

# Distinct patterns of surround modulation in V1 and hMT+

Gorkem Er<sup>a,+</sup>, Zahide Pamir<sup>\*,a,+</sup>, and Huseyin Boyaci<sup>a,b,c</sup>

<sup>a</sup>A.S. Brain Research Center; National Magnetic Resonance Research Center; Neuroscience Graduate Program, Bilkent University, Ankara, Turkey

<sup>b</sup>Department of Psychology, Bilkent University, Ankara, Turkey

<sup>c</sup>Department of Psychology, J.L. Gießen University, Gießen, Germany

<sup>+</sup>Equal contribution, alphabetical order

## Abstract

Modulation of a neuron's responses by the stimuli presented outside of its classical receptive field is ubiquitous in the visual system. This "surround modulation" mechanism is believed to be critical for efficient processing and leads to many well-known perceptual effects. The details of surround modulation, however, are still not fully understood. One of the open questions is related to the differences in surround modulation mechanisms in different cortical areas, and their interactions. Here we study patterns of surround modulation in primary visual cortex (V1) and middle temporal complex (hMT+) utilizing a well-studied effect in motion perception, where human observers' ability to discriminate the drift direction of a grating improves as its size gets bigger if the grating has a low contrast, and deteriorates if it has a high contrast. We first replicated the findings in the literature with a behavioral experiment using small and large (1.06 and 8.05 degrees of visual angle) drifting gratings with either low (2%) or high (99%) contrast presented at the periphery. Next, using functional MRI, we found that in V1 with increasing size cortical responses increased at both contrast levels, but they increased more at high contrast. Whereas in hMT+ with increasing size cortical responses remained unchanged at high contrast, and increased at low contrast. These findings show that surround modulation in V1 and hMT+ are distinct. Furthermore these findings provide evidence that the size-contrast interaction in motion perception is likely to originate in hMT+.

**Keywords:** surround modulation, surround suppression, surround facilitation, motion perception, primary visual cortex, middle temporal complex

## 1 Introduction

Visual neurons respond to stimuli only within their classical receptive fields (RF) when these stimuli are presented in isolation. If, however, the RF and its surround are stimulated together, the response patterns of the neurons alter. This surround manipulation is found in many levels of the visual hierarchy (Angelucci et al., 2017).

---

\*Present address: Schepens Eye Research Institute, Massachusetts Eye and Ear, Department of Ophthalmology, Harvard Medical School, Boston, MA, USA

37 Yet many questions about the mechanism remain open. Most importantly, even  
38 though surround modulation is heavily studied in primary visual cortex (V1), it is  
39 not clear whether the basic principles of the mechanism remains the same for a variety  
40 of stimuli in other visual areas. Here, to tackle this question we used a well-known  
41 perceptual effect in motion perception and investigated surround modulation in V1  
42 and human middle temporal complex (hMT+).

43 In the aforementioned perceptual effect, as the size of a drifting grating increases,  
44 discriminating its motion direction becomes harder if it has high contrast, but easier  
45 if it has low contrast (Tadin, Lappin, Gilroy, & Blake, 2003). This perceptual effect  
46 has been attributed to the surround modulation of neuronal populations in hMT+,  
47 which is one of the central cortical areas in motion processing ('MT-hypothesis', Tadin  
48 et al., 2003). According to the MT-hypothesis, discriminating motion direction of a  
49 high-contrast grating becomes harder owing to the suppressive effects of surround  
50 stimulation (i.e. surround suppression). For a low contrast grating, on the other  
51 hand, motion direction discrimination becomes easier as its size gets larger owing  
52 to facilitative effects of surround stimulation (i.e. surround facilitation). The MT-  
53 hypothesis has been later supported by functional magnetic resonance imaging (fMRI)  
54 findings (Turkcozer, Pamir, & Boyaci, 2016; Liu, Haefner, & Pack, 2016; Schallmo et  
55 al., 2018). Furthermore, disrupting hMT+ activity with application of TMS resulted  
56 in decreased surround suppression for high-contrast stimuli (Tadin, Silvanto, Pascual-  
57 Leone, & Battelli, 2011).

58 Previous studies, however, were not able to address the question of the role of other  
59 cortical areas in the observed perceptual effect. For example, neuronal correlates  
60 of spatial suppression and facilitation within the hMT+ might be inherited from  
61 earlier visual areas, most notably from V1. This possibility has not been systemati-  
62 cally analyzed in humans yet. One major cause of this was related to methodological  
63 limitations: Previous studies investigating the size-contrast interaction in motion per-  
64 ception used foveally presented stimuli, which activate neurons in the so-called foveal  
65 confluence, where borders of V1, V2, and V3 are difficult to draw using fMRI (e.g.  
66 Turkcozer et al., 2016; Schallmo et al., 2018). Thus, in those studies it was not pos-  
67 sible to confidently investigate the activity of V1 along with hMT+. In the current  
68 study we overcome this limitation by presenting the stimuli at the periphery, where it  
69 becomes straightforward to identify regions of interest in different early visual areas.

70 We first conducted a behavioral experiment to ensure that the perceptual effect per-  
71 sists when the stimulus is presented at the periphery. After ensuring that the per-  
72 ceptual effect is present, using fMRI we investigated the neuronal responses within  
73 hMT+ and V1 in response to drifting gratings in varying contrast and size levels.

## 74 **2 Experiment 1: Behavioral Experiment**

### 75 **2.1 Methods**

#### 76 **2.1.1 Participants**

77 Eleven participants, including the authors ZP and GE, participated in the experiment  
78 (seven female; age range: 19-28). All participants reported normal or corrected-to-

79 normal vision, and had no history of neurological or visual disorders. Prior to the  
80 experimental sessions participants gave their written informed consents. Experimental  
81 protocols and procedures were approved by the Human Ethics Committee of Bilkent  
82 University.

### 83 **2.1.2 Stimuli, Experimental Procedures, and Analyses**

84 Visual stimuli were presented on a CRT monitor (HP P1230, 22 inch,  $1600 \times 1200$  res-  
85 olution, 120 Hz). Participants were seated 75 cm from the monitor, and their heads  
86 were stabilized using a chin rest. Responses were collected via a standard computer  
87 keyboard. A gray-scale look-up table, prepared after direct measurements (Spectro-  
88 CAL, Cambridge Research Systems Ltd., UK), was used to ensure the presentation  
89 of correct luminance values. The experimental software was prepared by us using the  
90 Java programming platform.

91 Stimuli were horizontally oriented drifting sine wave gratings (spatial frequency: 1  
92 cycle per degree) weighted by two-dimensional isotropic Gaussian envelopes. Two  
93 size- and contrast-matched gratings were simultaneously and briefly presented on a  
94 mid-gray background ( $40.45 \text{ cd/m}^2$ ) at  $\pm 9.06$  degrees of horizontal eccentricity  
95 (the visual angle between the central fixation and the center of the gratings). Each  
96 grating drifted within the Gaussian envelope (starting phase randomized) at a rate  
97 of 4 degree/s either upward or downward. Participants reported whether or not the  
98 gratings drifted in the same direction, while maintaining fixation at the central fix-  
99 ation mark. After responding, participants received an auditory feedback (auditory  
100 tone of 200 ms duration, 300 Hz for correct and 3800 Hz for incorrect answers). Two  
101 size levels (small: 1.67, large: 8.05 degrees visual angle in diameter) and two contrast  
102 levels (2% and 99% Michelson contrast) were tested (4 experimental conditions in  
103 total). Each condition was blocked in a separate session of 160 trials, and the sessions  
104 were randomly ordered for each participant. Participants completed a short prac-  
105 tice session before beginning an experimental session. Based on their performance  
106 in the practice session, initial presentation duration parameter for the experimental  
107 sessions were selected per participant. For the ensuing trials, presentation duration  
108 was manipulated adaptively with a two interleaved 3-up 1-down staircase procedure.  
109 One staircase started from a relatively short duration, the other started from a longer  
110 duration. There were 80 trials in each staircase.

111 Psychometric functions were fit to the data using the Palamedes toolbox (Kingdom &  
112 Prins, 2010) in Octave (<http://www.octave.org>) for each observer and condition. Du-  
113 ration thresholds (79% success rate) and standard errors were estimated. Repeated-  
114 measures analysis of variances (ANOVA) with two factors (size and contrast) was  
115 performed to compare the thresholds at group level using SPSS Version 19 (SPSS  
116 Inc., Chicago, IL). Additionally, to facilitate drawing links between the behavioral  
117 results and fMRI findings, we calculated “sensitivity” values, defined as  $1/\text{threshold}$ .  
118 Next using the sensitivities for large and small Gabors,  $S_L$  and  $S_S$  respectively, we  
119 computed a size index ( $SI$ ) defined as  $SI = S_L - S_S$ . A positive  $SI$  means increased  
120 sensitivity with increasing size (spatial facilitation), a negative  $SI$  means decreased  
121 sensitivity with increasing size (spatial suppression). We compared the  $SI$  values  
122 to “0” by applying one-sample two-tailed Student’s t-test in SPSS. Also,  $SI$  values  
123 for low and high contrast conditions were compared using two-tailed paired-samples  
124 t-test.

## 125 2.2 Results

126 We measured duration thresholds for accurately judging the drift direction of Gabor  
127 patches presented at the periphery at two size (1.67 and 8.05 degrees) and contrast  
128 levels (2% and 99%). Results are shown in Figure 1. Analyses showed that main  
129 effect of contrast was statistically significant ( $F(1,10) = 12.16$ ,  $p < 0.01$ ) and main  
130 effect of size was close to significance ( $F(1,10) = 4.49$ ,  $p = 0.06$ ). Also, the interaction  
131 between contrast and size was statistically significant ( $F(1,10) = 33.96$ ,  $p < 0.001$ ).  
132 Bonferroni corrected pairwise comparisons showed that for the high-contrast gratings  
133 thresholds increased with size ( $t(10) = -4.1$ ;  $p < 0.01$ ; small stimuli:  $M = 39.83$ ,  $SEM$   
134  $= 2.24$ ; large stimuli:  $M = 99.8$ ,  $SEM = 13.02$ ) whereas for the low-contrast gratings  
135 the thresholds decreased with size ( $t(10) = 6.03$ ;  $p < 0.001$ ; small stimuli:  $M = 107.9$ ,  
136  $SEM = 6.23$ ; large stimuli:  $M = 81.18$ ,  $SEM = 6.22$ ). Consistent with the pairwise  
137 comparisons applied to the raw threshold values, one-sample two-tailed Student's  $t$ -  
138 tests showed that the size index ( $SI$ , see methods) was significantly lower than zero  
139 for the high-contrast gratings ( $t(10) = -5.53$ ;  $p < 0.001$ ;  $M = -0.014$ ,  $SEM = 0.002$ )  
140 whereas it was significantly higher than zero for the low-contrast gratings ( $t(10) =$   
141  $5.56$ ;  $p < 0.001$ ;  $M = 0.003$ ,  $SEM = 0.0006$ ). Also, two-tailed paired-samples Student's  
142  $t$ -tests showed that  $SI$  was significantly higher for low-contrast stimuli compared to  
143 that for high-contrast stimuli ( $t(10) = 6.97$ ;  $p < 0.001$ ). These results clearly replicate  
144 the size-contrast interaction in motion perception when stimuli is presented at the  
145 periphery.

## 146 3 Experiment 2: Functional MRI

### 147 3.1 Materials and Methods

#### 148 3.1.1 Participants

149 Six volunteers (age range: 23-26; mean age: 25; three male) participated in the  
150 experiment, three of whom also participated in the behavioral experiment. All par-  
151 ticipants had normal or corrected-to-normal vision and had no history of neurological  
152 or visual disorders. Participants gave their written informed consents prior to the  
153 fMRI sessions. Experimental protocols and procedures were approved by the Bilkent  
154 University Human Ethics Committee.

#### 155 3.1.2 Data Acquisition & Experimental Setup

156 MR images were collected in the National Magnetic Resonance Research Center (UM-  
157 RAM), Bilkent University on a 3 Tesla Siemens Trio MR scanner (Magnetom Trio,  
158 Siemens AG, Erlangen, Germany) with a 32-channel array head coil. MR sessions  
159 started with a structural run followed by two region of interest (ROI) localizer and  
160 four experimental functional runs, totaling approximately 1 hour in duration. One  
161 localizer run was used to identify the hMT+ region, the other one was for localizing  
162 the sub-regions of hMT+ and V1 that process the input from the visual field that cor-  
163 respond to the position and size of the small Gabors (see below 3.1.3 "Visual Stimuli  
164 & Experimental Design"). Structural data were acquired using a T1-weighted 3-D

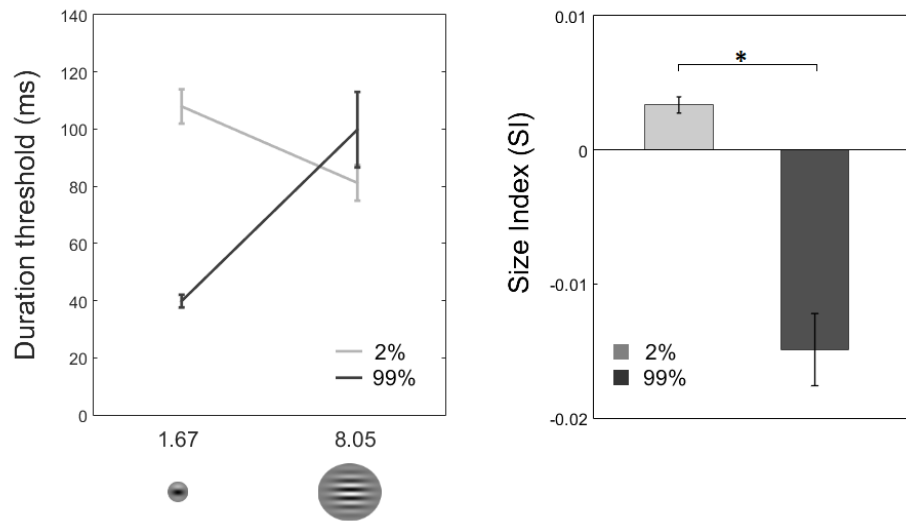


Figure 1: Left plot shows group mean ( $N = 11$ ) of duration thresholds. For low-contrast stimuli, discrimination threshold decreases as size gets bigger. On the contrary, for high-contrast stimuli, discrimination threshold increases as size gets bigger. Right plot shows mean size indices ( $SI$ s) for 2% and 99% contrast levels.  $SI$  is defined as the difference in sensitivity ( $1 / \text{threshold}$ ) between large and small Gabor patches. For low-contrast stimuli,  $SI$  is positive which indicates that sensitivity increases as size gets bigger, i.e. spatial facilitation. On the contrary, for high-contrast stimuli, sensitivity decreases as size gets bigger, i.e. spatial suppression. These results replicate the size-contrast interaction in motion perception when stimuli is presented at the periphery. Error bars represent  $\pm$ SEM. ( $*p < 0.001$ ).

165 anatomical sequence (TR: 2600 ms, spatial resolution:  $1 \text{ mm}^3$  isotropic, number of  
166 slices: 176). Functional images were acquired with a T2\*-weighted gradient-recalled  
167 echo-planar imaging (EPI) sequence (TR: 2000 ms; TE: 35 ms; spatial resolution:  
168  $3 \times 3 \times 3 \text{ mm}^3$ ; number of slices: 30; slice orientation: parallel to calcarine sulcus).  
169 Visual stimuli were presented on an MR-safe LCD Monitor (TELEMED PMEco, Is-  
170 tanbul, Turkey; 32 inch; resolution:  $1920 \times 1080$ ; vertical refresh: 60 Hz). The monitor  
171 was placed near the rear end of the scanner bore, and viewed by the participants  
172 from a distance of 165 cm via a mirror mounted on the head coil. The stimuli were  
173 generated and presented using Python and the Psychopy package (Peirce, 2009).

### 174 3.1.3 Visual Stimuli & Experimental Design

175 Visual stimuli were drifting Gabor patches as in the behavioral experiment. Two  
176 size (small: 1.67 degree, large: 8.05 degree) and two contrast levels (2% and 99%)  
177 were tested. Due to the limits of the visual display system, gratings were presented  
178 at  $\pm 8.02$  degrees of horizontal eccentricity (was 9.06 degrees in the behavioral  
179 experiment), and drifted with a rate of 6 degree/s (was 4 degree/s in the behavioral  
180 experiment) either upward or downward for the duration of 12 seconds. Both Gabor  
181 gratings drifted in the same direction simultaneously, and alternated direction every

182 two seconds to avoid motion adaptation.

183 A functional run was composed of “active” and “control” blocks, each lasting for 12  
184 seconds. In the active blocks, drifting Gabor patches and a central fixation mark were  
185 presented, whilst in the control blocks, only the fixation mark remained visible. In  
186 alternating active blocks, small and large drifting Gabor patches were shown, each  
187 repeated for 6 times in a run. Contrast level was kept constant within a run. Two  
188 experimental runs were conducted for each contrast level in a session. The runs  
189 started with an initial blank period of 24 seconds to allow hemodynamic response to  
190 reach a steady state. The total duration of a functional run was around 5 minutes.  
191 Figure 2 shows the schematic representation of an experimental run.

192 Both to ensure fixation and to control for spatial attention, participants were asked  
193 to perform a demanding fixation task throughout an entire functional run (all par-  
194 ticipants achieved a mean accuracy rate of over 90%, which was the predetermined  
195 threshold to discard the participant’s data). The color of the fixation mark (0.3-degree  
196 solid square) changed randomly from its original color (gray) to either red or yellow  
197 for the duration of 50 ms at the randomly designated interval between 250 to 1500  
198 ms. The participants’ task was to report the changes in the color of the fixation mark  
199 by pressing the designated button on an MR-safe response button-box (Fiber Optic  
200 Response Devices Package 904, Current Designs).

#### 201 **3.1.4 hMT+ Identification**

202 We identified the hMT+ complex in a separate run using the established methods  
203 in literature (Huk, Dougherty, & Heeger, 2002; Dukelow et al., 2001; Smith, Wall,  
204 Williams, & Singh, 2006). Specifically, we presented the participants fields of moving  
205 dots while acquiring functional MR images. The dot fields were comprised of 100 dots  
206 on a black background within an 8-degree diameter circular aperture. The centers of  
207 the fields were 8.02 degrees to the left and right of the fixation point. The dots moved  
208 in three different trajectories: radial (expanding - contracting), cardinal (left-right;  
209 up-down), and angular (clockwise-counterclockwise). Motion direction changed every  
210 two seconds to prevent adaptation. BOLD responses were collected for three types  
211 of configurations, each presented for 12-seconds: right field dynamic (left static), left  
212 field dynamic (right static), and both fields static. This cycle of the presentation was  
213 repeated eight times in a run. We used general linear model (GLM) to contrast the  
214 BOLD responses during dynamic and static presentations. Voxels that respond more  
215 strongly to contralateral dynamic compared to static stimuli at the ascending tip of  
216 the inferior temporal sulcus were identified as hMT+.

#### 217 **3.1.5 ROI Localization within hMT+ and V1**

218 Within hMT+ and V1, subregions which are selectively more responsive to the visual  
219 field that correspond to the locations of the small Gabors in the experiment were  
220 identified using another localizer run. This localizer run was composed of 12s active  
221 and rest blocks. In the active blocks participants saw drifting Gabor patches whose  
222 size, position and speed matched the small stimuli in the experimental runs. The  
223 contrast of the Gabor patches, however, was 60%. Throughout the entire run partic-  
224 ipants were required to maintain central fixation and perform a demanding fixation

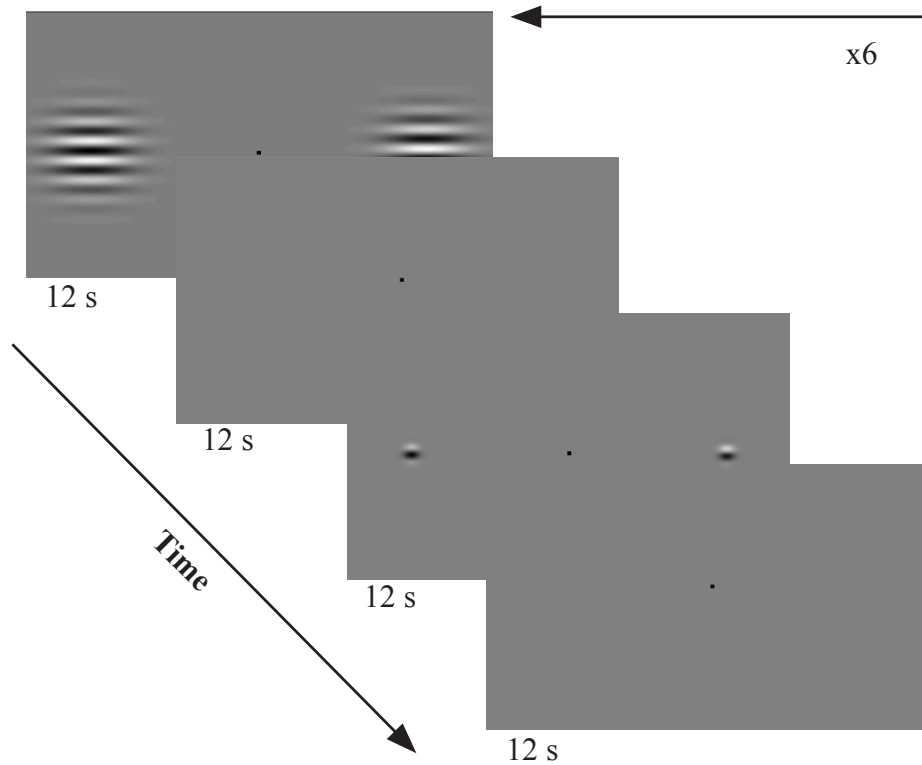


Figure 2: Schematic representation of the visual paradigm of a single cycle in the fMRI experiment. This cycle is repeated for 6 times within a run. Large (8.05 degree) and small (1.67 degree) drifting Gabor patches were presented in alternating active blocks. The contrast was kept constant in a run (2% or 99%), and two runs were conducted for each contrast. Participants were required to keep fixation at the central mark, and perform a demanding fixation task.

225 task as described before. All subsequent analyses were performed on the experimental  
226 data extracted from these ROIs.

227 *ROI within hMT+.* We first created masks using the hMT+ regions functionally  
228 identified as explained before. We then localized the hMT+ ROIs within these masked  
229 regions. Specifically we identified the voxels that are selectively more responsive to  
230 the drifting Gabor patch by contrasting the responses between active and control  
231 blocks using GLM.

232 *ROI within V1.* We identified V1 ROI using the data from the localizer run with the  
233 aid of anatomical landmarks. Specifically, using GLM, cortical regions responding  
234 preferentially more in the active blocks when contrasted with control blocks and  
235 located anteriorly in the calcarine sulcus were identified as the V1 ROI.

### 236 3.1.6 Analyses

237 Anatomical and functional data were preprocessed and analyzed using the BrainVoy-  
238 ager QX software (Brain Innovation, The Netherlands). Preprocessing steps for the  
239 functional images included head motion correction, high-pass temporal filtering and  
240 slice scan time correction. T1-weighted structural images were transformed into the  
241 AC-PC plane, and aligned with the functional images. For each brain, the border be-  
242 tween white matter and cortex was drawn, and an inflated three-dimensional model  
243 of the cortex was generated. Functional maps were projected onto the inflated cortex  
244 to aid the visualization of subsequent analyses. hMT+, and subregions within hMT+  
245 and V1 were identified with GLM as described before using BrainVoyager.

246 Statistical analyses were performed on BOLD responses computed by using the beta  
247 weights calculated with GLM within the ROIs. To further quantify the changes in  
248 BOLD response evoked by an increase in stimulus size, and to draw links with be-  
249 havioral results, we calculated an fMRI size index ( $SI$ ) defined as  $SI = B_L - B_S$ ,  
250 where  $B_L$  and  $B_S$  are the BOLD responses for large and small gratings, respectively.  
251 A positive  $SI$  denotes an increased BOLD response with increase in size (surround  
252 facilitation), and a negative  $SI$  denotes decreased BOLD response (surround suppres-  
253 sion). To compare the BOLD responses at the group level, a 2 X 2 repeated-measures  
254 ANOVA was conducted with contrast level (low and high) and size of the stimuli  
255 (small and large) as factors. We applied one-sample two-tailed Student's t-test to  
256 compare  $SI$ s to zero at each contrast level. We also performed paired samples t-test  
257 to compare the  $SI$ s to each other (low versus high contrast). Statistical analyses were  
258 conducted using JASP Version 0.8.5 (JASP Team, 2018).

## 259 3.2 Results

260 In this experiment, we recorded and analyzed BOLD responses in hMT+ and V1 while  
261 the observers viewed peripherally presented drifting Gabor patches. We compared the  
262 magnitudes of BOLD responses between small and large Gabors at two contrast levels  
263 within predefined ROIs that correspond to the location and size of small stimuli (i.e.  
264 "center"). Because our goal here is to measure the modulatory effect of surround  
265 stimulation on the responses of the neurons whose classical RF centers are inside the  
266 visual space that correspond to the small Gabor patch. BOLD response differences  
267 evoked by presenting the large and small-sized stimuli would highlight the suppressive  
268 or facilitative influence of the surround on the center.

### 269 3.2.1 hMT+

270 Figure 3 (left plot) shows the results from hMT+ ROI. Increasing the size of the stim-  
271 ulti resulted in increased BOLD response, when stimuli had low contrast. Conversely,  
272 increasing size of the stimuli resulted in no change in BOLD response when stimuli  
273 had high contrast. To test whether stimulus contrast and size affect the magnitude of  
274 BOLD response significantly, we applied a  $2 \times 2$  repeated-measures ANOVA for the  
275 contrast (2% and 99%) and size (1.67 and 9.05 degree) as factors. Results revealed a  
276 no main effect of contrast ( $F(1,5) = 3.69, p = 0.113$ ), while we found a main effect of  
277 size ( $F(1,5) = 19.41, p = 0.007$ ), as well as an interaction between size and contrast  
278 ( $F(1,5) = 10.257, p = 0.024$ ) at hMT+.



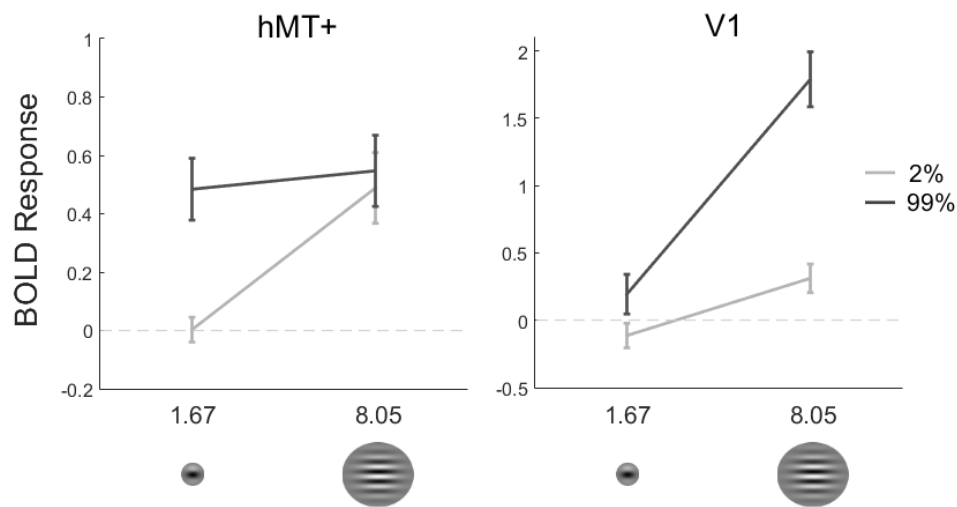


Figure 3: Group mean ( $N = 6$ ) of BOLD responses from hMT+ and V1. BOLD responses to small and large stimuli with high and low contrast were extracted from ROIs that correspond to the location and size of the small stimuli (identified with an independent functional localizer run). Average responses from the ROIs were computed for each subject, then the group mean was calculated. Error bars represent SEM.

279 To further investigate the patterns of results, we computed size indices ( $SI$ s), defined  
280 as the difference between the BOLD responses to large and small Gabors (see Meth-  
281 ods). Figure 4 shows individual  $SI$  values for all participants, as well as the mean  
282  $SI$ . We found that  $SI$  was significantly different (greater) than zero at low contrast  
283 ( $M_{SI} = 0.486$ ,  $SEM = 0.110$ ; one-sample t-test,  $t(5) = 4.42$ ,  $p = 0.007$ ). On the other  
284 hand, the  $SI$  at high contrast was not significantly different than zero ( $M_{SI} = 0.063$ ,  
285  $SEM = 0.066$ ; one-sample t-test,  $t(5) = 0.95$ ,  $p = 0.385$ ). Furthermore, based on the  
286 results in literature (Turkozer et al., 2016; Schallmo et al., 2018), we expected a larger  
287  $SI$  for low contrast compared to high contrast. Indeed, paired sample t-test results  
288 revealed that the  $SI$ s were statistically significantly different, the  $SI$  for low-contrast  
289 being greater than that for the high-contrast ( $t(5) = 3.20$ ,  $p = 0.024$ ).

### 290 3.2.2 V1

291 We next analyzed how the size and contrast of stimuli affect BOLD responses within  
292 the V1 ROI. Figure 3 (right plot) shows the responses for each condition. Results  
293 showed that BOLD responses increased significantly with size both at high and low  
294 contrast conditions. Critically, this increase was greater when the stimuli had high  
295 contrast compared to low contrast. This pattern was inconsistent with the perceptual  
296 effect, and surprisingly it was different than the pattern observed in hMT+. We  
297 applied two-way repeated-measures ANOVA with the contrast (low and high) and size  
298 (small and large) as factors to investigate the role of contrast and size on the BOLD  
299 response. ANOVA revealed a main effect of contrast, ( $F(1,5) = 17.83$ ,  $p = 0.008$ ),  
300 and size ( $F(1,5) = 148.94$ ,  $p < 0.001$ ), as well as the interaction between size and

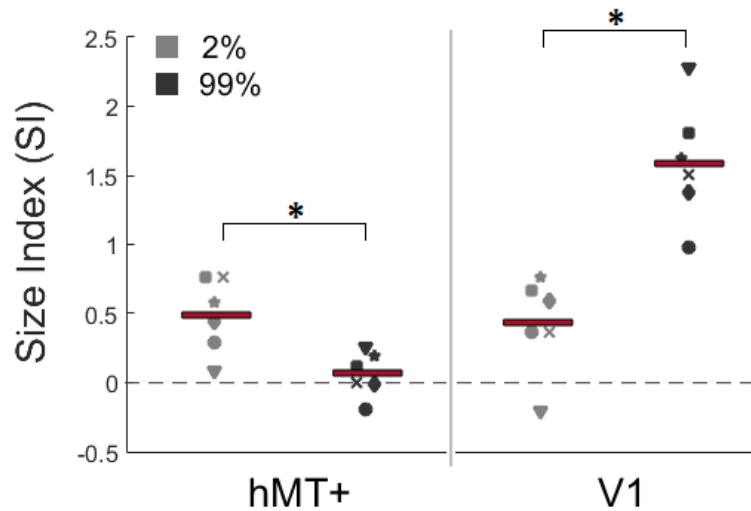


Figure 4: fMRI size indices ( $SI$ ) in hMT+ and V1 for individual participants, and the group mean (red bars).  $SI$  is defined as the difference in BOLD response between large and small stimuli. A positive  $SI$  indicates surround facilitation, a negative one surround suppression. Compare to behavioral results in Figure 1, right plot. ( $*p < 0.05$ ).

301 contrast ( $F(1,5) = 17.96$ ,  $p = 0.008$ ).

302 As we did for the hMT+ data, here too we performed further analyses on  $SI$ s. Figure  
303 4 shows  $SI$ s plotted for individual participants, as well as the group mean. At low  
304 contrast, group mean of  $SI$  was significantly greater than zero ( $M_{SI} = 0.425$ ,  $SEM$   
305  $= 0.143$ ; one sample t-test,  $t(5) = 2.98$ ,  $p = 0.031$ ). Average  $SI$  value was positive at  
306 high contrast, as well ( $M_{SI} = 1.596$ ,  $SEM = 0.178$ ; one sample t-test,  $t(5) = 8.98$ ,  
307  $p < 0.001$ ). Furthermore, we performed paired-sample t-test, and found that the  $SI$   
308 value for low contrast was significantly lower than the  $SI$  value for high contrast ( $t(5)$   
309  $= -4.24$ ,  $p = 0.004$ ).

### 310 3.3 Linking Behavioral and fMRI Results

311 To draw a link between the behavioral and fMRI results, we first assume that the  
312 amount of neuronal responses can be approximated by a monotonically increasing  
313 function of sensitivity (1/duration threshold). Further, we assume a linear relation  
314 between neuronal activity and the BOLD fMRI response (see e.g. Boynton, Demb,  
315 Glover, & Heeger, 1999). Thus, if an area is involved in the processes related to the  
316 perceptual effect, we expect an increase in fMRI BOLD response in that area as the  
317 behavioral sensitivity gets better. Comparing the duration thresholds (Figure 1, left  
318 plot), and BOLD responses (Figure 3), we see that the hMT+ activity captures the  
319 behavioral results for low contrast stimulus; for high contrast stimulus, however, there  
320 seems to be a slight disagreement (see Discussion for possible reasons for this). The  
321 V1 activity, on the other hand, is completely inconsistent with the behavioral results.

322 This pattern becomes more clear when the size indexes are compared (Figure 1 right  
323 plot, and Figure 4). Overall the response patterns in hMT+, but not in V1, agree  
324 well with the behavioral results.

## 325 4 Discussion

326 We demonstrated that surround modulation found in hMT+ complex, but not in V1,  
327 agrees with the size-contrast interaction in motion perception. First in a behavioral  
328 experiment we measured the temporal thresholds for successfully detecting the direc-  
329 tion of motion of drifting Gabor patches presented at the periphery. We found that  
330 the thresholds decreased with size (i.e. increased sensitivity, spatial facilitation) if  
331 the target has low contrast (2%). Conversely, we found that the thresholds increased  
332 with size (i.e. decreased sensitivity, spatial suppression) if the target has high con-  
333 trast (99%). These results were in good agreement with literature (e.g., Tadin et al.,  
334 2011). Next, we recorded BOLD fMRI responses while participants viewed high- and  
335 low-contrast, small and large drifting Gabor patches presented at the periphery. In  
336 hMT+ we found that BOLD responses significantly increased with size for the low  
337 contrast Gabor patches (surround facilitation), and remained unchanged for the high  
338 contrast Gabor patches. In V1, however, BOLD responses increased for both high  
339 and low contrast Gabor patches; and critically the increase was stronger for the high  
340 contrast stimuli. Overall, we contend that the activity patterns in hMT+, but not in  
341 V1, reflect the perceptual effect.

342 The BOLD responses we found in hMT+ for high-contrast stimuli may at first seem  
343 inconsistent with the behavioral results. Specifically, behaviorally the sensitivity de-  
344 creases with size, whereas the BOLD responses in hMT+ remains unchanged instead  
345 of also decreasing. We believe that there may be several reasons for this. Firstly,  
346 because we were interested in the modulatory effects of the surround on the neurons  
347 that were processing the center, we sought to identify regions of cortex that process  
348 the part of the visual field that correspond to the position and size of the small stimuli  
349 (center). To do this, we identified the voxels that responded more strongly to small  
350 drifting Gabors compared to a blank screen. Given the large number of neurons in  
351 an fMRI voxel (about a million), and their possible heterogeneity, it is likely that  
352 neurons with larger receptive fields that respond directly to both small (center) and  
353 large (center + surround) stimuli could have been included in our ROI. This scenario  
354 is especially likely in hMT+ where RF sizes are usually much larger. Those neurons  
355 could have responded more strongly during the presentation of the large stimulus,  
356 hiding the effect of response reduction in the neurons that respond only to the cen-  
357 ter. Alternatively, the observed discrepancy could be because of saturation of BOLD  
358 responses at high contrasts (Tootell et al., 1995; Buracas & Boynton, 2007), or be-  
359 cause of the commonly observed non-linear relations between the stimulus energy and  
360 BOLD responses (Logothetis & Wandell, 2004). Considering these possibilities, we  
361 believe that assessing the neuronal responses using the Size Index (*SI*) is more ap-  
362 propriate, because *SI* better represents the overall pattern of interaction between size  
363 and contrast. Based on the *SIs*, we see that hMT+ activity agree with the perceptual  
364 sensitivity. The pattern in V1, however, is completely at odds with the perceptual  
365 effect.

366 Surround modulation in hMT+ has been previously claimed to underlie the size-

367 contrast interaction in motion perception (Tadin et al., 2003). Two recent neuroimag-  
368 ing results landed support for this hypothesis (Turkozer et al., 2016; Schallmo et al.,  
369 2018). Specifically, in both studies, the response patterns in hMT+ were found to re-  
370 flect the size-contrast interaction. The neuronal activity in hMT+, however, could be  
371 inherited from earlier areas, particularly the primary visual cortex (V1). To examine  
372 this possibility required investigating the responses in earlier areas. This could not be  
373 done previously with fMRI due to methodological limitations. In both studies stimuli  
374 were presented at the fovea (Turkozer et al., 2016; Schallmo et al., 2018). This part  
375 of the visual field is mapped onto the so-called foveal confluence at the occipital pole,  
376 where it is difficult to reliably distinguish V1, V2, and V3 using the standard retino-  
377 topic mapping techniques (e.g. Engel, Glover, & Wandell, 1997). Therefore neither of  
378 the previous studies could argue strongly that the size-contrast interaction in motion  
379 perception did not originate in earlier visual areas (Turkozer et al., 2016; Schallmo  
380 et al., 2018). In the present study we have successfully avoided this limitation by  
381 presenting the stimuli at the periphery, which allowed us to confidently localize ROIs  
382 in V1, as well as hMT+.

383 Our results in hMT+ agree with the results of the two studies introduced above  
384 (Turkozer et al., 2016; Schallmo et al., 2018). Turkozer et al. (2016) did not report  
385 results from other visual areas, but Schallmo et al. (2018) found surround suppression  
386 for all contrast levels in early visual cortex (EVC), which is defined as the sum of V1,  
387 V2 and V3 in the foveal confluence. This seems to stand in contradiction to our  
388 findings. The major difference between our study and Schallmo et al. (2018) was the  
389 position of the stimuli (periphery vs. fovea). Owing to the methodological limitations  
390 described before, Schallmo et al. (2018) could report only the aggregated activity of  
391 V1, V2, and V3 (EVC). Reporting the averaged activity from EVC, however, may  
392 have obscured the facilitation in V1. This is probable, because suppression has been  
393 shown to be progressively stronger in V2 and V3 than in V1 (Zenger-Landolt &  
394 Heeger, 2003). Alternatively, the difference in surround modulation at the fovea and  
395 periphery in V1 could have caused the differences between our results and Schallmo  
396 et al. (2018) results (Xing & Heeger, 2000). This will be elaborated further below.

397 In their study, Schallmo et al. (2018) argue that a single computational mechanism,  
398 namely divisive normalization (Heeger, 1992; Reynolds & Heeger, 2009; Carandini &  
399 Heeger, 2012), can successfully account for both surround facilitation and suppression  
400 (also see Schallmo et al., 2019). Borrowing this idea, we have also tested how well  
401 the divisive normalization model could explain our results with the estimated RF  
402 center and the surround size at the periphery. We found good agreement between  
403 the model and hMT+ responses. But the model could not predict the responses in  
404 V1. This could be because of the complex role of eccentricity in mediating surround  
405 modulation in V1. For example, divisive normalization may be a good choice to for-  
406 mulate the effects of feedforward mechanisms, but surround modulation, particularly  
407 at the periphery, may involve more than these feedforward connections, and include  
408 feedback and horizontal mechanisms, as well (Nurminen & Angelucci, 2014; Nurmi-  
409 nen, Merlin, Bijanzadeh, Federer, & Angelucci, 2018). There is little doubt that as  
410 part of a vastly interconnected network V1 receives feedback from other visual areas  
411 (Shao & Burkhalter, 1996), including those for motion processing (Hupé et al., 1998;  
412 Ponce, Lomber, & Born, 2008; Paffen, van der Smagt, te Pas, & Verstraten, 2005).  
413 Such top-down influences may need to be factored in a computational model for a  
414 fuller understanding of surround modulation in V1. Moreover, such a model should  
415 also be able to incorporate the effects of attention, which is shown to interact with

416 eccentricity in surround modulation in V1 (Reynolds & Heeger, 2009).

417 It is worth noting that there seems to be a disagreement between the surround mod-  
418 ulation in V1 found using cell-recording methods on animal models and fMRI on  
419 humans. Using cell-recording techniques, surround suppression has been routinely  
420 shown for high-contrast stimuli in V1 (e.g. Jones, Grieve, Wang, & Sillito, 2001; An-  
421 gelucci & Shushruth, 2013; Angelucci et al., 2017). Similar to our findings here, Press,  
422 Brewer, Dougherty, Wade, and Wandell (2001) reported only facilitation in V1 us-  
423 ing flickering checkerboard patterns (but see Zenger-Landolt & Heeger, 2003). These  
424 inconsistencies may indicate a species difference, or differences between the methods  
425 (i.e. fMRI vs. cell recording, and differences in experimental procedures). Reconcil-  
426 ing these differences would require careful and systematic comparison of results using  
427 the same experimental conditions.

## 428 **5 Conclusion**

429 Our results provide further evidence that size-contrast interaction in motion percep-  
430 tion likely originates in hMT+. In a broader context our results show that surround  
431 modulation can be distinct in different components of the biological visual system.  
432 The good agreement of hMT+ response patterns in periphery (this study) and fovea  
433 (Turkozer et al., 2016; Schallmo et al., 2018) suggests that surround modulation in  
434 hMT+ is relatively stable across eccentricity. But this may not be true for the V1  
435 neurons. Overall, these results show that V1 and hMT+ may exhibit different char-  
436 acteristics of surround modulation.

## 437 **6 Author Contributions**

438 ZP and HB conceived the original idea. ZP designed, implemented and conducted  
439 the behavioral experiments. ZP and GE designed; GE implemented and conducted  
440 the fMRI experiments. GE, ZP, and HB wrote the manuscript.

## 441 **7 Declaration of Interest**

442 Declarations of interest: none

## 443 **8 Acknowledgments**

444 We thank Michael-Paul Schallmo and his colleagues for letting us use their code for  
445 testing the divisive normalization model. We thank Halide Bilge Turkozer for her  
446 contribution to designing the experimental paradigm. Author ZP was supported by  
447 TÜBİTAK (National Scholarship Program for PhD Students, scholarship id: 2211-  
448 E).

## 449 References

- 450 Angelucci, A., Bijanzadeh, M., Nurminen, L., Federer, F., Merlin, S., & Bressloff,  
451 P. C. (2017). Circuits and Mechanisms for Surround Modulation in Visual  
452 Cortex. *Annual Review of Neuroscience*, *40*(1), 425–451. doi: 10.1146/annurev-  
453 neuro-072116-031418
- 454 Angelucci, A., & Shushruth, S. (2013). Beyond the classical receptive field: surround  
455 modulation in primary visual cortex. In J. S. Werner & L. M. Chalupa (Eds.),  
456 *The new visual neurosciences* (pp. 425–444). Cambridge, MA: The MIT Press.
- 457 Boynton, G. M., Demb, J. B., Glover, G. H., & Heeger, D. J. (1999, jan). Neu-  
458 ronal basis of contrast discrimination. *Vision Research*, *39*(2), 257–269. doi:  
459 10.1016/S0042-6989(98)00113-8
- 460 Buracas, G. T., & Boynton, G. M. (2007). The effect of spatial attention on contrast  
461 response functions in human visual cortex. *Journal of Neuroscience*, *27*(1),  
462 93–97.
- 463 Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural compu-  
464 tation. *Nature Reviews Neuroscience*, *13*(1), 51–62. doi: 10.1038/nrn3136
- 465 Dukelow, S. P., DeSouza, J. F., Culham, J. C., van den Berg, A. V., Menon, R. S.,  
466 & Vilis, T. (2001). Distinguishing subregions of the human mt+ complex using  
467 visual fields and pursuit eye movements. *Journal of Neurophysiology*, *86*(4),  
468 1991–2000.
- 469 Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization  
470 in human visual cortex and the spatial precision of functional MRI. *Cerebral*  
471 *Cortex*, *7*(2), 181–192.
- 472 Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual*  
473 *neuroscience*, *9*(2), 181–197.
- 474 Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional  
475 subdivision of human areas mt and mst. *Journal of Neuroscience*, *22*(16), 7195–  
476 7205.
- 477 Hupé, J., James, A., Payne, B., Lomber, S., Girard, P., & Bullier, J. (1998). Cortical  
478 feedback improves discrimination between figure and background by v1, v2 and  
479 v3 neurons. *Nature*, *394*(6695), 784.
- 480 JASP Team. (2018). *JASP (Version 0.8.5)[Computer software]*. Retrieved from  
481 <https://jasp-stats.org/>
- 482 Jones, H., Grieve, K., Wang, W., & Sillito, A. (2001). Surround suppression in  
483 primate v1. *Journal of neurophysiology*, *86*(4), 2011–2028.
- 484 Kingdom, F. A. A., & Prins, N. (2010). *Psychophysics: a practical introduction*.  
485 Academic Press.
- 486 Liu, L. D., Haefner, R. M., & Pack, C. C. (2016). A neural basis for the spatial sup-  
487 pression of visual motion perception. *Elife*, *5*, e16167. doi: 10.7554/eLife.16167
- 488 Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD Sig-  
489 nal. *Annual Review of Physiology*, *66*(1), 735–769. doi: 10.1146/an-  
490 nurev.physiol.66.082602.092845
- 491 Nurminen, L., & Angelucci, A. (2014). Multiple components of surround modula-  
492 tion in primary visual cortex: Multiple neural circuits with multiple functions?  
493 *Vision Research*, *104*, 47–56. doi: 10.1016/j.visres.2014.08.018
- 494 Nurminen, L., Merlin, S., Bijanzadeh, M., Federer, F., & Angelucci, A. (2018, dec).  
495 Top-down feedback controls spatial summation and response amplitude in pri-

- 496 mate visual cortex. *Nature Communications*, 9(1), 2281. doi: 10.1038/s41467-  
497 018-04500-5
- 498 Paffen, C. L., van der Smagt, M. J., te Pas, S. F., & Verstraten, F. A. (2005). Center-  
499 surround inhibition and facilitation as a function of size and contrast at multiple  
500 levels of visual motion processing. *Journal of Vision*, 5(6), 8–8.
- 501 Peirce, J. (2009). Generating stimuli for neuroscience using psychopy. *Frontiers in*  
502 *Neuroinformatics*, 2, 10. doi: 10.3389/neuro.11.010.2008
- 503 Ponce, C. R., Lomber, S. G., & Born, R. T. (2008). Integrating motion and depth  
504 via parallel pathways. *Nature neuroscience*, 11(2), 216.
- 505 Press, W. A., Brewer, A. A., Dougherty, R. F., Wade, A. R., & Wandell, B. A. (2001,  
506 May). Visual areas and spatial summation in human visual cortex. *Vision*  
507 *Research*, 41(10-11), 1321–1332. doi: 10.1016/S0042-6989(01)00074-8
- 508 Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention.  
509 *Neuron*, 61(2), 168–185.
- 510 Schallmo, M. P., Kale, A. M., Millin, R., Flevaris, A. V., Brkanac, Z., Edden, R. A., ...  
511 Murray, S. O. (2018). Suppression and facilitation of human neural responses.  
512 *eLife*, 7, 1–23. doi: 10.7554/eLife.30334
- 513 Schallmo, M.-P., Kolodny, T., Kale, A. M., Millin, R., Flevaris, A. V., Edden, R. A.,  
514 ... Murray, S. O. (2019). Weaker neural suppression in autism. *bioRxiv*,  
515 645846.
- 516 Shao, Z., & Burkhalter, A. (1996). Different balance of excitation and inhibition  
517 in forward and feedback circuits of rat visual cortex. *Journal of Neuroscience*,  
518 16(22), 7353–7365.
- 519 Smith, A., Wall, M., Williams, A., & Singh, K. D. (2006). Sensitivity to optic flow  
520 in human cortical areas mt and mst. *European Journal of Neuroscience*, 23(2),  
521 561–569.
- 522 Tadin, D., Lappin, J. S., Gilroy, L. A., & Blake, R. (2003). Perceptual consequences  
523 of centre-surround antagonism in visual motion processing. *Nature*, 424(6946),  
524 312.
- 525 Tadin, D., Silvanto, J., Pascual-Leone, A., & Battelli, L. (2011). Improved motion  
526 perception and impaired spatial suppression following disruption of cortical area  
527 mt/v5. *Journal of Neuroscience*, 31(4), 1279–1283.
- 528 Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J.,  
529 ... Belliveau, J. W. (1995). Functional analysis of human mt and related  
530 visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*,  
531 15(4), 3215–3230.
- 532 Turkozer, H. B., Pamir, Z., & Boyaci, H. (2016). Contrast affects fmri activity in mid-  
533 dle temporal cortex related to center-surround interaction in motion perception.  
534 *Frontiers in Psychology*, 7, 454. doi: 10.3389/fpsyg.2016.00454
- 535 Xing, J., & Heeger, D. J. (2000). Center-surround interactions in foveal and peripheral  
536 vision. *Vision Research*, 40(22), 3065–3072.
- 537 Zenger-Landolt, B., & Heeger, D. J. (2003). Response suppression in v1 agrees with  
538 psychophysics of surround masking. *Journal of Neuroscience*, 23(17), 6884–  
539 6893.