**A backward progression of attentional effects in the ventral stream**

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1073/pnas.0907658106">http://dx.doi.org/10.1073/pnas.0907658106</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>National Academy of Sciences</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Thu Jan 24 01:31:28 EST 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/61394">http://hdl.handle.net/1721.1/61394</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
A backward progression of attentional effects in the ventral stream

Elizabeth A. Buffalo,1, Pascal Fries,2,3, Rogier Landman,4, Huaiou Liang5, and Robert Desimone

*Yerkes National Primate Research Center and Department of Neurology, Emory University School of Medicine, Atlanta, GA 30329; †Donors Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, 6525 EN Nijmegen, The Netherlands; ‡Ernst Strungmann Institute in Cooperation with Max Planck Society, 60528 Frankfurt, Germany; §McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139; and †School of Biomedical Engineering, Drexel University, Philadelphia, PA, 19104

Edited by Charles G. Gross, Princeton University, Princeton, NJ, and approved November 4, 2009 (received for review July 12, 2009)

The visual processing of behaviorally relevant stimuli is enhanced through top-down attentional feedback. One possibility is that feedback targets early visual areas first and the attentional enhancement builds up at progressively later stages of the visual hierarchy. An alternative possibility is that the feedback targets the higher-order areas first and the attentional effects are communicated “backward” to early visual areas. Here, we compared the magnitude and latency of attentional enhancement of firing rates in V1, V2, and V4 in the same animals performing the same task. We found a reverse order of attentional effects, such that attentional enhancement was larger and earlier in V4 and smaller and later in V1, with intermediate results in V2. These results suggest that attentional mechanisms operate via feedback from higher-order areas to lower-order ones.

Results

We recorded from 134, 93, and 116 neurons in V1, V2, and V4, respectively, in two rhesus monkeys performing a task of directed spatial attention. On each trial, two slowly drifting achromatic gratings were presented for several seconds in the parafoveal visual field, and on alternating blocks of trials, the monkey attended to the stimulus either inside or outside the recorded neuron’s RF (Fig. 1F). Because of the blocked trial design, the monkey knew the location of the behaviorally relevant stimulus before it appeared, although we cannot say when the animal actually began to attend to that location. The monkey was rewarded for releasing a bar when it detected a subtle color change in the attended stimulus, while ignoring any change in the unattended stimulus. The color change could occur at any time

Author contributions: E.A.B. and R.D. designed research; E.A.B., P.F., and R.L. performed research; E.A.B., R.L., and H.L. analyzed data; and E.A.B. and R.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence should be addressed. E-mail: elizabeth.buffalo@emory.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0907658106/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.0907658106

PNAS | January 5, 2010 | vol. 107 | no. 1 | 361–365
Latency of attentional modulation of firing rate in areas V1, V2, and V4. (A–C) Red traces represent spike density plots of the average response in each area with attention directed into the neuron’s RF (as illustrated by the “spotlight” of attention in the drawing in the top panel of F). Blue traces represent responses with attention directed out of the RF (F, Bottom). Responses are shown for area V1 (A), V2 (B), and V4 (C). Responses were aligned to the stimulus onset (0) and were smoothed with a Gaussian window of 30 ms. Shaded areas represent SEM. Vertical black lines represent the onset of the attentional modulation in each area. (D) Distributions of the latency of attentional modulation are shown for each area. Red: V4; Blue: V2; Green: V1. Arrows denote median latencies for each visual area. (E) Cumulative distribution plot of the latency of attentional modulation in each area; colors represent the three areas as in the previous plot. (F) Cartoon depicting the stimuli in the blocked design task.

We then calculated the latency of attentional effects for each cell individually (see Methods). The distributions of individual latencies for all cells with significant attentional effects at any time are shown in Fig. 1D, and the cumulative distributions are shown in Fig. 1E. The median latencies in V4 (n = 101), V2 (n = 37), and V1 (n = 24) were 270 ms, 780 ms, and 1375 ms, respectively. A one-way nonparametric ANOVA revealed a significant main effect across areas ($\chi^2(2) = 38.64; P < 0.0001$), and post hoc analyses (Tukey-Kramer) revealed significant differences for all pairwise combinations except for the comparison between areas V1 and V2 (all significant P values < 0.05). Some cells in V4 (8% of total cells) showed significant effects of attention before the onset of the stimulus. This presumably occurred because in the blocked trial design the animals knew the location where the relevant stimulus would appear in advance.

To compare the magnitude of attentional effects across areas, we computed a contrast index of attentional effects from the period 1,000–3,000 ms after stimulus onset in the attention IN vs. attention OUT condition. As described above, this interval avoided the period just after stimulus onset when cells showed differences in attentional latencies. The attention index was computed according to the following formula: attention IN − attention OUT/attention IN + attention OUT for all cells. Fig. 2 shows the distribution of the index for all recorded cells across the three areas, which had a median value of 0.09 in V4, followed by 0.05 in V2 and 0.02 in V1 (sign test, all P values < 0.05). These indices correspond to an overall increase in firing rate with attention of 23%, 19%, and 5%, respectively, in the three areas. Thus, the magnitude of attentional effects across areas follows the same “backwards” trend as the distribution of latencies. A small number of cells in each area showed significant reductions between 500 ms and 5,000 ms after stimulus onset, thus requiring the monkey to sustain attention for a long period. Neuronal responses were compared during trials when attention was directed to the stimulus located inside (attention IN) vs. outside (attention OUT) the RF. The sensory conditions were identical across attention conditions.

The percentages of cells showing sustained attentional effects on their firing rates in V4, V2, and V1 were 73% (n = 85), 44% (n = 41), and 26% (n = 35), respectively. To then estimate the earliest effects of attention in the population, we calculated the normalized average response histograms for all of these cells with attentional effects (positive or negative), which are shown in Fig. 1. The population latencies in V4, V2, and V1 were 170 ms, 440 ms, and 860 ms, respectively (see Methods). A bootstrap method was used to sample (n-1 latencies per sample) the latencies in the total cell population histogram, which were 65 ms, 40 ms, and 35 ms in V4, V2, and V1, respectively.

We then calculated the latency of attentional effects for each cell individually (see Methods). The distributions of individual latencies for all cells with significant attentional effects at any time are shown in Fig. 1D, and the cumulative distributions are shown in Fig. 1E. The median latencies in V4 (n = 101), V2 (n = 37), and V1 (n = 24) were 270 ms, 780 ms, and 1375 ms, respectively. A one-way nonparametric ANOVA revealed a significant main effect across areas ($\chi^2(2) = 38.64; P < 0.0001$), and post hoc analyses (Tukey-Kramer) revealed significant differences for all pairwise combinations except for the comparison between areas V1 and V2 (all significant P values < 0.05). Some cells in V4 (8% of total cells) showed significant effects of attention before the onset of the stimulus. This presumably occurred because in the blocked trial design the animals knew the location where the relevant stimulus would appear in advance.

To compare the magnitude of attentional effects across areas, we computed a contrast index of attentional effects from the period 1,000–3,000 ms after stimulus onset in the attention IN vs. attention OUT condition. As described above, this interval avoided the period just after stimulus onset when cells showed differences in attentional latencies. The attention index was computed according to the following formula: attention IN − attention OUT/attention IN + attention OUT for all cells. Fig. 2 shows the distribution of the index for all recorded cells across the three areas, which had a median value of 0.09 in V4, followed by 0.05 in V2 and 0.02 in V1 (sign test, all P values < 0.05). These indices correspond to an overall increase in firing rate with attention of 23%, 19%, and 5%, respectively, in the three areas. Thus, the magnitude of attentional effects across areas follows the same “backwards” trend as the distribution of latencies. A small number of cells in each area showed significant reductions between 500 ms and 5,000 ms after stimulus onset, thus requiring the monkey to sustain attention for a long period. Neuronal responses were compared during trials when attention was directed to the stimulus located inside (attention IN) vs. outside (attention OUT) the RF. The sensory conditions were identical across attention conditions.

The percentages of cells showing sustained attentional effects on their firing rates in V4, V2, and V1 were 73% (n = 85), 44% (n = 41), and 26% (n = 35), respectively. To then estimate the earliest effects of attention in the population, we calculated the normalized average response histograms for all of these cells with attentional effects (positive or negative), which are shown in Fig. 1. The population latencies in V4, V2, and V1 were 170 ms, 440 ms, and 860 ms, respectively (see Methods). A bootstrap method was used to sample (n-1 latencies per sample) the latencies in the total cell population histogram, which were 65 ms, 40 ms, and 35 ms in V4, V2, and V1, respectively.
in response with attention to the RF stimulus (V1: 4% of total cells; V2: 13%; and V4: 7%).

We also examined whether there was a relationship between the magnitude and the latency of attentional effects. As shown in Fig. 3, across the population, there was no systematic relationship between the latency of attentional modulation and the magnitude of the attentional effect for individual cells (Pearson correlation, \( r = -0.15, P > 0.05 \)). Additionally, there were no systematic differences in the monkeys’ eye position across attention conditions that could have contributed to the attentional modulation (Fig. S2). We tested for transient attentional effects on just the ‘peak’ of the initial sensory response to the stimulus in V1 and V2, similar to the analysis of McAlonan et al. (21) in the LGN and thalamic reticular nucleus (TRN), but, on average, there was no attentional enhancement of the initial peak response in either area.

**Discussion**

Is the enhanced processing of attended stimuli mediated by feedback that targets early visual areas first and is passed on to higher-order areas, or by feedback that targets higher-order areas first and is fed back to earlier areas with diminished effect and longer latency? Here, we found that neuronal responses were enhanced by spatially selective attention all along the ventral stream, but these effects were earlier and larger in V4 and progressively later and smaller in V2 and V1. The results in all three areas were obtained in the same monkeys performing the same task, which minimized the variance due to task and monkey variables. The results thus support the idea of a “backwards” progression of attentional feedback within the ventral stream.

We cannot say whether the backward progression of effects is actually mediated by direct feedback anatomic pathways from V4 to V2 to V1. In principle, all three areas could receive inputs from the same higher-level attentional structures, but with weaker and later feedback to V2 and V1. However, given the large proportion of cells showing attentional effects in V4, it seems likely that feedback from these cells would have a direct effect on visual responses in V2 and that cells showing an attentional effect in V2 would have an impact on V1. Cells in both superficial and deep layers of V4 and V2 project heavily back to the superficial and deep layers of V2 and V1, respectively (28–31). V4 itself receives direct feedback projections from frontal eye field and the posterior parietal cortex, both of which are implicated in attentional control (9–11), but this feedback projection is much weaker in V2 and may be nonexistent in V1 (32). If V1 (or the LGN) receives direct feedback from structures mediating attentional control, at high spatial specificity, it is not clear what would be the anatomic source of this feedback (see discussion of the LGN, below). Finally, it should be noted that lesions of V4 cause attentional impairments (33, 34), consistent with the idea that it plays a key role in mediating attentional feedback in the ventral stream.

The backward progression of attentional effects from V4 to V1 is consistent with the findings of Mehta et al. (24), who used a cross-modal attention design with a diffuse visual stimulus in monkeys, as well as the findings of Martinez et al. (35) based on the timing of attentional effects on event-related potentials in humans. Other monkey studies have found smaller attentional effects in V1 than in V4 (3, 4, 6), although the latencies of attentional effects were not reported in those studies. Human neuroimaging studies also typically find smaller attentional effects in V1 compared with higher-order areas.

The “normalization model of attention” (36) suggests that the effects of attention on a stimulus-evoked response can vary between a contrast gain and a response gain function, depending on the relationship between the size of the attended stimulus and the size of the receptive field. In the present study, the size of the stimulus was fixed but the relationship between the stimulus size and receptive field size varied among V1, V2, and V4, which may have influenced the attentional effects. However, with the high-contrast stimulus used in the present study, the model does not
predict the relative magnitudes of attentional effects that we found across areas, nor does it explain the longer attentional latencies in earlier visual areas found here. The backward anatomic progression of attentional effects seems a more likely explanation for the present results.

We are more confident of the relative order rather than absolute latencies of attentional effects that we found across the areas, because differences in the type of task, task difficulty, and stimulus variables will likely cause variations in absolute latencies. Local field potential (LFP) measures have been shown to have shorter absolute latencies (24). Future investigations of the relative sensitivity of single units vs. LFPs are needed to clarify this issue. However, a long latency of attentional effects in V1 could explain why some studies found little or no significant average enhancement of response with attention in V1, because these studies generally used short stimulus presentations (3, 6, 37, 38). Likewise, a long latency attentional effect in V1 provides an obvious explanation for why some imaging studies find robust effects of attention in V1 (13). The blood oxygen level–dependent signal in functional MRI averages activity changes over long intervals, which could easily include the late attentional effects observed in V1.

Recently, McAlonan et al. (21) found that attention serves to modulate visual signals at very short latencies in the LGN and the TRN. However, the short-latency attention effects found in the TRN and LGN were very transient, lasting no more than 100 ms. The LGN also showed more sustained attention effects at a longer latency of approximately 250 ms. The authors suggested that the short-latency, transient effects were mediated by top-down signals affecting TRN in advance of the target stimulus, because the monkey knew the location of the expected target in advance. Just as in the present study, attention was directed to the target location long before stimulus onset, which could have contributed to these very early effects in the TRN and LGN.

They proposed that the transient inputs from TRN caused enhanced responses in the LGN (through a reduction of inhibition), and that these enhanced responses were then passed on to V1. However, because the attentional effects in TRN were so short lived, the longer latency attentional effects in the LGN must have resulted from a different mechanism. An obvious possibility is direct feedback from V1. If so, the long-latency attention effects found in V1 to V2 must have resulted from a different mechanism. An obvious possibility is that this late attentional modulation in V1 and V2 is important in tasks for which very fine details are important. Coarse information at the attentional focus might be processed in tasks requiring discrimination of fine details. If so, future studies might detect that attention affects the processing of fine details in V4 at long latencies.

What is the purpose of the attentional modulation of V1 and V2 cells if this modulation occurs later than in V4? One possibility is that this late attentional modulation in V1 and V2 is important in tasks for which very fine details are important. Coarse information at the attentional focus might be processed in tasks requiring discrimination of fine details. If so, future studies might detect that attention affects the processing of fine details in V4 at long latencies.

Methods

Surgical Procedures. Experiments were performed in areas V1, V2, and V4 in four hemispheres of two male rhesus monkeys (Macaca mulatta), which followed the guidelines of the National Institutes of Health.

Two adult male rhesus monkeys were surgically implanted with a head post, a scleral eye coil, and recording chambers. Surgery was conducted under deep sedation with isoflurane anesthesia, and antibiotics and analgesics were administered postoperatively. Preoperative MRI was used to identify the stereotaxic coordinates of V1, V2, and V4. V4 recording chambers were placed over the prelunate gyrus. Additional plastic recording chambers were used for V1 and V2 recordings, centered 15 mm lateral and 15 mm dorsal to the occipital pole. The skull remained intact during the initial surgery, and small holes (3 mm in diameter) were later drilled within the recording chambers under ketamine anesthesia and xylazine analgesic to expose the dura for electrode penetrations.

Recording Techniques. In each recording session, four to eight tungsten microelectrodes (impedances of 1 to 2 MΩ) were advanced separately at a very slow rate (1.5 um/s) to minimize deformation of the cortical surface by the electrode ("dimpling"). Recordings were performed in each of the three visual areas in separate recording sessions. Electrode tips were separated by 650 or 900 μm. Data amplification, filtering, and acquisition were done with a Multichannel Acquisition Processor (Plexon). The signal from each electrode was passed through a headstage with unit gain and an output impedance of 240 Ω. The signals were filtered with a passband of 100–8,000 Hz, further amplified, and digitized with 40 kHz. A threshold was set interactively, and spike waveforms were stored for a time window from 150 μs before to 700 μs after threshold crossing. The threshold clearly separated spikes from noise but was chosen to include multunit activity. Off-line, we performed a principal component analysis of the waveforms and plotted the first against the second principal component. Those waveforms that corresponded to artifacts were excluded. Spikes were sorted into single units. When this was not possible, multunits were accepted. The times of threshold crossing were kept and downsampled to 1 kHz. RF position and neuronal stimulus selectivity were as expected for the target part of each visual area.

Visual Stimulation and Experimental Paradigm. Stimuli were presented on a 17-inch cathode ray tube monitor 0.57 m from the monkey’s eyes that had a resolution of 800 × 600 pixels and a screen refresh rate of 120 Hz interlaced. Stimulus generation and behavioral control were accomplished with the CORTEX software package (www.cortex.salk.edu). A trial began when the monkey touched a bar and directed its gaze within 0.7° of the fixation spot on the computer screen (Fig. 1F). After achieving fixation for 300 ms, the stimuli were presented. The stimuli consisted of two circular patches of drifting square-wave luminance grating (100% contrast, 2–3° diameter, 1–2° drift rate, 1–2 cycles per degree of spatial frequency). One stimulus was positioned inside the recorded neurons’ RF, and the other was at an equal eccentricity in an adjacent visual field quadrant. The task of the monkey was to release the bar between 150 and 650 ms after a change in stimulus color (i.e., a change of the white stripes of the grating to photometrically isoluminant yellow). That change in stimulus color occurred at an unpredictable moment between 500 and 5,000 ms after stimulus onset. All times during this period were equally likely for the color change. Successful trial completion was rewarded with four drops of diluted apple juice. If the monkey released the bar too early or if it moved its gaze out of the fixation window, the trial was immediately aborted and followed by a timeout.

On the 50% of the trials in which the distracter changed before the target, the target nevertheless changed later on in the trial. Blocks consisted of 20 trials. The first two trials in a block were instruction trials in which only one of the two stimuli was shown and the monkey performed the task on that stimulus. The location of that stimulus was the target location for that block. For the remainder of the block, both stimuli were shown together without
any further cue. Thus, in the block design, the different attention conditions were physically identical. We recorded 100–300 correctly performed trials per attention condition.

**Data Analysis.** All data analyses were performed using custom programming in Matlab (The MathWorks). Sustained attentional affects for each cell were calculated by averaging the firing rates across the period from 1,000 to 3,000 ms after stimulus onset for each trial. A test was then used to compare these firing rates across attentional conditions. To calculate the latency of the attentional modulation on firing rate for both the population average and for individual cells, we compared firing rates (averaged over 0.5 ms nonoverlapping bins) for the two attentional conditions using a two-tailed t test. The first of three consecutive significant (P < 0.05) bins defined the onset of consistent attentional modulation.

**ACKNOWLEDGMENTS.** We thank T. Buschman, D. Heister, and P. Su for help with animal training. Supported by National Institutes of Health Grants EY017292 and EY017921 (to R.D.) and MH072034 (to H.L.).