J Neurophysiol 130: 990–998, 2023.
First published September 14, 2023; doi:10.1152/jn.00194.2023

american physiological society

JNP JOURNAL OF NEUROPHYSIOLOGY

RESEARCH ARTICLE

Sensory Processing

Attention preserves the selectivity of feature-tuned normalization

Michaela Klímová,[1,2] Ilona M. Bloem,[1,2,3] and Sam Ling[1,2]

[1]*Department of Psychological and Brain Sciences, Boston University, Boston, Massachusetts, United States;* [2]*Center for Systems Neuroscience, Boston University, Boston, Massachusetts, United States; and* [3]*Department of Psychology, New York University, New York City, New York, United States*

Abstract

Attention and divisive normalization both contribute to making visual processing more efficient. Attention selectively increases the neural gain of relevant information in the early visual cortex, resulting in stronger perceived salience for attended regions or features. Divisive normalization improves processing efficiency by suppressing responses to homogeneous inputs and highlighting salient boundaries, facilitating sparse coding of inputs. Theoretical and empirical research suggest a tight link between attention and normalization, wherein attending to a stimulus results in a release from normalization, thereby allowing for an increase in neural response gain. In the present study, we address whether attention alters the qualitative properties of normalization. Specifically, we examine how attention influences the feature-tuned nature of normalization, whereby suppression is stronger between visual stimuli whose orientation contents are similar, and weaker when the orientations are different. Ten human observers viewed stimuli that varied in orientation content while we acquired fMRI BOLD responses under two attentional states: attending toward or attending away from the stimulus. Our results indicate that attention does not alter the specificity of feature-tuned normalization. Instead, attention seems to enhance visuocortical responses evenly, regardless of the degree of orientation similarity within the stimulus. Since visuocortical responses exhibit adaptation to statistical regularities in natural scenes, we conclude that while attention can selectively increase the gain of responses to attended items, it does not appear to alter the ecologically relevant correspondence between orientation differences and strength of tuned normalization.

NEW & NOTEWORTHY The magnitude of visuocortical BOLD responses scales with orientation differences in visual stimuli, with the strongest response suppression for collinear stimuli and least suppression for orthogonal, in a way that appears to match natural scene statistics. We examined the effects of attention on this feature-tuned property of suppression and found that while attending to a stimulus increases the overall gain of visuocortical responses, the qualitative properties of feature-tuning remain unchanged, suggesting attention preserves tuned normalization properties.

attention; divisive normalization; fMRI; vision

INTRODUCTION

Although our senses are constantly flooded with stimulation, they are surprisingly effective at filtering incoming information. To do so, the visual system deploys a number of strategies, some of which involve top-down, endogenous guidance, and others that are bottom-up and largely automatic. One such bottom-up process is divisive normalization, wherein the magnitude of a neuron's response to a stimulus is divided by the pooled responses of itself and neighboring neurons. Normalization is widely believed to be a canonical neural computation, as it operates throughout the visual processing hierarchy (1). Divisive normalization models offer a computational account of numerous commonly observed behaviors of early visual neurons, one of which is surround suppression (1, 2).

Surround suppression, where the neural response to a visual stimulus is reduced by the presence of surrounding or flanking stimuli, has been observed with electrophysiology (2–5), as well as with large-scale population measures such as fMRI (6–8) and EEG (9). Interestingly, not all surrounding stimuli and configurations are created equal: normalization appears to be feature-tuned, wherein the magnitude of suppression depends on the degree of feature similarity between

Correspondence: M. Klímová (mklimova@bu.edu).
Submitted 12 May 2023 / Revised 22 August 2023 / Accepted 11 September 2023

the center and surrounding stimuli. Specifically, surrounding stimuli whose features (e.g., orientation) differ from the central stimulus are less suppressive than surrounding stimuli that share identical properties with the center (4, 5, 10–22). This feature-tuned component of normalization is proposed to play a key functional role in efficient coding during natural scene perception. According to efficient coding approaches, the visual system uses feature-tuned normalization to compress representations of redundant information and enhance the salience of unique contours, which aids in segregating figures from background (23, 24). Interestingly, the selectivity of feature-tuned suppression within the human visual cortex appears to be nicely tuned to the image statistics of natural environments: by characterizing the bandwidth of orientation-tuned suppression in early visual areas V1–V3, we found that bandwidths fell between 20° and 30° orientation difference (25), closely matching the distribution of orientation co-occurrences in natural scenes (26).

In addition to spatial vision mechanisms such as divisive normalization, top-down processing also plays a key role in the moment-to-moment filtering of information. Visual attention helps parse scenes and select relevant stimuli for preferential processing, at the expense of unattended stimuli. This essential top-down mechanism allows us to regulate between the overabundance of sensory stimulation in our environment and the limited processing resources available to the brain at any given moment (27). Directing attention to a stimulus is well known to increase the perceptual salience of its representation (28–30)—a behavioral consequence of attention-driven changes in the gain of neural responses (31–33). Although divisive normalization and attention independently regulate incoming visual information, evidence gathered in recent years suggests that they may be closely intertwined. Normalization models of attention propose that attention elicits gain changes by altering the balance between excitation and inhibition, effectively providing a release from suppression (34–37). Functional neuroimaging in the human visual cortex supports the link between tuned normalization and attention, wherein subpopulations exhibiting stronger tuned normalization also show stronger attentional modulation (11). Likewise, psychophysical work found weaker attentional effects for stimuli that evoked weaker surround suppression [orthogonal arrangement as opposed to collinear; (38)].

Given this tight relationship between attention and normalization, could attention, in turn, also enhance the specificity of tuned normalization? Although the normalization model of attention proposes that attentional benefits emerge from a release from normalization, it remains unclear whether this also applies to the tuned component of normalization. If attention enhanced the specificity of orientation-tuned normalization, this would lead to tighter orientation-tuned normalization bandwidth. On the perceptual level, we may see heightened sensitivity to smaller changes in orientation, which could enhance scene processing. Such a modulation would qualitatively change the way feature differences in a scene are processed and perceived. Alternatively, attending to the stimuli could result in no changes to orientation-tuned normalization bandwidth. Work from our group and others (25, 26, 39) has supported the notion that normalization bandwidths are tuned to the statistical regularities of our visual environments—larger differences in orientation and other visual features are highlighted (by inducing larger neural responses), while continuities in scenes and objects are discounted (23). Given that tuned normalization already serves to faithfully adapt visual responses to the predominant distribution of orientations in the environment, perhaps attending to a stimulus should not augment this tuned function. In this study, we used fMRI to examine whether visual attention qualitatively modulates the profile of orientation-tuned normalization. We leveraged full-field stimuli previously designed to measure the bandwidth of orientation-tuned suppression (25) under two conditions: attending to the oriented stimuli and attending away (attending to a task at fixation). Our results revealed an overall enhancement of the blood oxygenation level-dependent (BOLD) response when the stimulus was attended to, but no discernible qualitative attentional effects on the orientation-difference tuning bandwidth, suggesting that while attention increases the neural responses to a stimulus, it does not alter properties of visual processing, which are likely titrated to natural scene statistics.

MATERIALS AND METHODS

Observers

Ten observers (4 female) took part in the study. All received monetary compensation for their time, except three who were also the authors of the study. The imaging for the attentional task was conducted in a single 2-h session. An additional fMRI session was conducted before the study to obtain a high-resolution T1-weighted scan and to conduct population receptive field (pRF) mapping to identify early visual areas V1–V3. All observers reported their vision as normal or corrected-to-normal and were between the ages of 18 and 40. All provided written informed consent and the study was approved by the Boston University Institutional Review Board.

Apparatus and Stimuli

Visual stimuli were generated with Psychophysics Toolbox in MATLAB (R2015b) on a MacBook Pro (OS X 10.7) and displayed in the scanner on a rear-projection screen using a linearized projection system (VPixx Technologies PROPixx DLP LED projector). The stimulus (Fig. 1*B*), subtending 3–17 degrees of visual angle (dva) in diameter was circular and composed of 20 wedges (two interleaved sets of 10 with variable orientation content). The spatial pattern filling each wedge was created by band-pass filtering white noise. The filter had an orientation bandwidth of 10° (5° on either side of the target orientation), and a spatial frequency bandwidth of 2–3 cycles per degree (cpd). Neighboring wedges could contain one of six orientation differences, spanning between 0° (collinear) and 90° (orthogonal), in steps of 18°. Furthermore, each orientation difference could be offset by one of five "base orientations," which spanned between 9° and 81° (also in steps of 18°). All stimuli were initially generated with one of the wedge sets containing 0° (vertical) orientation and the other orientation corresponded to the orientation difference; for example, in the 18° orientation difference condition, one set of wedges was 0° and the other 18°. On each trial, the entire stimulus was additionally rotated by the base

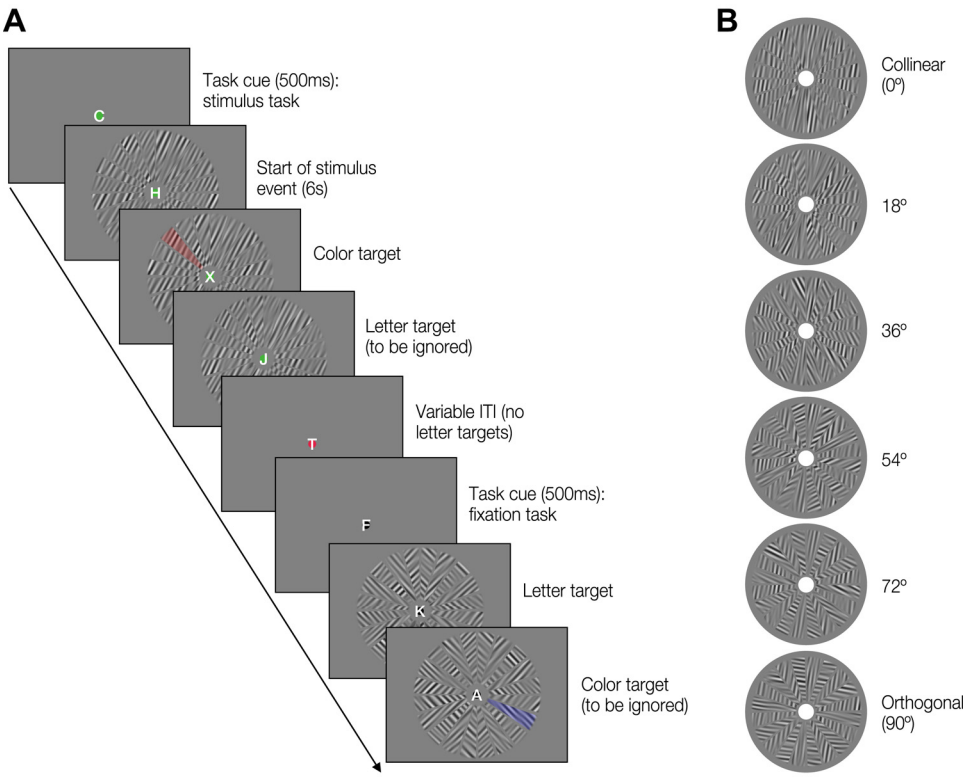

Figure 1. *A*: illustration of the experimental procedure, depicting a sample of two stimulus event blocks, with a blank interval in between. Here, the observer is shown a color task cue, followed by a stimulus event with a collinear condition. After an intertrial interval with no stimulus, a fixation task cue was shown, followed by a fixation task block with an orthogonal condition. Note that the only visual change on the screen between the two conditions is the fixation dot color. In this depiction, the opacity of the color patches has been increased for visibility, as has the size of the fixation dot and the letter. The larger white fixation circle is not depicted in this illustration. *B*: experimental stimuli. Stimuli are all depicted with a starting orientation of 9° from vertical. For illustration purposes, all stimuli are rendered at a lower spatial frequency and a higher contrast. ITI, intertrial interval.

orientation value for that trial. In the example aforementioned, if the base orientation was 9°, the orientations of the two-component sets would now be 9° and 27°, still maintaining the 18° offset. All stimuli had 50% Michelson contrast and were presented on a mean luminance background. Each fMRI run contained 36 stimulus presentations, and each condition (combination of orientation difference and attentional state) was presented three times within a run.

Base orientations were fully counterbalanced across every 10 runs, such that each base orientation was presented six times for each orientation and attention condition across the 10 runs. Furthermore, we counterbalanced which set of wedges contained the base orientation and which would contain the offset orientation (3 of 6 presentations each).

Experimental Procedure

Each observer completed 9–11 scan runs of the main task. The main task [408 TRs, 1-s TR (repetition time)] was a slow event-related design with a 6-s stimulus duration and variable ITI (intertrial interval; between 4 and 14 s). Event scheduling was done using FreeSurfer's Optseq2 scheduling tool (40). The event schedule in each run was identical for all observers in terms of orientation difference and attention condition, but the order of base orientations on each trial differed between observers. Each main task run began with an 8-s baseline period, only showing the mid-luminance screen and a rapid letter stream within the fixation circle (see next paragraph). Observers were asked to maintain central fixation throughout the experiment.

A small fixation point was presented in the center of a white fixation circle (3 dva diameter). In front of the fixation point, white letters (letter size 0.75 dva) were presented in a rapid visual stream at 5 Hz [RSVP (rapid serial visual presentation) task]. The fixation circle was always present, as was the letter stream; both remained white throughout the scan. The color of the fixation point differed between the baseline, the attend-to-stimulus condition, and the attend-away condition, serving as a cue to the observer to indicate which task to perform. Throughout baseline/null periods, the fixation point was red. Five hundred milliseconds before each 6-s task trial, the fixation point changed color to indicate the task for the upcoming trial. There was also a 500-ms letter cue, interrupting the RSVP stream, before each trial. A green fixation point and the cue letter "C" indicated that in the upcoming 6-s trial, the observer will be asked to detect transparent color patches within the stimulus display and ignore the RSVP letter stream so that covert attention is directed to the oriented stimulus. A black fixation point and the cue letter "F" told the observer that they should attend to the letter stream at fixation and monitor for target letters, thus passively viewing the stimulus while directing their attention away from it. The fixation point kept its task-specific color throughout the 6-s task block, to minimize the possibility that observers might erroneously focus on the wrong task.

During the attended trials (Fig. 1*A*, top four time points), observers were to maintain central fixation but monitor the entire stimulus for the appearance of a faint color patch that could appear over any of the wedges. The color patch appeared gradually, ramping up from α (opacity) = 0 to each observer's individual α level (see paragraph six in this section for average alpha value) and then back to 0 in the span of 1 s; the ramp-up took 400 ms, maximum α remained for 200 ms, followed by 400-ms ramp-down. A color target appeared with a 60% probability after two consecutive nontarget stimulus refresh cycles (100 ms). This led to between 0 and 2 color targets appearing in each 6-s task block. The participants held a two-button box

throughout the experiment with instructions to press the left key if they saw a red patch, and the right key if they saw a blue patch.

In the passive viewing trials (Fig. 1*A*, bottom three time points), participants monitored the central fixation letter stream, in which a new letter was presented every 200 ms. Upon seeing the letter "J," participants were to press the left button, and if the letter "K" was detected in the stream, the right button. A target letter appeared with a 40% probability after a minimum of 5 nontarget letters. Typically, between 2 and 3 letter targets appeared during each task event. During the null and baseline intervals, the fixation dot was red, and the letter stream continued but never contained target letters. Observers were instructed to simply maintain fixation during these periods.

Importantly, visual stimulation was identical throughout the experiment; i.e., the letter stream containing targets was presented within every block regardless of the task, as were the color patches. This ensured that differences in BOLD signal between attention conditions would not be a result of changes in visual stimulation but rather a result of the amount of attention directed to the stimulus.

One concern with the color patch detection/identification task was the opacity level; we wanted this task to be challenging such that observers would be motivated to continuously attend to the stimulus. Therefore, immediately after getting settled in the scanner (before fMRI data collection), observers completed two runs of a staircasing task in which the opacity of the patch was adjusted. The visual stimuli and response key mapping were identical to the main task, with the exception that every block contained the color task. The initial α (opacity) value was set to 0.1, and an adaptive staircase (41) converged after 40 trials in each run to determine the threshold for 90% accuracy. We took the lower of the two estimates as the observer's individual α value for the main task. We also monitored observers' performance on the color task throughout the scan session and adjusted the α as needed, to compensate for performance improving over time. On average, the α level was 0.053 (±0.007).

Observers found the color task to be more challenging compared with the fixation task, with mean letter RSVP accuracy 92.79 ± 1.36% and color patch identification accuracy 82.99 ± 1.26%. Across participants, the task performance difference reached statistical significance [paired t test, $t(9) = 6.9$, $P < 0.001$). In the color task, we also checked whether perceptual difficulty differed between the six orientation offsets; a one-way ANOVA indicated that perceptual difficulty was not changed by the orientation offset of the wedges.

After the completion of the color opacity thresholding and before starting the main runs, we also collected two runs of a functional localizer, presented in a block design (14 s on, 14 s off, 182 TRs at 1-s TR, starting and ending with an off block). The localizer stimulus had inner and outer eccentricities identical to the main task stimulus, but was a solid annulus (i.e., no separation into wedges) consisting of a 100% Michelson contrast solid spatial pattern generated by combining radial and spiral gratings, with 10-Hz contrast-reversing flicker. The inner and outer bounds of the localizer were identical to the experimental stimulus (inner diameter 3 dva, outer diameter 17 dva). Data from the functional localizer analysis were later used for voxel selection.

Eye-Tracking Data Acquisition

Eye-tracking data were collected for all observers with an MR-compatible EyeLink 1000 (SR Research, ON, Canada) eye tracker system. The sampling rate was 500 Hz. Due to a technical error, all runs from one observer and three runs in another observer were accidentally recorded at 1,000 Hz and terminated prematurely (after ~350 s). For these runs, we cropped the data after 350 s and included the resulting data files in the analysis.

Although the behavioral performance suggests that observers were maintaining steady fixation, we also compared eye movements between the two conditions to ensure there were no significant differences. Blinks were removed (padding 125 ms on either side of a detected blink, resulting in an average of 14.8 ± 2.2% of the eye-tracking data removed per participant), and the absolute distance from fixation was calculated for each attention condition (across orientation differences). The average distance was 0.39 dva (±0.04) in the fixation task and 0.4 (±0.04) in the color task, with no significant difference between the two conditions [paired t test, $t(9) = 0.66$, $P = 0.53$]. Moreover, in both conditions, the average distance from fixation was well under the radius of the fixation circle (1.5 dva).

MRI Data Acquisition

All MRI data were collected using a Siemens Prisma 3.0 Tesla scanner (Siemens, Erlagen, Germany) with a 64-channel head coil at the Boston University Cognitive Neuroimaging Center. fMRI data were acquired with simultaneous multi-slice echoplanar T2*-weighted imaging. The fMRI acquisition field of view (FOV) was oriented perpendicular to the calcarine sulcus [FOV = 60 × 112 × 172 mm, TR = 1,000 ms, TE (echo time) = 35 ms, FA (flip angle) = 80°, voxel size = 2 mm isotropic]. A whole brain anatomical scan used to register the functional data was acquired in a separate session (see *Population receptive field mapping*) using a T1-weighted multi-echo MPRAGE sequence (FOV = 256 × 256 × 176 mm, 36 slices, TR = 2,530 ms, TE = 1.69 ms, FA = 7°, voxel size = 1 mm isotropic).

Population receptive field mapping.

Prior to the experimental session, each observer completed a 1.5–2 h population receptive field (pRF) mapping session using standard stimuli and procedures provided by the analyzePRF toolbox (42). pRF mapping results were used to manually draw the boundaries of early visual areas V1, V2, and V3, identified as reversals in polar angle preference.

Anatomical data analyses.

The whole brain T1-weighted anatomical image (voxel size 1 mm isotropic) acquired at the start of the pRF mapping session was analyzed in FreeSurfer (v.5.3) using the recon-all pipeline. The anatomical scan was used to register the functional data with boundary-based registration.

fMRI Data Analyses and Voxel Selection

Before preprocessing the fMRI data, EPI (echo-planar imaging) distortion correction was applied with FSL (43) using the reverse phase-encoding method (44). Data were then preprocessed using FS-FAST (45) with standard motion

correction, Siemens slice timing correction, and boundary-based registration (46). Following this, we implemented robust rigid registration (47) for voxel-to-voxel alignment between individual runs. No spatial smoothing was applied in preprocessing (full width at half maximum = 0 mm).

Preprocessed functional localizer data were analyzed using a standard GLM analysis in FreeSurfer to identify voxels most responsive to the localizer stimulus. Preprocessed main task runs were analyzed via a univariate voxel-wise deconvolution in FSL. The window size was set to 24 TRs, with a prestimulus window of 4 TRs. For each voxel, we thus obtained 24 β weights for each combination of orientation difference and attentional state. This data was further analyzed with custom MATLAB scripts, and statistical comparisons between attention condition and visual areas were conducted by means of repeated-measures ANOVA in R. We report variance in our data as standard error of the mean, unless otherwise specified.

Only the top 40% of localizer-responsive voxels were selected for further analysis within each visual area, based on their ranking of the localizer GLM significance values. Out of this pool, we excluded any voxels whose pRF eccentricity estimates fell outside the stimulus bounds (1.5 dva inner radius and 8.5 dva outer radius). We also excluded voxels with poor goodness-of-fit of their pRF models (pRF $R^2 < 10\%$). Following this initial voxel selection, the average number of voxels submitted to the fitting procedure (see next section) was 476 (SD ± 65) in V1, 343 (±35) in V2, and 264 (±39) in V3.

Fitting the orientation-tuned suppression function. To obtain the tuned normalization functions for each observer, we first averaged β weights across all voxels and conditions within each visual area and determined the TR of the peak of this time course. We then took three TRs on either side of this peak as the end-points of the averaging window for each voxel for that observer and visual area, resulting in a 7-TR temporal averaging window (equal across conditions). This step aimed to account for differences in the timing of the hemodynamic response between observers and conditions. The resulting functions of β weights against orientation difference were fit with a half Gaussian function on a voxel-wise basis using the *fmincon* function in MATLAB, as follows:

$$\text{Response} = b + Ae^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

The fitting procedure was conducted separately for the two attention conditions and converged on a solution for all voxels. The μ parameter represents the mean of the function and designates where it will be centered and was fixed to zero. The bandwidth is denoted by the σ parameter, or standard deviation, and was constrained between 8° (the minimum possible difference between the oriented components given the orientation filter bandwidth and minimum orientation difference) and 90°. b represents the baseline β weight (BOLD signal gain) and accounts for the offset of the BOLD response; we left this free parameter unconstrained. A is the gain of the Gaussian and was constrained with an upper bound of 10. Since the mean was set as 0, the fitting procedure largely resulted in an inverted Gaussian (with the lowest response being in the collinear condition), and the gain parameter is negative across voxels for most observers. The gain represents the difference between the fitted response to the orthogonal and collinear conditions. Following parameter estimation, we also calculated R^2 of the Gaussian model fit for each voxel. We found that in each observer and visual area, a subset of voxels had bandwidth estimates that fell along the upper boundary of the constraints for fitting (between 85° and 90°) in at least one condition. A comparison of the R^2 values of the Gaussian fits between voxels whose bandwidth fell above 85° and the remainder revealed significantly lower quality of fits in the high-bandwidth voxels, based on two-factor within-subjects ANOVA with bandwidth group and attention condition as factors [$F(1,9) = 79.92$, $P < 0.001$]. In the fixation task, the upper-boundary voxels had R^2 values of 34% (±5.5%) in V1, 42% (±4.1%) in V2, and 34% (±4.7%) in V3, whereas the remaining voxels' R^2 values were 60% (±3.8%) in V1, 65% (±4.7%) in V2, and 57% (±6.1%) in V3. The values were similar in the attend stimulus color task: voxels with higher bandwidth estimates had R^2 values of 35.6% (±4%) in V1, 39.7% (±3.6%) in V2, and 37.9% (±4.3%) in V3, and voxels with lower bandwidths 56.5% (±4.5%) in V1, 58.5% (±4.8%) in V2, and 60% (±4.2%) in V3. We therefore opted for their removal from the data set. Across observers, we removed 52 ± 2.7% voxels in V1, 45.9 ± 2.8% in V2, and 51.6 ± 1.9% in V3, and the following analyses were conducted using this voxel subset. Before voxel exclusion, there was a significant R^2 difference between visual areas in a repeated-measures (RM) ANOVA with factors visual area and attention condition [$F(2,18) = 4.923$, $P = 0.02$], driven by higher R^2 values in V2 as compared with V1 and V3 [V1 vs. V2: $t(19) = -3.18$, $P = 0.015$; V2 vs. V3: $t(19) = 3.15$, $P = 0.016$, after Bonferroni correction]; this difference was not observed in the subset of voxels that were kept for further analysis [$F(2,18) = 1.374$, $P = 0.28$]. Importantly, in either set of voxels, we did not find significant differences in R^2 between the two attention conditions.

RESULTS

Orientation-Tuned Normalization with and without Attention

In line with our previous results, we observed lower BOLD responses for collinear configurations and larger responses for orthogonal stimuli. For both attentional states, the response gradually increased as a function of increasing orientation similarity, indicating stronger neural suppression as a function of orientation similarity. Figure 2 depicts the observer-averaged BOLD responses (β weights) as a function of orientation difference under the two attention conditions, for all visual areas. Although this general shape can be seen in both tasks, attending to the stimulus evoked a clear overall increase in BOLD response.

Voxel-wise tuned suppression strength (defined as orthogonal unattended minus collinear unattended BOLD response) was positive for the majority of voxels for all but one observer (Fig. 3*A*). We summarized this measure by taking the median across voxels for each observer/visual area and averaging across observers per visual area. In every visual area, the suppression strength measure was significantly different from zero [one-sample t test, V1: $t(9) = 2.84$, $P = 0.019$; V2: $t(9) = 4.7$, $P = 0.001$; V3: $t(9) = 4.0$, $P = 0.003$]. Overall, suppression strength did not differ between visual areas [$F(2,18) = 2.36$, $P = 0.123$].

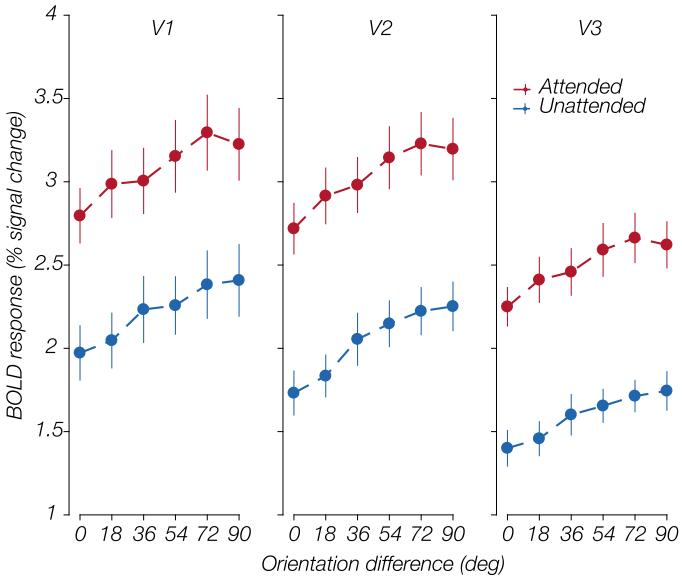

Figure 2. Observer-averaged (n = 10) BOLD response (means ± SE) as a function of stimulus orientation difference in the attended (color task; red) and unattended (fixation/letter task; blue) conditions, in each early visual area. The plots show a gradual increase in BOLD response as a function of orientation offset in both attention conditions, and a notable upward shift of the orientation-tuned normalization function in the attended condition, compared with unattended. BOLD, blood oxygenation level dependent.

Voxel-wise attentional modulation (the difference between β weights in the attended collinear condition minus the unattended collinear condition) was positive across all observers and visual areas (Fig. 3*B*), and significantly different from zero [one-sample t test, V1: $t(9) = 6.93$, $P < 0.001$; V2: $t(9) = 9.16$, $P < 0.001$; V3: $t(9) = 8.37$, $P < 0.001$], again confirming that visual attention increased the BOLD response. A repeated-measures ANOVA with attention condition and visual area as factors revealed a significant difference between visual areas [$F(2,18) = 3.95$, $P = 0.038$]; however, when following up with Bonferroni-corrected t tests, we found that none of the between-visual area differences reached significance. The main effect appeared to be driven by higher attentional modulation in V2 compared with V1 [$t(9) = -2.58$, $P = 0.03$ before correction, $P = 0.089$ after correction].

Quantifying Effects of Attention

As a means of quantifying changes to the tuned normalization function with attention, we described each voxel's response by fitting it with a half-Gaussian function, as done previously (25). The bandwidth of tuned normalization was denoted by the standard deviation (σ) parameter of the Gaussian. The gain parameter (A) represents the peak of the function. Note that the function is centered on at 0° orientation difference (the mean of the function was set to zero), resulting in an inverted half-Gaussian function where a negative gain represents a lower response for collinear than orthogonal stimulus configurations. We also included an overall response offset parameter (b) to account for the upward shift in responses when the stimuli were attended.

Following the voxel-wise fitting and voxel exclusion (see METHODS), The Gaussian parameters: bandwidth, gain, and baseline were submitted to a two-factor within-subjects ANOVA, with the factors of visual area (V1, V2, V3) and attentional state (attend to stimulus or attend to fixation). As can be seen in Fig. 4*A*, baseline estimates were elevated when observers attended to the stimulus as opposed to attending to the central fixation task; the ANOVA confirmed a significant main effect of attention in this direction [$F(1,9) = 85.17$, $P < 0.001$]. The average b estimate (in % signal change) in the fixation condition was 2.21 ± 0.23 in V1, 2.14 ± 0.15 in V2, and 1.64 ± 0.12 in V3, whereas in the color condition the estimates were 2.91 ± 0.25 in V1, 2.99 ± 0.2 in V2, and 2.51 ± 0.18 in V3. There was also a main effect of the visual area [$F(2,18) = 10.71$, $P < 0.001$]. Bonferroni-corrected post hoc tests showed that this effect was driven by significantly higher BOLD response in V1 and V2, compared with V3 [V1 vs. V3 $t(19) = 4.66$, $P < 0.001$; V2 vs. V3 $t(19) = 10.5$, $P < 0.001$]. The interaction term was marginally nonsignificant [$F(2,18) = 3.488$, $P = 0.053$].

Crucially, the bandwidth of tuned normalization (Fig. 4*B*) did not differ between attentional states [$F(1,9) = 0.308$, $P = 0.593$] or visual areas [$F(18,2) = 0.414$, $P = 0.667$], with a nonsignificant interaction term [$F(2,18) = 0.004$, $P = 0.996$], indicating lack of modulation in orientation-difference tuning as a result of attending to the stimuli. Likewise, changes in the gain parameter (Fig. 4*C*) were also nonsignificant, both between attention conditions [$F(1,9) = 0.018$, $P = 0.897$] and between visual areas [$F(2,18) = 2.404$, $P = 0.119$], with no interaction [$F(2,18) = 0.534$, $P = 0.595$]. In the unattended condition, the average bandwidth was 20.6° (±1.35°) in V1, 20.8° (±0.84°) in V2, and 20.23° (±0.94°) in V3, whereas in the attended condition, bandwidths were 21.96° (±2.35°) in V1, 22.06° (±1.85°) in V2, and 21.5° (±1.93°) in V3.

To verify that our voxel exclusion criteria based on the bandwidth estimates did not bias the overall statistical results, we repeated the above tests having skipped the voxel exclusion step. The results were consistent between the two

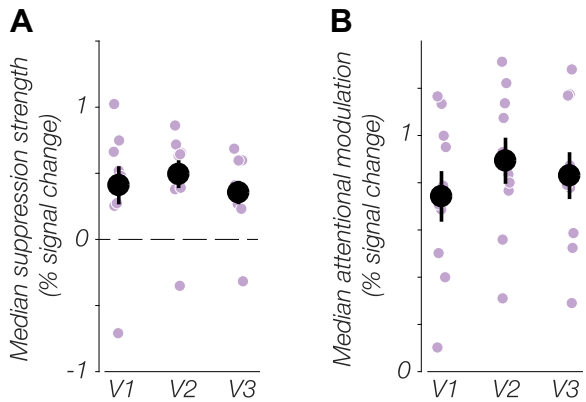

Figure 3. *A*: observer-averaged (n = 10) median orientation-tuned suppression strength in areas V1–V3, defined as voxel-wise difference between β weight in orthogonal minus collinear unattended conditions. The plot shows that the BOLD responses of 9 of 10 observers were higher in the orthogonal condition compared with collinear. There were no significant differences across visual areas [$F(2,18) = 2.36$, $P = 0.123$]. *B*: observer-averaged median attentional modulation, computed for each voxel as the difference between β weight in collinear attended condition minus collinear unattended condition. Attentional modulation was higher in V2 compared with V1; however, this difference did not survive multiple comparisons correction [$t(9) = -2.58$, $P = 0.089$ after correction). Error bars represent means ± SE. BOLD, blood oxygenation level dependent.

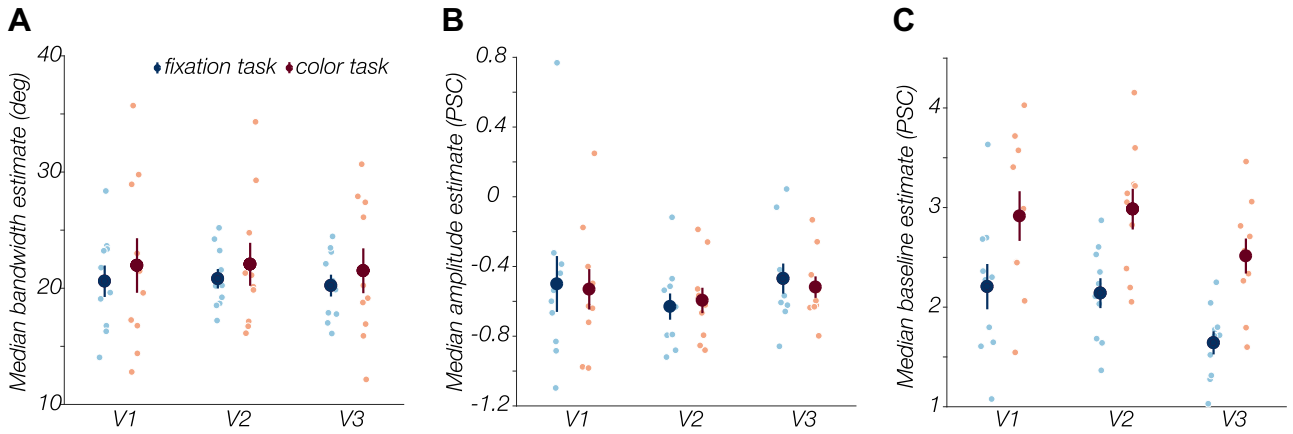

Figure 4. Observer-averaged ($n = 10$) Gaussian parameters describing the tuned normalization functions across attention conditions in each visual area. *A*: baseline of the orientation-tuned normalization function was significantly higher in the attended vs. unattended condition across visual areas [$F(1,9) = 85.17$, $P < 0.001$]. Neither bandwidth (*B*) nor amplitude (*C*) of the orientation-tuned normalization function showed differences between the attended and unattended conditions [$F(1,9) = 0.308$, $P = 0.593$, and $F(1,9) = 0.018$, $P = 0.897$, respectively]. Error bars represent ± 1 SE. PSC, % signal change.

voxel subsets; the only exception besides the R^2 measures detailed in METHODS is the interaction term between visual area and attention condition for the offset (b) parameter, which shifted from marginally nonsignificant [$F(2,18) = 3.488$, $P = 0.053$] to marginally significant when all initially fitted voxels were submitted to the ANOVA [$F(2,18) = 3.689$, $P = 0.0454$].

DISCUSSION

Our data show an overall significantly increased BOLD signal in the visuocortical responses to oriented stimuli when stimuli are attended to, compared with attention withdrawn away. Importantly, the absence of significant shifts in the bandwidth and gain of orientation-tuned suppression further points to a lack of qualitative change in the response magnitude to varying feature differences, suggesting that attending to a visual stimulus did not alter the sensitivity to orientation differences in early visual areas.

We designed our attention manipulation to ensure participants were directing covert, endogenous attention to the entire oriented stimulus (by virtue of the unpredictability of the target location), and to avoid any exogenous attention or pop-out effects (through the use of individual opacity thresholds for each observer and constant monitoring of performance). Covert visual attention and its effects on sensory gain have most frequently been studied in the context of how it affects neural responses to varying stimulus intensity (37, 48–50). A common finding in electrophysiological experiments is a multiplicative gain increase in the neural contrast response (51, 52). In contrast, in the fMRI literature, a majority of findings report an additive increase in attended states over unattended (49, 53, 54). To explain this discrepancy, it has been suggested that additive baseline effects seen in fMRI could simply be the result of insensitivity of the BOLD signal to modifications of stimulus-driven activity (55). A recent study from our laboratory presented evidence that attention in fact produces nonadditive effects in contrast response functions (CRF) measured with BOLD fMRI (56). It is important to note that in this study, we did not manipulate stimulus intensity, and instead changed only the orientation offsets at a fixed 50% contrast level. Therefore, the increase in our orientation-tuned normalization baseline parameter should not be confused with the baseline shifts sometimes found in attention studies that measure CRF (49, 53, 54). Our results could be driven by either multiplicative or additive modulation in the domain of contrast; within the confines of recording at a single contrast level, we can only comment on attentional modulation of the orientation-tuned suppression function. Similarly, we are also not measuring orientation tuning, but rather tuning to the orientation differences in our stimulus, a higher-order statistic.

A limitation of our study design and stimulus choice is the absence of a no-surround condition, which precludes the measurement of a BOLD response baseline without surround suppression. This leaves open the possibility of an alternative explanation for our results. The increased BOLD response in orthogonal compared with collinear configuration reflects facilitation of the BOLD response by the presence of multiple orientations, thereby exciting two distinct subpopulations of neurons selective for the two orientations, whereas in the collinear configuration only one subpopulation is stimulated. However, the facilitation explanation is inconsistent with findings of subadditivity, where superimposing a grating stimulus over another results in a response that is lower than would be predicted by the linear addition of two separate responses to the two components of the superimposed stimulus (11). Subadditivity is present also for orthogonal stimuli, suggesting that suppression is taking place upon presenting an overlay, annular or flanking stimulus (12, 16, 21).

Our findings square with theories of efficient coding, which posit that sensory neurons are adapted to the signals they are most exposed to, and in particular, that signal processing in the visual cortex is optimized to the prevalent image statistics we find in our natural environments (23, 57, 58). For instance, the estimated orientation-tuned suppression bandwidths in this study are comparable with the parameter estimates from our previous work, in which average bandwidth estimates were between 23.1° (V1), 24.4° (V2), and 22.8° (V3) (25), and the pattern of suppression strength as a function of orientation similarity bears resemblance to the

prevalence of orientation differences in nearby regions of a scene (26). Although it has been established that attention alters subjective appearance in ways that enhance the selective processing of attended areas, objects, or features (28, 29), a change in tuned normalization bandwidth would imply a change in the processing of second-order statistics (i.e., a change in the subjective salience of feature differences), which is to some degree dependent upon our visual system's apparent adaptation to the typical composition of orientations and other basic visual features in our environments (57).

DATA AVAILABILITY

Source data for this study are available at https://doi.org/10.17605/osf.io/6tq8h.

ACKNOWLEDGMENTS

We thank the members of the Ling Lab for providing helpful feedback and comments on this work.

GRANTS

This work was funded by NIH Grant EY028163 (to S. Ling). The equipment used was funded by NSF Major Instrumentation Grant 1625552.

DISCLOSURES

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

AUTHOR CONTRIBUTIONS

M.K., I.M.B., and S.L. conceived and designed research; M.K. and I.M.B. performed experiments; M.K. analyzed data; M.K. and S.L. interpreted results of experiments; M.K. prepared figures; M.K. drafted manuscript; M.K., I.M.B., and S.L. edited and revised manuscript; M.K., I.M.B., and S.L. approved final version of manuscript.

REFERENCES

1. **Carandini M**, **Heeger DJ.** Normalization as a canonical neural computation. *Nat Rev Neurosci* 13: 51–62, 2011 [Erratum in *Nat Rev Neurosci* 14: 152, 2013]. doi:10.1038/nrn3136.
2. **Cavanaugh JR**, **Bair W**, **Movshon JA.** Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J Neurophysiol* 88: 2530–2546, 2002. doi:10.1152/jn.00692.2001.
3. **DeAngelis GC**, **Robson JG**, **Ohzawa I**, **Freeman RD.** Organization of suppression in receptive fields of neurons in cat visual cortex. *J Neurophysiol* 68: 144–163, 1992. doi:10.1152/jn.1992.68.1.144.
4. **Self MW**, **Lorteije JAM**, **Vangeneugden J**, **van Beest EH**, **Grigore ME**, **Levelt CN**, **Heimel JA**, **Roelfsema PR.** Orientation-tuned surround suppression in mouse visual cortex. *J Neurosci* 34: 9290–9304, 2014. doi:10.1523/JNEUROSCI.5051-13.2014.
5. **Webb BS**, **Dhruv NT**, **Solomon S**, **Tailby C**, **Lennie P.** Early and late mechanisms of surround suppression in striate cortex of macaque. *J Neurosci* 25: 11666–11675, 2005. doi:10.1523/JNEUROSCI.3414-05.2005.
6. **Flevaris AV**, **Murray SO.** Attention determines contextual enhancement versus suppression in human primary visual cortex. *J Neurosci* 35: 12273–12280, 2015. doi:10.1523/JNEUROSCI.1409-15.2015.
7. **Joo SJ**, **Boynton GM**, **Murray SO.** Long-range, pattern-dependent contextual effects in early human visual cortex. *Curr Biol* 22: 781–786, 2012. doi:10.1016/j.cub.2012.02.067.
8. **Zenger-Landolt B**, **Heeger DJ.** Response suppression in V1 agrees with psychophysics of surround masking. *J Neurosci* 23: 6884–6893, 2003. doi:10.1523/JNEUROSCI.23-17-06884.2003.
9. **Schallmo M-P**, **Kale AM**, **Murray SO.** The time course of different surround suppression mechanisms. *J Vis* 19: 12, 2019. doi:10.1167/19.4.12.
10. **Bloem IM**, **Watanabe YL**, **Kibbe MM**, **Ling S.** Visual memories bypass normalization. *Psychol Sci* 29: 845–856, 2018. doi:10.1177/0956797617747091.
11. **Bloem IM**, **Ling S.** Normalization governs attentional modulation within human visual cortex. *Nat Commun* 10: 5660, 2019. doi:10.1038/s41467-019-13597-1.
12. **Chen C-C.** Partitioning two components of BOLD activation suppression in flanker effects. *Front Neurosci* 8: 149, 2014. doi:10.3389/fnins.2014.00149.
13. **Ling S**, **Pratte MS**, **Tong F.** Attention alters orientation processing in the human lateral geniculate nucleus. *Nat Neurosci* 18: 496–498, 2015. doi:10.1038/nn.3967.
14. **McDonald JS**, **Seymour KJ**, **Schira MM**, **Spehar B**, **Clifford CWG.** Orientation-specific contextual modulation of the fMRI BOLD response to luminance and chromatic gratings in human visual cortex. *Vision Res* 49: 1397–1405, 2009. doi:10.1016/j.visres.2008.12.014.
15. **Petrov Y**, **Carandini M**, **McKee S.** Two distinct mechanisms of suppression in human vision. *J Neurosci* 25: 8704–8707, 2005. doi:10.1523/JNEUROSCI.2871-05.2005.
16. **Pihlaja M**, **Henriksson L**, **James AC**, **Vanni S.** Quantitative multifocal fMRI shows active suppression in human V1. *Hum Brain Mapp* 29: 1001–1014, 2008. doi:10.1002/hbm.20442.
17. **Poltoratski S**, **Ling S**, **McCormack D**, **Tong F.** Characterizing the effects of feature salience and top-down attention in the early visual system. *J Neurophysiol* 118: 564–573, 2017. doi:10.1152/jn.00924.2016.
18. **Schallmo M-P**, **Murray SO.** Identifying separate components of surround suppression. *J Vis* 16: 2, 2016. doi:10.1167/16.1.2.
19. **Shushruth S**, **Nurminen L**, **Bijanzadeh M**, **Ichida JM**, **Vanni S**, **Angelucci A.** Different orientation tuning of near- and far-surround suppression in macaque primary visual cortex mirrors their tuning in human perception. *J Neurosci* 33: 106–119, 2013. doi:10.1523/JNEUROSCI.2518-12.2013.
20. **Trott AR**, **Born RT.** Input-gain control produces feature-specific surround suppression. *J Neurosci* 35: 4973–4982, 2015. doi:10.1523/JNEUROSCI.4000-14.2015.
21. **Williams AL**, **Singh KD**, **Smith AT.** Surround modulation measured with functional MRI in the human visual cortex. *J Neurophysiol* 89: 525–533, 2003. doi:10.1152/jn.00048.2002.
22. **Xing J**, **Heeger DJ.** Center-surround interactions in foveal and peripheral vision. *Vision Res* 40: 3065–3072, 2000. doi:10.1016/S0042-6989(00)00152-8.
23. **Coen-Cagli R**, **Kohn A**, **Schwartz O.** Flexible gating of contextual influences in natural vision. *Nat Neurosci* 18: 1648–1655, 2015. doi:10.1038/nn.4128.
24. **Schwartz O**, **Simoncelli EP.** Natural signal statistics and sensory gain control. *Nat Neurosci* 4: 819–825, 2001. doi:10.1038/90526.
25. **Klímová M**, **Bloem IM**, **Ling S.** The specificity of orientation-tuned normalization within human early visual cortex. *J Neurophysiol* 126: 1536–1546, 2021. doi:10.1152/jn.00203.2021.
26. **Sigman M**, **Cecchi GA**, **Gilbert CD**, **Magnasco MO.** On a common circle: natural scenes and Gestalt rules. *Proc Natl Acad Sci USA* 98: 1935–1940, 2001. doi:10.1073/pnas.98.4.1935.
27. **Desimone R**, **Duncan J.** Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18: 193–222, 1995. doi:10.1146/annurev.ne.18.030195.001205.
28. **Carrasco M**, **Ling S**, **Read S.** Attention alters appearance. *Nat Neurosci* 7: 308–313, 2004. doi:10.1038/nn1194.
29. **Carrasco M**, **Barbot A.** Spatial attention alters visual appearance. *Curr Opin Psychol* 29: 56–64, 2019. doi:10.1016/j.copsyc.2018.10.010.
30. **Yeshurun Y**, **Rashal E.** Precueing attention to the target location diminishes crowding and reduces the critical distance. *J Vis* 10: 16–16, 2010. doi:10.1167/10.10.16.
31. **Kastner S**, **De Weerd P**, **Desimone R**, **Ungerleider LG.** Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 282: 108–111, 1998. doi:10.1126/science.282.5386.108.
32. **Somers DC**, **Dale AM**, **Seiffert AE**, **Tootell RBH.** Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proc Natl Acad Sci USA* 96: 1663–1668, 1999. doi:10.1073/pnas.96.4.1663.

33. **Treue S.** Neural correlates of attention in primate visual cortex. *Trends Neurosci* 24: 295–300, 2001. doi:10.1016/S0166-2236(00)01814-2.
34. **Lee J**, **Maunsell JHR.** A normalization model of attentional modulation of single unit responses. *PLoS One* 4: e4651, 2009. doi:10.1371/journal.pone.0004651.
35. **Ni AM**, **Ray S**, **Maunsell JHR.** Tuned normalization explains the size of attention modulations. *Neuron* 73: 803–813, 2012. doi:10.1016/j.neuron.2012.01.006.
36. **Ni AM**, **Maunsell JHR.** Spatially tuned normalization explains attention modulation variance within neurons. *J Neurophysiol* 118: 1903–1913, 2017. doi:10.1152/jn.00218.2017.
37. **Reynolds JH**, **Heeger DJ.** The normalization model of attention. *Neuron* 61: 168–185, 2009. doi:10.1016/j.neuron.2009.01.002.
38. **Kınıklıoğlu M**, **Boyaci H.** Increasing the spatial extent of attention strengthens surround suppression. *Vision Res* 199: 108074, 2022. doi:10.1016/j.visres.2022.108074.
39. **Phillips DJ**, **McDougall TJ**, **Dickinson JE**, **Badcock DR.** Motion direction tuning in centre-surround suppression of contrast. *Vision Res* 179: 85–93, 2021. doi:10.1016/j.visres.2020.11.001.
40. **Dale AM.** Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8: 109–114, 1999. doi:10.1002/(SICI)1097-0193(1999)8:2/3<109::AID-HBM7>3.0.CO;2-W.
41. **Watson AB**, **Pelli DG.** Quest: a Bayesian adaptive psychometric method. *Percept Psychophys* 33: 113–120, 1983. doi:10.3758/BF03202828.
42. **Kay KN**, **Winawer J**, **Mezer A**, **Wandell BA.** Compressive spatial summation in human visual cortex. *J Neurophysiol* 110: 481–494, 2013. doi:10.1152/jn.00105.2013.
43. **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.
44. **Andersson JLR**, **Skare S**, **Ashburner J.** How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* 20: 870–888, 2003. doi:10.1016/S1053-8119(03)00336-7.
45. **Fischl B.** FreeSurfer. *NeuroImage* 62: 774–781, 2012. doi:10.1016/j.neuroimage.2012.01.021.
46. **Greve DN**, **Fischl B.** Accurate and robust brain image alignment using boundary-based registration. *NeuroImage* 48: 63–72, 2009. doi:10.1016/j.neuroimage.2009.06.060.
47. **Reuter M**, **Rosas HD**, **Fischl B.** Highly accurate inverse consistent registration: a robust approach. *NeuroImage* 53: 1181–1196, 2010. doi:10.1016/j.neuroimage.2010.07.020.
48. **Carrasco M.** Visual attention: the past 25 years. *Vision Res* 51: 1484–1525, 2011. doi:10.1016/j.visres.2011.04.012.
49. **Itthipuripat S**, **Sprague TC**, **Serences JT.** Functional MRI and EEG index complementary attentional modulations. *J Neurosci* 39: 6162–6179, 2019. doi:10.1523/JNEUROSCI.2519-18.2019.
50. **Martínez-Trujillo J**, **Treue S.** Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron* 35: 365–370, 2002. doi:10.1016/S0896-6273(02)00778-X.
51. **McAdams CJ**, **Maunsell JHR.** Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19: 431–441, 1999. doi:10.1523/JNEUROSCI.19-01-00431.1999.
52. **Williford T**, **Maunsell JHR.** Effects of spatial attention on contrast response functions in macaque area V4. *J Neurophysiol* 96: 40–54, 2006. doi:10.1152/jn.01207.2005.
53. **Buracas GT**, **Boynton GM.** The effect of spatial attention on contrast response functions in human visual cortex. *J Neurosci* 27: 93–97, 2007. doi:10.1523/JNEUROSCI.3162-06.2007.
54. **Murray SO.** The effects of spatial attention in early human visual cortex are stimulus independent. *J Vis* 8: 2.1–2.11, 2008. doi:10.1167/8.10.2.
55. **Itthipuripat S**, **Ester EF**, **Deering S**, **Serences JT.** Sensory gain outperforms efficient readout mechanisms in predicting attention-related improvements in behavior. *J Neurosci* 34: 13384–13398, 2014. doi:10.1523/JNEUROSCI.2277-14.2014.
56. **Foster JJ**, **Ling S.** Feature-based attention multiplicatively scales the fMRI-BOLD contrast-response function. *J Neurosci* 42: 6894–6906, 2022. doi:10.1523/JNEUROSCI.0513-22.2022.
57. **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.
58. **Vinje WE**, **Gallant JL.** Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287: 1273–1276, 2000. doi:10.1126/science.287.5456.1273.