougherty K, Maier A. V1 microcircuit dynamics: altered signal propagation suggests intracortical origins for adaptation in response to visual repetition. *J Neurophysiol* 121: 1938–1952, 2019. doi:10.1152/jn.00113.2019.

Winawer J, Witthoft N. Identification of the ventral occipital visual field maps in the human brain. *F1000 Res* 6: 1526, 2017. doi:10.12688/f1000research.12364.1.

Wissig SC, Kohn A. The influence of surround suppression on adaptation effects in primary visual cortex. *J Neurophysiol* 107: 3370–3384, 2012. doi:10.1152/jn.00739.2011.

Zhang N, Zhu XH, Chen W. Investigating the source of BOLD nonlinearity in human visual cortex in response to paired visual stimuli. *Neuroimage* 43: 204–212, 2008. doi:10.1016/j.neuroimage.2008.06.033.

Zhou J, Benson NC, Kay KN, Winawer J. Compressive temporal summation in human visual cortex. *J Neurosci* 38: 691–709, 2018. doi:10.1523/JNEUROSCI.1724-17.2017.