

# Attention Selectively Gates Afferent Signal Transmission to Area V4

Iris Grothe,<sup>1,2\*</sup> David Rotermund,<sup>3\*</sup> Simon David Neitzel,<sup>1</sup> Sunita Mandon,<sup>1</sup> Udo Alexander Ernst,<sup>3</sup> Andreas K. Kreiter,<sup>1</sup> and Klaus Richard Pawelzik<sup>3</sup>

<sup>1</sup>Brain Research Institute, University of Bremen, 28359 Bremen, Germany, <sup>2</sup>Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, 60528 Frankfurt, Germany, and <sup>3</sup>Institute for Theoretical Physics, University of Bremen, 28359 Bremen, Germany

Selective attention allows focusing on only part of the incoming sensory information. Neurons in the extrastriate visual cortex reflect such selective processing when different stimuli are simultaneously present in their large receptive fields. Their spiking response then resembles the response to the attended stimulus when presented in isolation. Unclear is where in the neuronal pathway attention intervenes to achieve such selective signal routing and processing. To investigate this question, we tagged two equivalent visual stimuli by independent broadband luminance noise and used the spectral coherence of these behaviorally irrelevant signals with the field potential of a local neuronal population in male macaque monkeys' area V4 as a measure for their respective causal influences. This new experimental paradigm revealed that signal transmission was considerably weaker for the not-attended stimulus. Furthermore, our results show that attention does not need to modulate responses in the input populations sending signals to V4 to selectively represent a stimulus, nor do they suggest a change of the V4 neurons' output gain depending on their feature similarity with the stimuli. Our results rather imply that selective attention uses a gating mechanism comprising the synaptic "inputs" that transmit signals from upstream areas into the V4 neurons. A minimal model implementing attention-dependent routing by gamma-band synchrony replicated the attentional gating effect and the signals' spectral transfer characteristics. It supports the proposal that selective interareal gamma-band synchrony subserves signal routing and explains our experimental finding that attention selectively gates signals already at the level of afferent synaptic input.

**Key words:** area V4; information routing; monkey; neural network model; selective attention

## Significance Statement

Depending on the behavioral context, the brain needs to channel the flow of information through its networks of massively interconnected neurons. We designed an experiment that allows to causally assess routing of information originating from an attended object. We found that attention "gates" signals at the interplay between afferent fibers and the local neurons. A minimal model demonstrated that coherent gamma-rhythmic activity ( $\sim 60$  Hz) between local neurons and their afferent-providing input neurons can realize the gating. Importantly, the attended signals did not need to be amplified already in an earlier processing stage, nor did they get amplified by a simple output response modulation. The method provides a useful tool to study mechanisms of dynamic network configuration underlying cognitive processes.

## Introduction

Selective attention is required to focus on parts of the continuous stream of incoming information about the environment, thereby

allowing for advanced processing like shape perception of individual stimuli embedded in cluttered visual scenes (Rock and Gutman, 1981; Rock et al., 1992; Wolfe and Bennett, 1997). In a now classical experiment aiming to demonstrate such selective processing on the neuronal level, Moran and Desimone (1985) placed either one, or two stimuli in a neuron's receptive field (RF). They found for neurons in visual areas V4 and IT that

Received Aug. 6, 2017; revised Feb. 14, 2018; accepted Feb. 15, 2018.

Author contributions: A.K.K. and K.R.P. designed research; I.G., D.R., S.D.N., S.M., and U.A.E. performed research; I.G., D.R., and U.A.E. analyzed data; I.G., D.R., U.A.E., A.K.K., and K.R.P. wrote the paper.

This work was supported by the BMBF (Bernstein Group for Computational Neuroscience Bremen, Grant 01GQ0705, Innovationswettbewerb Medizintechnik, Grant 01 EZ 0867, Bernstein Award Udo Ernst, Grant 01GQ1106), the DFG (Priority Program 1665, Grant ER 324/3), the University of Bremen's Research-Focus Neurotechnology, Creative Unit I-See, and Zentrum fuer Kognitionswissenschaften, and the Leibniz Graduate School for Primate Neurobiology (to I.G.). We thank K. Thoß, R. Hakizimana, and K. Taylor for monkey care and training.

The authors declare no competing financial interests.

\*I.G. and D.R. contributed equally to this work.

Correspondence should be addressed to Andreas K. Kreiter, Brain Research Institute, Hochschulring 16a, University of Bremen, 28359 Bremen, Germany. E-mail: kreiter@brain.uni-bremen.de.

DOI:10.1523/JNEUROSCI.2221-17.2018

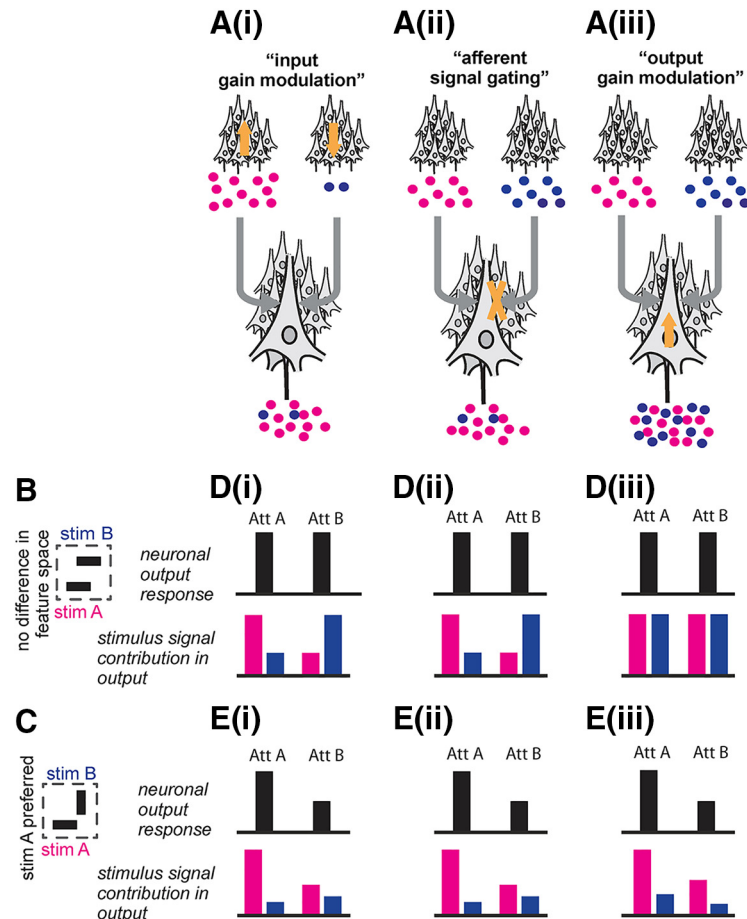
Copyright © 2018 the authors 0270-6474/18/383441-12\$15.00/0

placing a preferred and a nonpreferred stimulus together in the RF resulted in an intermediate firing rate response of the neurons compared with the responses when the stimuli were presented alone. In contrast, when allocating attention to one of the two simultaneously presented stimuli, the neuron's firing rate resembled its response when the stimulus was presented alone. Succeeding studies (Treue and Maunsell, 1996; Luck et al., 1997; Chelazzi et al., 1998; Reynolds et al., 1999) confirmed for various visual areas that the neuronal firing rates can reflect even less preferred stimuli while optimal stimuli are present at the same time. This raises the question where in the neuronal circuitry attention intervenes into signal routing to achieve selective representation of the attended stimulus.

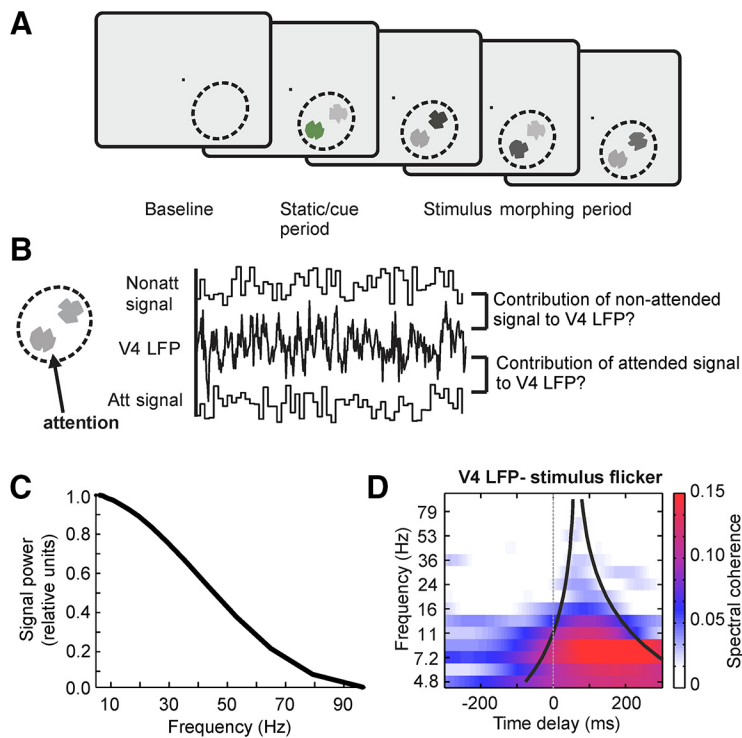
An early intervention point could be the upstream area sending the stimulus signals (Fig. 1*Ai*). If the RFs of the sender area are sufficiently small, the two stimuli will activate two separate populations. Strong differences in the response strengths of these two input populations would prevent signals caused by non-attended stimuli to be conveyed to the receiving neurons in substantial amounts. Indeed, areas like V4 or MT are receiving signals from V1 neurons which have much smaller RFs, however, they often show only small or no attention-dependent modulations of their responses if activated by a single stimulus (Motter, 1993; Luck et al., 1997; McAdams and Maunsell, 1999; Mehta et al., 2000).

Alternatively, two sites further downstream have been proposed for attention-dependent intervention. The first (Fig. 1*Aii*) comprises the synaptic inputs where transmission of signals arriving via the afferent fibers from the upstream area into the V4 neurons occurs. Differential modulation of input signal transmission, for example by modulation of effective connectivity with the input population, would result in local neuronal processing being predominately determined by the attended stimulus (Reynolds et al., 1999; Hamker, 2004; Mishra et al., 2006; Masuda, 2009).

The second downstream site at which selective attention may intervene into stimulus processing is at the output of the V4 neurons (Fig. 1*Aiii*). Modulation of the output gain, for example by changing the neurons' firing thresholds, has been proposed to depend on the correspondence between features (including location) of the attended stimulus and the neurons' stimulus selectivity and RF location (Treue and Martínez Trujillo, 1999; Spratling and Johnson, 2004; Deco and



**Figure 1.** Possible stages at which selective attention intervenes with processing and the predictions resulting from them. *Ai*, Early intervention point. Directing attention could differentially modulate the responses of the separate sender populations (“input gain modulation”) as indicated by the orange arrows. The result would be that much stronger signals originating from the attended object would leave the sender populations and thus result in a strong imbalance between stimulus-related signals processed in the receiver population. *Aii*, Intermediate intervention point. Alternatively, attention could modulate the interactions between receiver and sender populations (“afferent signal gating”) as indicated by the orange cross. The result would be a strong imbalance in favor of processing the attended stimulus' signals in the receiver population, without requiring an imbalance originating from the sender populations. *Aiii*, Late intervention point. Attention could act at an even later stage by modulating the response of the receiver population, in dependence of its tuning properties, at its output stage (“output gain modulation”) as indicated by the orange arrow. This would change the strength, but not the proportion of the signals of both stimuli processed in the receiver population. *B*, Sketch of a RF with two stimuli presented that drive the neuron(s) equally well like in our experimental design. This configuration holds for the predictions outlined in *Di*, *Dii*, and *Diii*. *C*, Sketch of an often used experimental paradigm in which two stimuli are presented in the RF but one is preferred, here stimulus A. This configuration holds for the predictions outlined in *Ei*, *Eii*, and *Eiii*. *D*, According to input gain modulation (*Di*), neuronal responses in the receiver population would be the same (black bars; e.g., single neuron firing rates, for a better comparison normalized to the maximum response in the corresponding prediction scheme), no matter which stimulus is attended, because the stimuli drive the neuron(s) similarly. The same holds true for afferent signal gating (*Dii*) and output gain (*Diii*). The induced neuronal responses are therefore not a particularly suitable measure for differentiating between the three mechanisms. However, if one would be able to measure the stimulus signals' contribution to the signals generated by the receiver population directly (colored bars; normalized to the maximum contribution in the corresponding prediction scheme), one would expect from the output gain modulation scenario (*Diii*) that both stimulus signals would be equally well represented, whereas for the afferent signal gating (*Dii*) mainly the signals from the attended object would be contained. Note, that it is not possible to differentiate strictly between afferent signal gating (*Dii*, *Eii*) and input gain modulation (*Di*, *Ei*) when only measuring the receiver population. *E*, Also in the more common experimental setup, in which the two stimuli drive the neuron(s) differently, the neuronal responses (black bars) will not be able to differentiate between the three scenarios. In contrast to *D*, the receiver population would always show a stronger response when Stimulus A is attended because it prefers Stimulus A. Again, when simultaneously measuring both stimulus signals' contribution to the receiver population output (colored bars) one can differentiate between afferent signal gating (*Eii*) and output gain modulation (*Eiii*). The proportion between the signal contributions of the preferred and not-preferred stimulus would stay the same for the output gain modulation scenario, as it is merely the gain that changes absolute output strength. In contrast for the afferent signal gating scenario, switching attention would strongly change the proportion between the signal contributions because attention would effectively block the non-attended signals while letting attended stimulus signals through.



**Figure 2.** Experimental paradigm and analysis approach. **A**, Outline of the task. After a baseline period, two shapes were presented together in a V4 RF (dashed lines; not visible on the monitor) and started a continuous morphing sequence into other shapes after one of them was cued (initial green coloring) to be attended. The monkeys had to respond to the reappearance of the initially cued shape. Both shapes changed luminance on each visual frame (flicker), independently of each other. Note that the luminance fluctuations were much faster than the morphing speed of the shapes and that they were not relevant for the task. **B**, For each trial we acquired three signals: the flicker signals of the attended and non-attended stimulus and the V4 LFP. We performed spectral coherence analysis to estimate how much each of the flicker signals contributed to the V4 LFP. **C**, Power spectrum of the broad band luminance fluctuations imposed on the stimuli. **D**, The average SC TdF spectrum over all recording sites from Monkey B (single stimulus inside the RFs) shows the typical cone-like shape. For later site-based analyses we averaged the SC within the cone (black lines).

Rolls, 2005; Maunsell and Treue, 2006; Sripathi and Johnson, 2006; Buia and Tiesinga, 2008). Also in this scenario, neuronal firing rate responses in V4 would be reflecting the attended stimulus.

We aimed to identify the site where attention selectively intervenes into routing and processing of the neuronal signals caused by simultaneously presented stimuli. To do so, we introduced a new approach to analyze the causal relation between the stimuli and their corresponding neuronal signals: we imposed independent, broadband luminance modulations on both stimuli (“flicker tags”; Fig. 2B). This temporal tagging allowed us to infer simultaneously and independently to which extent each of the two stimuli within the RF exerted a causal influence on the neuronal activity and thus how much of each stimulus’ signal was routed through local processing of the V4 neurons. Moreover, our approach allowed investigating the characteristics of attention-dependent selective processing for identical stimuli, because the method does not depend on detecting differences in response strength expected for stimuli distinguished by features matching the neurons’ stimulus selectivity to different degrees (Fig. 1).

## Materials and Methods

**Surgical procedures, behavioral task.** Details about the surgical preparation and behavioral task have been reported previously by Grothe et al. (2012). In short, two adult male rhesus monkeys (*Macaca mulatta*) were implanted under aseptic conditions with a post to fix the head and a recording chamber placed over area V4. All procedures and animal care

were in accordance with the regulation for the welfare of experimental animals issued by the federal government of Germany and were approved by the local authorities. Before chamber implantation, the monkeys had been trained on a demanding shape-tracking task. The task (Fig. 2A) required fixation throughout the trial within a fixation window (diameter 1–1.5°) around a fixation point in the middle of the screen. After a baseline period, the monkeys had to covertly attend to the one of two statically presented, closely spaced stimuli (shapes) that was cued (static/cue period). Both shapes started morphing into other shapes and the cued initial shape reappeared at pseudorandomly selected positions in the sequence of shapes. The monkeys were trained to respond with a lever release to the reoccurrence of the initial shape in the cued stream. A reappearance of the initial shape in the distracter stream had to be ignored.

**Stimuli.** The shapes were placed at equal eccentricity. Care was taken to achieve approximately equal responses from the V4 recording site for both stimuli to make sure they were placed at similarly responsive positions of the RF. The stimuli were presented with a refresh rate of 100 Hz on a 22 inch CRT monitor containing  $1152 \times 864$  pixels (Monkey B) or  $1024 \times 768$  pixels (Monkey F), which was placed at a distance of 92 cm (Monkey B) or 87 cm (Monkey F) in front of the animal. Eye position was monitored using a video-based eye-tracking system (Monkey B: custom made, Monkey F: IScan).

To be able to track two input signals (the 2 stimuli) simultaneously and independently in one output signal (the recorded V4 LFP), we used filled shapes and imposed broadband luminance fluctuations (“flicker”) on the stimuli (“tagging”). For this purpose, we changed the luminance of the shapes by choosing a random, integer gray pixel value with each frame update of the display. For Monkey F, the values were drawn from an interval [128, 172], and for Monkey B, from the full range [0, 255], corresponding to luminance fluctuations in a range of 6.9–12.5 and 0.02–38.0  $\text{Cd/m}^2$ , respectively. The background was black (0.00 and 0.02  $\text{Cd/m}^2$ , respectively). To achieve a high signal-to-noise ratio of the flicker tags in the data, we intended to use the full range of luminance levels, however, the range for Monkey F needed to be adjusted to keep his performance similar to his performance without flickering stimuli. Both shape streams had their own independent flicker time series of luminance values. Note that the flickering of the stimuli was not relevant to complete the task. A few trials were included in which only one stimulus was presented for off-line controlling of response strength to individual stimuli.

**Recording.** Local field potentials (LFPs) were measured from V4 with 1–3 epoxy-insulated tungsten microelectrodes (125  $\mu\text{m}$  diameter; 1–3  $\text{M}\Omega$  at 1 kHz; Frederic Haer) in the upper cortical layers. Reference and ground electrodes for Monkey F were platinum-iridium wires below the skull at frontal and lateral sites. The reference for Monkey B was a platinum-iridium wire placed posteriorly below the skull, and the ground was a titanium pin at the posterior end of the skull. For Monkey B, the V4 recording sites were re-referenced to a local reference recorded by a conventional, low impedance electrode placed on top of the dura near the V4 recording site on a daily basis. Each electrode’s signal was amplified 1000 $\times$  (Monkey B) or 5000 $\times$  (Monkey F; PGA 64, Multi-Channel Systems GmbH) and digitized at 25 kHz. For positioning stimuli RFs were mapped manually while the monkey was fixating centrally, followed by an automated mapping procedure consisting of rapid presentations of circular dots.

**Data preparation.** Recorded LFP signals were bandpass filtered from 1 to 150 Hz (FIR filter) and downsampled to 1 kHz. Forward and backward filtering was applied to avoid phase shifts. To get a frequency-resolved representation of the data, we applied a wavelet transformation. In the following, we will consider as signals  $x(t)$  the recorded V4 LFP as well as the input flicker time series. We first removed the mean  $\mu$  and normalized by the signal's SD  $\sigma$ , which were computed individually for each trial in the time window of interest  $T(k)$ .  $T(k)$  comprises all morph cycles in a trial  $k$ , excluding the last cycle leading to the target shape in the attended sequence, which is associated with the correct behavioral response.

$$y(t) = \frac{x(t) - \mu}{\sigma}.$$

The normalized signal  $y(t)$  was then convolved with complex Morlet wavelets  $\omega(t, f_0)$  (Kronlandt-Martinet et al., 1987) to obtain the wavelet coefficients  $a(t, f_0)$  via:

$$a(t, f_0) = \int \omega(\tau, f_0) y(t - \tau) d\tau.$$

The spacing of frequency bands was logarithmic between 5 and 100 Hz, chosen as  $f(l) = \Omega^{l-1} f_0$  for  $l = 1, \dots, 16$  frequency bands starting at  $f_0 = 4.84$  Hz. For a sufficiently tight coverage of frequency space, we set  $\Omega = 1.221$ . Because of the wavelet's overlap in frequency space, data from two neighboring frequency bands are not completely independent.

**RF mapping.** If time allowed, we performed an automated mapping of the RF. Monkeys performed a simple fixation task, in which they needed to fixate a central spot for several seconds until the spot changed its luminance, which the monkey had to detect to receive reward. Filled circles were flashed in the periphery while the monkey was fixating and had to be ignored. In the relevant lower quadrant at 394 positions between 1 and 5 degrees of eccentricity, white filled circles were presented on a black background. Each circle was presented for 300 ms and was followed by a blank screen for 100 ms. A pseudorandom presentation avoided sequential presentations at nearby locations. Data were recorded and prepared as described above. At each location the average evoked response (as defined by broadband power between 5 and 160 Hz in a time window of 50–150 ms after disk onset) was calculated and space between presentation points was interpolated. The RF centers were defined as the coordinates of maximum activation. RF sizes were calculated by measuring the area occupied within the contour of half the maximum response and size is defined as the square root of that area.

**Spectral coherence.** Central to our analysis is the estimation of the contribution of each flicker tag to the V4 LFP. To do so we first calculated a time delay ( $\tau$ )-frequency (Tdf) representation of the absolute value of the complex spectral coherence (SC) between the LFP and the flicker signals. It was computed on the basis of the complex wavelet coefficients  $a_{i,k}(t, f_0)$  of signal  $i$  in trial  $k$ , by:

$$c_{i,j}(f_0, \tau) = \frac{|\sum_{k,t} a_{i,k}(t + \tau, f_0) \overline{a_{j,k}(t, f_0)}|^2}{(\sum_{k,t} |a_{i,k}(t, f_0)|^2)(\sum_{k,t} |a_{j,k}(t, f_0)|^2)}$$

where the sums over time include all values of  $t$  for which *both*,  $t$  and  $t + \tau$ , are contained within the time window of interest  $T(k)$ .  $\bar{a}$  Denotes the complex conjugate of  $a$ , and the values of  $c_{i,j}(f_0, \tau)$  lie in the range of [0, 1]. Note that  $\tau = 0$  represents the point of zero time delay between the external flicker stimulus and its arrival in the V4 LFP.

**Normalized coherence.** If a signal  $y$  is a linear mixture of two components A and B with mixing coefficient  $\alpha$ , given by  $y(t, f_0) = \alpha y_A(t, f_0) + (1 - \alpha) y_B(t, f_0)$ , we would like the (non-lagged) SC between  $y$  and  $y_A$  to be linearly related to  $\alpha$ . Such a relation allows to directly judge “how much” of signal  $y_A$  is contained within the mixture  $y$ , and makes (differences between) SC values easier to interpret. It can be achieved by a nonlinear normalization of the SC  $c_{y,y_A}$ .

We consider the complex wavelet coefficients for a specific frequency band  $f_0$  of the signals  $y_A$  and  $y_B$  in their polar representation  $a_A(t) = X_A(t) \exp(i\varphi_A(t))$  and  $a_B(t) = X_B(t) \exp(i\varphi_B(t))$ . Furthermore, we assume the signals to be statistically independent with zero mean

and equal SD. For infinitely long time series, these conditions imply that  $\Gamma \equiv \langle a_A(t) a_B(t) \rangle_t = 0$  and  $Z \equiv \sum_t X_A(t)^2 = \sum_t X_B(t)^2 = \text{const.}$ , respectively. By inserting these expressions in the spectral coherence  $c_{y,y_A}$ , we obtain:

$$c_{y,y_A} = \frac{|\alpha^2 \sum_t X_A(t)^2 \exp(i\varphi_A(t) - i\varphi_A(t)) + \Gamma|^2}{N}$$

with a normalization factor  $N$  given by the following:

$$N = [\alpha^2 \sum_t X_A(t)^2 \exp\{i\varphi_A(t) - i\varphi_A(t)\}] \times [\alpha^2 \sum_t X_A(t)^2 \exp\{i\varphi_A(t) - i\varphi_A(t)\}] + (1 - \alpha)^2 \sum_t X_B(t)^2 \exp\{i\varphi_B(t) - i\varphi_B(t)\} + \Gamma + \bar{\Gamma},$$

which simplifies to

$$c_{y,y_A} = \frac{\alpha^4 z^2}{\alpha^2 Z (\alpha^2 Z + (1 - \alpha)^2 Z)} = \left( 1 + \left( \frac{1 - \alpha}{\alpha} \right)^2 \right)^{-1}.$$

Solving this equation with respect to  $\alpha \in [0, 1]$  yields the mixing coefficient:

$$\alpha = \left( 1 + \sqrt{\frac{1}{c_{y,y_A}} - 1} \right)^{-1}.$$

As the conditions above are fulfilled for the flicker signals, we have normalized the previously obtained SC according to this equation:

$$C_{i,j}(f_0, \tau) = \left( 1 + \sqrt{\frac{1}{c_{i,j}(f_0, \tau)} - 1} \right)^{-1}.$$

We pooled the Tdf spectra of V4 LFPs with stimulus A and B according to the attentional conditions. Subsequently, for a recording site-based comparison of the SC between the attended and non-attended signals we averaged the SC values in a specific time delay range yielding a mean SC  $\bar{C}(f_0)$ . The range was fixed as a frequency-dependent “cone” around an onset delay of  $\tau = 60$  ms that fits known V4 latency values (Motter, 1993; Lee et al., 2007) and the onset latencies observed for the evoked LFP. The cone was centered on half a wavelength after onset delay, with its borders at  $\pm 7\%$  wavelengths around the center to account for spread of the signal due to wavelet size (Fig. 2). This wavelength-dependent asymmetry of the cone was chosen based on empirical observations of a corresponding frequency-dependent shift of responses in independent data taken from the stimulus onset response.

**Empirical design and statistical analysis.** Included in the study are the data of the two adult male macaque monkeys in which the experiments have been performed. A recorded V4 site was included in all further analyses when individual presentation of the stimuli induced responses of similar strengths. The rationale of the study requires the two stimuli to provide similar input to the local V4 population. Therefore, the two stimuli were very similar and closely spaced. To verify similarity of activation of the local V4 population, we calculated the stimulus induced gamma-band activity as described by Grothe et al., 2012. In short, for trials in which only one stimulus was presented, the power spectral density was computed by taking the square of the absolute value of the result of the wavelet convolution and dividing it by the Nyquist frequency (500 Hz). Visual response strength was calculated as the average gamma-band response during all morph cycles with respect to the baseline (Monkey B: 800 ms, Monkey F: 650 ms before stimulus onset) by first subtracting baseline activity from stimulus driven activity and then dividing by baseline. V4 sites with sustained responses to the two different shape stimuli differing by more than a factor of 2.5 were excluded from the analysis.

In general, for assessing significance of the time delay averaged SC  $\bar{C}$  between the V4 LFP and a flicker signal, we used surrogate data. In particular, we computed a distribution of surrogate SC  $\bar{C}^s$  between the same LFP data, but different flicker signals that were generated by using different initial seeds for the random number generator. In total, we computed 1000 surrogate samples  $\bar{C}^s$ , which were then ranked according to their value. For a (frequency-dependent) spectral coherence  $\bar{C}$  to be

significantly different from zero with an error probability  $<1\%$ , its value had to surpass the 990th largest surrogate sample  $\bar{C}^S$ . For judging whether the difference  $D$  between two spectral coherence values  $\bar{C}_1$  and  $\bar{C}_2$  is significant, we first computed the absolute value of all differences  $D_{1,2}^S$  between all pairs  $\{\bar{C}_1^S, \bar{C}_2^S\}$  of surrogate samples. Again, to be significantly different with error probability of  $1\%$ ,  $D$  had to surpass the 99% quantile of this distribution.

First, as proof of principle, we assessed whether our method worked and we could quantify the “flicker tag transfer” independently of the attentional condition. To do so, we estimated for each frequency bin if more recording sites than expected by chance ( $p < 0.01$ , one-sided binomial test with  $p_0 = 0.01$ ) were showing significant SC values  $\bar{C}(f_0)$  in that frequency bin for at least one attentional condition (Monkey F:  $n = 23$  sites, Monkey B:  $n = 12$  sites; findings reported in Results).

For the main analysis of attention-dependent routing we first tested, for each monkey individually, for a significant attentional gating effect. For this purpose, we computed whether there is a larger number of sites having a higher SC in the attended condition than expected by chance ( $p < 0.01$ , one-sided binomial test with  $p_0 = 0.5$ ). Subsequently, we tested significance of the attentional effect size (see Fig. 4A) for each site individually by determining whether the difference between attended and non-attended SCs is significantly  $>0$  (test statistics  $D_{1,2}^S$  described above).

For the analysis in Figure 4B data of the two monkeys were treated separately to appreciate individual variability (Monkey F:  $n = 23$  sites, Monkey B:  $n = 12$  sites). Its purpose is to highlight the frequency dependence of the effect and therefore the significance value (determined as described assessing significance above) is plotted along with the results.

**Control analysis.** A control analysis was performed in one monkey carrying a V1 array that had been used in previous experiments (Grothe et al., 2012). The V1 array of Monkey F consisted of 96 tungsten electrodes with 2 mm spacing and was placed intracortically. Although the rationale of the study requires the two stimuli to provide similar input to the local V4 population at a given V1 recording site, only one stimulus should induce a response. Therefore, we required a difference of at least a factor of 2.5 in induced gamma power between the stimuli when presented individually. In addition, we required a minimum response strength (above baseline) of one times the baseline activity (Grothe et al., 2012). Applying these selection criteria resulted in  $n = 15$  V1 sites. The experimental V1 data were gathered in the same recording sessions as the V4 data and accordingly further experimental procedures, data acquisition, data analysis and statistical procedures were as described in the corresponding methods sections above, i.e., they were exactly the same as for the main experiment in V4.

**Model.** The model consists of three modules (see Fig. 5), each of which is a threshold linear function of the respective inputs. Two modules represent neural “sender” or input populations, such as in V1, driven by Stimulus A and Stimulus B, respectively. The third module represents a neural “receiver” population, such as in V4, which receives common input from both sender populations. We assumed each module is subjected to gamma oscillations. For simulating (in-phase and anti-phase) gamma synchrony between populations, the gamma oscillations were derived from a common “master clock”, which itself generates a jittered sinusoidal signal. In particular, the cycles  $k$  of the master clock were generated by randomly drawing a sequence of intervals from a normal distribution with mean 16 ms and SD 2.5 ms, which is equivalent to a 62.5 Hz oscillation, similar to the induced gamma band observed in our monkeys with this paradigm. The core assumptions of the model have found clear experimental support. Attention has been shown to modulate local gamma-band synchronization (Sokolov et al., 1999; Fries et al., 2001; Bichot et al., 2005; Taylor et al., 2005; Jensen et al., 2007; Womelsdorf et al., 2007; Rotermund et al., 2009, 2013) and recent experiments have shown that a receiver neuronal population can selectively synchronize in the gamma band with one of multiple input populations (Bosman et al., 2012; Grothe et al., 2012).

The sequence generated by the master clock defines times  $t_k$  where the master clock is in phase  $\varphi = 0$ . Gamma oscillations  $\gamma(t) = \sin(\varphi(t))$  were derived by first adding normally distributed time jitter  $\Delta t_k$  with 0 mean and SD  $\sigma_\gamma$  to the times  $t_k$ , and by then performing a cubic interpolation

for defining a time-dependent phase  $\varphi(t)$  for all other times  $t$  from the supporting points  $\varphi(t_k + \Delta t_k) = 2\pi k + \varphi_0$ .  $\varphi_0$  Realizes a constant average phase shift between the master clock and any derived gamma oscillation.

For the simulations, the gamma oscillations in the V1 population driven by Stimulus A and in the V4 population were always derived from the master clock with phase shift zero (after taking into account the conduction delays from V1 to V4). According to our working hypothesis this phase relation reflects attention being directed to Stimulus A. For the gamma oscillations in the second V1 population driven by the non-attended Stimulus B, two procedures were used: first, derivation from the master clock with phase shift  $\varphi_0 = \pi$ , and second, derivation from an independent realization of the master clock. These two procedures are equivalent with assuming that the V1 population representing the non-attended stimulus either goes into anti-phase synchronization with V4, or goes into a random phase-relationship with V4. For modeling situations in which the V1 populations only intermittently goes into anti-phase with V4, we introduced a control parameter  $\mu$ . In the corresponding simulations, a fraction  $\mu$  of all trials then used the independent gamma oscillation for the V1 population driven by the non-attended flicker signal, and the remaining fraction  $1 - \mu$  of all trials used the anti-phase gamma oscillation. As the spectral coherence is obtained by summing over trials and time, this procedure yielded the same results as if there would be an intermittent change between anti- and random-phase periods during a trial, which is more difficult to generate numerically in a controlled manner.

Activity in each population was described by  $v$ , which was modeled as a linear superposition of the feedforward input  $I$ , weighted by a factor  $\lambda$ , and the local gamma oscillation  $\gamma$ , passed through a threshold-linear gain function  $g$  and additive noise  $\eta$  with amplitude  $c$ .  $\eta(t)$  Was realized by drawing a random number from a uniform distribution in the interval  $[-1, 1]$  for every time  $t$  (for a time discretization of  $\delta t = 1$  ms):

$$v(t) = g_{a,b}[\lambda I(t) + \gamma(t)] + c\eta(t),$$

$$\text{with } g_{a,b}[z] = \begin{cases} a(z - b), & z \geq b \\ 0, & z < b \end{cases}.$$

Because we measured the LFP experimentally, we assumed that recordings will acquire a smoothed and noisy version  $x$  of the activity  $v$ . We modeled  $x$  as a convolution of  $v$  with an exponentially decaying kernel  $L = 0.03 \exp(-t/\tau_0)$  with time constant  $\tau_0 = 30$ ms, plus “external” noise:

$$x(t) = (v \otimes L)(t) + d\eta(t).$$

Internal and external noise sources were realized by independent random processes for all three populations. The feedforward input was defined as:

$$I_{V1,A}(t) = F_A(t - \tau_{V1}); I_{V1,B}(t) = F_B(t - \tau_{V1});$$

$$I_{V4}(t) = \frac{1}{2}\{v_{V1,A}(t - \tau_{V4}) + v_{V1,B}(t - \tau_{V4})\},$$

where  $F_A$  and  $F_B$  are the flicker signals, and  $\tau_{V1} = 50$  ms and  $\tau_{V4} = 10$  ms the signal conduction delays from the retina to V1 and from V1 to V4, respectively. Further model parameters were set to the following values:

$$\text{for V1: } a = 6, b = 0.2, c = 1, d = 1.75, \lambda = 0.2$$

$$\text{for V4: } a = 2.5, b = 0.8, c = 0, d = 0.8, \lambda = 0.35.$$

The parameters  $a$  and  $b$  were adjusted such that the corresponding layers effectively clip the “troughs” of the total synaptic input, and the noise levels chosen such that the LFP signals resemble real physiological signals. Because we do not pass output from layer V4 on to subsequent areas, we subsume the total impact of internal and observation noise in parameter  $d$  (i.e.,  $c = 0$ ). After choosing an initial set of parameters, their values were fine-tuned (in particular the noise levels and mixing constants  $\lambda$ ) to match SCs in the model to the physiological data.

We performed 100 trials with a fixed trial length of  $T = 6300$  ms, and normalized spectral coherences  $C$  computed according to the procedure outlined in the data analysis section. This amount of data typically is available from one recording site, which implicitly determines significance levels of the SC. To obtain smoother curves for the model graphs, the corresponding results were averaged  $>10,000$  simulations with different seeds for the random number generator.

The gating strength  $G$  was defined as the ratio of the SC (averaged in a symmetric cone around a delay of 60 ms for a frequency range up to 11 Hz) of the LFP with the attended signal and the LFP with the non-attended signal. The synchronization ratio  $S$  was obtained from the corresponding ratio of average SCs between the simulated V4 and input population activity in a frequency range between 40 Hz and 80 Hz, thus quantifying the increase in gamma synchronization between V1 and V4 under attention.

## Results

Two monkeys performed a demanding shape-tracking paradigm (Taylor et al., 2005; Grothe et al., 2012; Fig. 2A) in which two similar visual stimuli, one attended and one non-attended were placed close together, to ensure a strong need for attentional stimulus selection (eccentricity of the RFs: mean  $2.64^\circ$ , SD  $0.35^\circ$ , range  $2.24$ – $3.52^\circ$ ; sizes of the RFs: mean  $1.93^\circ$ , SD  $0.15^\circ$ , range  $1.73$ – $2.26^\circ$ ,  $n = 12$ ). On top, we independently tagged both stimuli via broadband luminance modulations. On each display frame, the luminance of each shape was changed to a random value, resulting in a rapid “flickering” of the stimuli (Fig. 2A,B). These fast luminance fluctuations were not necessary to perform the task and were not informative in any way for the monkeys. Further, the values were changed independently for both stimuli, which made the time-varying signals unique for each stimulus (Fig. 2B). To characterize the transmission of the tags of each stimulus to the LFPs recorded in V4, we computed the normalized absolute value of the complex SC between the luminance modulation of the stimuli and V4 LFP (see Materials and Methods). This allowed inferring simultaneously and independently to which extent both stimuli were causally influencing neuronal activity recorded in V4. After having learned the shape-tracking task, monkeys quickly adapted to the luminance modulations and performed well on the task, despite the flickering of the stimuli. Excluding trials with a fixation error, Monkey F performed with 79.1% correct responses, 9.9% false alarms, and 11.0% misses, whereas Monkey B performed with 90.9% correct responses, 7.9% false alarms, and 1.2% misses averaged over all sessions included in the analyses (Monkey F: 26.9% fixation errors, Monkey B: 30.7% fixation errors). V4 LFP data were acquired in 16 sessions with one up to three electrodes per session that provided 35 recording sites (Monkey F: 23 sites, Monkey B: 12 sites) fulfilling the inclusion criteria (see Materials and Methods) which were used in the subsequent analyses.

### Attention gates input signals to V4 LFP

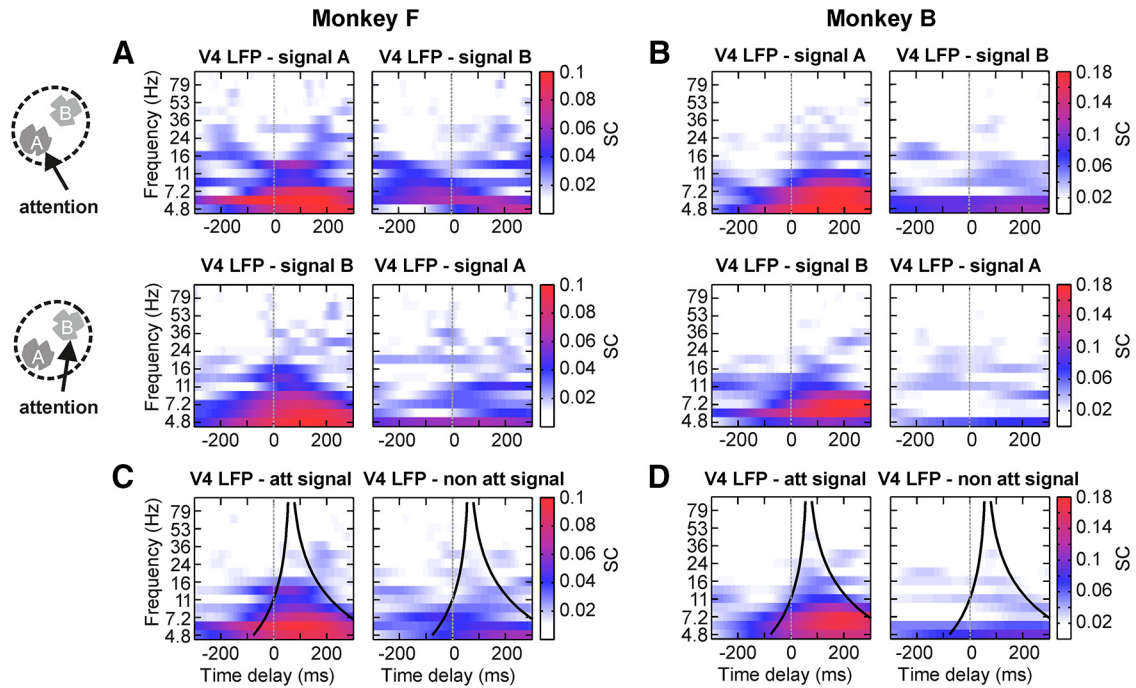
The external signals, i.e., the random luminance modulations of the stimuli, contained a broad range of frequencies with their main spectral power  $<50$  Hz (Fig. 2C). TdF representations of the SC show that specific components of these external signals causally influenced the LFP in V4, as illustrated by the average TdF spectrum of multiple recording sites (Fig. 2D; single shape presentations only, all sites of Monkey B included). It shows that the influence shaped the TdF spectrum in a cone-like fashion: first, it is wider for lower frequencies reflecting the temporally extended correlation structure of lower frequency signals. Second, it is shifted to the right because of the time delay of the neuronal signals. Due to the asymmetry of the cone, the delay at

which SC reached its maximum depended on the frequency, ranging from  $\sim 75$  ms for 24 Hz up to  $\sim 200$  ms for 5 Hz. The SC between flicker signal and V4 LFP dropped with increasing frequency indicating that most of the signal components processed in V4 have frequencies  $<20$ – $30$  Hz (Fig. 2D). Note that the frequency-dependent declines in SC between flicker signals and V4 LFP are not attributable to the cutoff of the input power spectrum which reaches  $-3$  dB not before 45 Hz (Fig. 2C). For determining the upper frequency limit for significant signal transfer, we estimated for each frequency bin if more recording sites than expected by chance were showing significant SC values  $\bar{C}(f_0)$  in that frequency bin (all attentional conditions included). For both monkeys, all frequency bins up till including the bin centered on 24 Hz showed significantly more sites than expected by chance ( $p < 0.01$ ) with a significant SC, thereby setting the frequency border up to which we can observe signal transfer at 24 Hz.

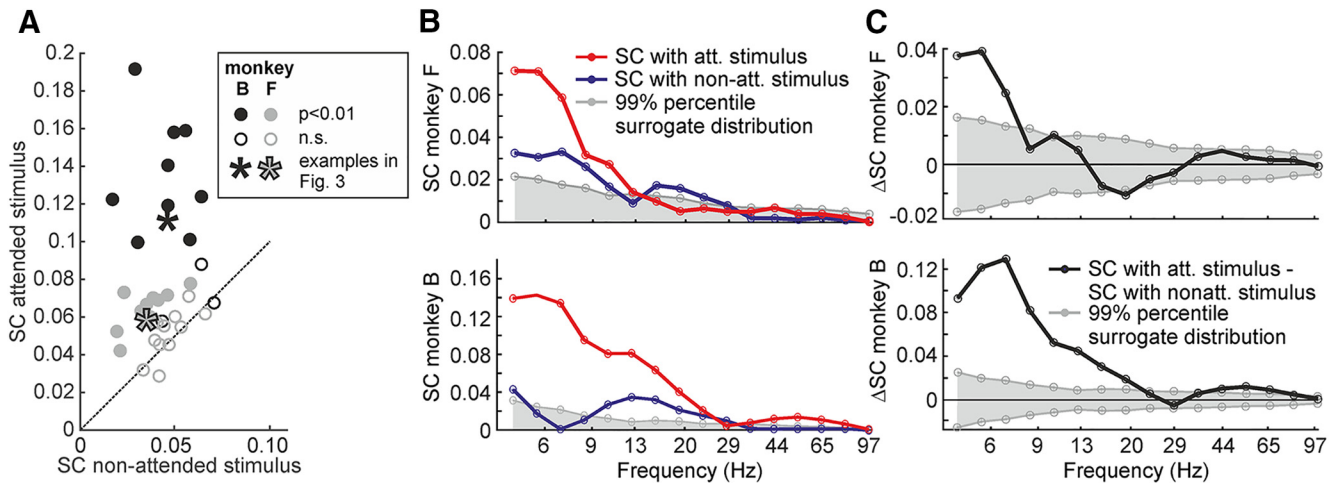
After establishing that we could read out the external signals in V4, we used the SC measure to investigate the attention-dependent signal routing. We placed stimuli such that both were present in the RF of the local V4 population. Therefore, the neurons received synaptic input corresponding to both stimuli. Within our paradigm gating of inputs and output gain modulation predict two different results. Input gating prevents (most of the) transmission of signals caused by the non-attended stimulus into local processing within V4. Therefore, a clear dominance of the flicker tag of the attended stimulus is expected, also if the stimuli are not distinguished by features for which the neurons are selective (Fig. 1Aii,Dii,Eii). In contrast, output gain modulation predicts that afferent input signals of both stimuli enter into neuronal processing independently of the direction of attention. If the two stimuli would differ in features for which the neurons are selective, the flicker tag of the preferred stimulus would dominate, however, the proportion of the two flicker tags should not change if attention is directed from one stimulus to the other. If the two stimuli do not differ with respect to features for which the neurons are selective, flicker tags are expected to be present to a similar amount, with no influence of switching attention (Fig. 1Aiii,Diii,Eiii).

In our experiment, the presented stimuli did indeed not differ systematically with respect to their features and appeared at equivalent positions within the RF. As a result, both tagging signals were expected to contribute similarly to the V4 LFP if attention would not selectively gate the signals at the input level.

However, our results show that when attention was allocated to Stimulus A, Stimulus A predominantly contributed signals to the V4 LFP, whereas simultaneously very little transfer of Stimulus B was observed. When attention was allocated to Stimulus B, accordingly Stimulus B was the major signal contributor to the V4 LFP. This is illustrated by example recording sites for both monkeys (Fig. 3A,B). For all recording sites, we quantified this effect by pooling the SC TdF spectra of the V4 LFP with Stimuli A and B when they were attended and compared it with the pooled SC TdF spectra when the stimuli were not attended (pooling examples for one site shown in Fig. 3C,D): we calculated the average SC up to the border frequency of 24 Hz, and within the cone around a time delay of 60 ms (Fig. 3C,D; Materials and Methods). For almost all sites in both monkeys (11 of 12 for Monkey B; 20 of 23 for Monkey F), we obtained a higher SC of the LFP with the attended stimulus, thus demonstrating a significant gating of the attended signal compared with the non-attended signal (one-sided binomial test:  $p < 0.00322$  for Monkey B;  $p < 0.00024$  for Monkey F). For quantifying the magnitude of this



**Figure 3.** Example recording sites of attention-dependent signal gating. **A**, Tf representations of spectral coherence for one exemplary recording site of Monkey F. When attention was allocated to Stimulus A, the SC between the V4 LFP and the flicker signal of Stimulus A was high, and low between the V4 LFP and the flicker signal of Stimulus B (top row). This effect reversed when Stimulus B was attended (bottom row). **C**, Same recording site as in **A**, but data for the two attended signals (left column) and the two non-attended signals (right column) were pooled for further analysis. **B** and **D** show the same result for one exemplary recording site of Monkey B.



**Figure 4.** Attention-dependent signal gating across recording sites and its frequency dependence. **A**, Attention-dependent modulation of SC. Each circle or star represents a single recording site. Filled circles show recordings sites with a significant ( $p < 0.01$ ) attentional effect calculated over the frequency range from 5 Hz up to 11 Hz, whereas open circles show no significant effect. Stars correspond to the examples depicted in Figure 3. Black and gray symbols depict recordings sites of Monkeys B and F, respectively. **B**, These plots show the frequency-dependent profiles of SC after pooling V4 sites with a significant gating effect (**A**, filled circles and stars) for both monkeys separately. The greatest differences between the SC with the attended stimulus (red lines) and the non-attended stimulus (blue lines) occur for the lowest frequencies. SC values above the black lines are significantly different from chance. **C**, Difference between the attentional conditions in **B**, directly plotted.

effect, we next compared the SCs averaged up to frequencies of 11 Hz where signal transfer is strongest (Fig. 2D), and within the cone around a time delay of 60 ms. Across recording sites and monkeys, we observed strong signal gating by attention (Fig. 4A), similar to the examples shown in Figure 3. From the 35 V4 recording sites, 20 (9 of 12 for Monkey B; 11 of 23 for Monkey F) showed a significant difference between the SC of the LFP with the attended signal and the SC of the LFP with the non-attended signal. Without exception, recording sites with a significant attention-dependent difference in signal contribution to the LFP

show that it is the attended signal that is contributing stronger to the V4 LFP. The mean SC of the V4 LFP with the attended flicker tag was more than a factor of 2.4 higher than that of the SC with the non-attended signal (mean SC with attended signal: 0.096, SD: 0.042; mean SC with non-attended signal: 0.039, SD: 0.014). Differences in eye position or jitter were very small between the attentional conditions and could not explain the observed gating effects: we repeated the SC analysis on a subset of our data by selecting trials from both attentional conditions until a Kolmogorov–Smirnov test was unable to detect a difference in the corre-

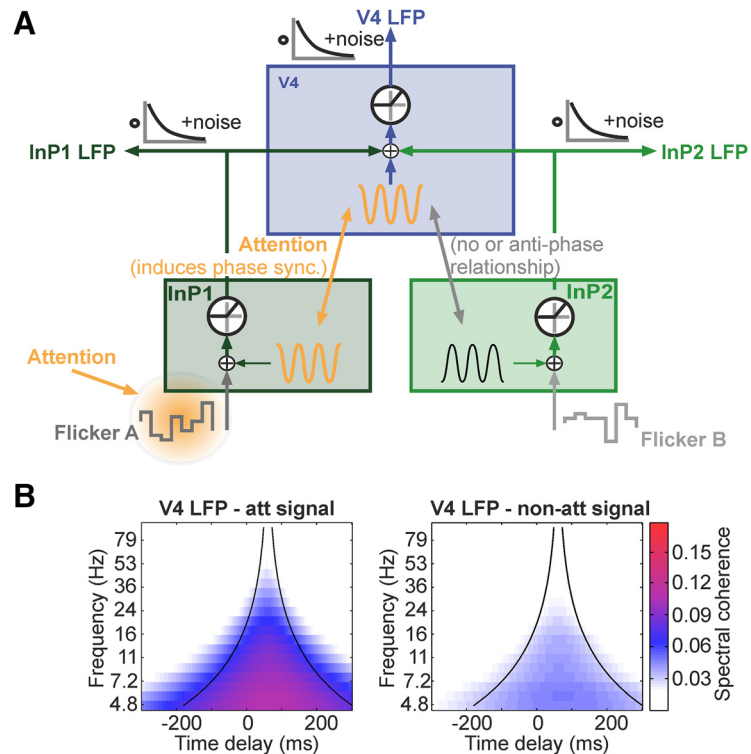
sponding distributions of eye positions and eye jitter ( $p > 0.05$ ). We obtained SCs almost identical to those displayed in Figure 4A, demonstrating that eye movements had a negligible impact on the main gating effect.

For assessing the frequency dependence of input signal gating we again computed the SC within the cone around a time delay of 60 ms for each attentional condition separately, after pooling all recording sites with a significant effect of attention for each monkey. This revealed that for a broad range of frequencies, the signal transfer was weaker for the non-attended than the attended stimulus (Fig. 4B,C), with the difference being largest for lower frequencies, where also signal transfer is strongest. No correlation of the total power differences with the differences of spectral coherences for the two attention conditions for both monkeys was observed (data not shown). A control analysis in Monkey F showed that the SC between the flicker tags and the LFPs of the two upstream V1 populations providing the signals to V4 was slightly but significantly smaller when attention was directed to the stimulus within the V1 RF. In total, 7 of 15 sites showed a significant difference by on average 22.5%, exhibiting a mean SC of 0.089 (SD: 0.060) if the stimulus was attended, and a mean SC of 0.106 (SD: 0.062) if the stimulus was not attended.

### A model explaining routing of signals

With essentially identical task requirements, but without luminance modulations, we recently demonstrated that selective attention goes hand in hand with highly selective gamma-band synchronization between V4 and V1 (Grothe et al., 2012). In short, we previously showed that receiver neurons in V4 synchronize with only one of multiple input populations in V1: the population representing the attended stimulus. This suggests that selective inter-areal synchronization in the gamma-band could underlie selective stimulus processing as observed in the present experiments. Based on these results, we built a simple model to test whether the idea of signal routing by inter-areal gamma-band synchrony can explain the current experimental findings (see Materials and Methods). Figure 5A presents the structure of this model, consisting of two input populations receiving two separate flicker signals and projecting to one common population representing a local group of neurons in V4. For accounting for the observed mean reduction in SC of the V1 population representing the attended signal with the V4 LFP, we reduced the corresponding flicker input strength  $\lambda$  by 15%. Each population was subjected to gamma oscillations and we simulated gamma synchrony between populations to be in or out of phase. For more realistic models where oscillations and synchrony between populations emerge from neuronal interactions see (Harnack et al., 2015; Palmigiano et al., 2017).

Interestingly, the typical shape of the TdF spectra as well as the characteristic decay in SC with increasing frequency closely matches the experimental data (compare Figs. 5B with 3, 4B).

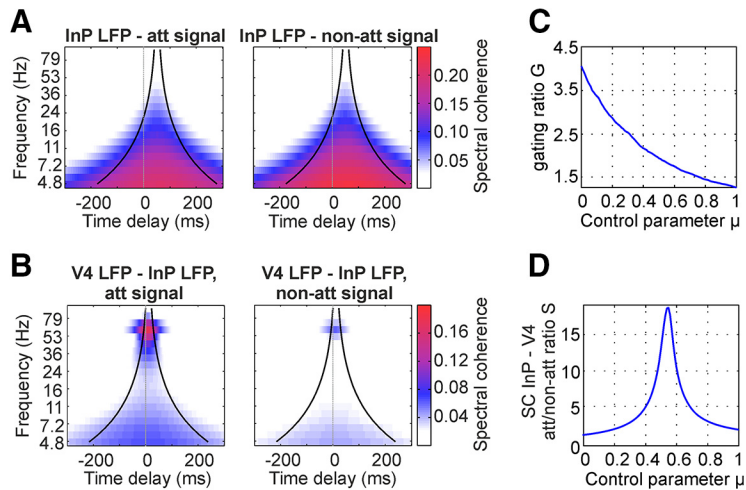


**Figure 5.** A model of signal routing by gamma-band synchrony reproduces gating characteristics as observed in the experiments. **A**, The model consists of three neuronal populations with threshold-linear gain functions (circles). One input population (InP) such as V1 receives the flicker tag A as feedforward input, and a second input population is driven by the flicker tag B. The third population represents V4, which receives a superimposed signal from both input populations. Each population contains a noisy gamma oscillator whose signal is added to the feedforward input. LFP measurements are modeled by convolving the neural signal with an exponentially decaying kernel. Both internal dynamics and external observation are subject to additional, independent noise. Attention was modeled by synchronizing the V4 gamma oscillator with the gamma oscillator of the population receiving input from the attended stimulus. The gamma oscillator of the other input population, processing the non-attended stimulus, is set into anti-phase (or to a random phase) with respect to the V4 population. **B**, The model quantitatively reproduces the SC in the attended (left) and non-attended (right) conditions (compare to experimental data in Figs. 3, 4).

Moreover, we found that the SCs between the flicker signals and the simulated V4 LFP were fully consistent with the SCs calculated for the experimentally recorded V4 LFP. In the model, the “attended” signal, for which the sender population was in phase with the receiver population, was reflected much stronger in the simulated V4 LFP. While the model results suggest that signal routing by inter-areal gamma-band synchrony can explain the observed attentional gating findings, it also predicts how the SCs between the input population’s LFPs and the flicker time series as well as between input population’s LFPs and the V4 LFPs would look like (Fig. 6A,B).

Further, it allowed us to test the adequacy of two different scenarios for reducing the influence from the non-attended stimulus on V4 activity (Fries, 2009). The first scenario assumes a strict anti-phase relationship between gamma activity in V1 and V4, while the second scenario assumes a random phase relationship. In the model, we quantified the consequences of both possibilities on signal gating strength and the ratio of the synchrony between V4 and its two V1 input populations. In particular, we introduced a control parameter  $\mu$  which allowed a continuous transition between the extremes of having either perfect anti-phase ( $\mu = 0$ ) or perfect random phase relationship ( $\mu = 1$ ). For example,  $\mu = 1/3$  would represent a situation in which in 66.6% of the duration of a trial, V1 is in antiphase with V4, while taking a random phase relationship in the remaining time. To analyze the relation between selective synchronization and selective trans-





**Figure 6.** Predictions of the model. **A**, The model predicts that the SCs between the flicker tag and the sender populations do not depend on the attention condition. The cone is centered around an assumed delay between flicker signal and input populations of 50 ms. **B**, The model predicts that the SC between sender populations [input Populations (InP)] and simulated receiver V4 activity is strong in the gamma range and modulated by attention. The cone is centered around an assumed delay of 10 ms between sender populations and V4. **C**, Gating strength expressed in terms of the gating ratio  $G$  in dependence on the parameter  $\mu$ .  $\mu$  Controls the randomness in the phase relationship between V4 and the input population representing the non-attended flicker signal. The model predicts that  $G$  monotonously decreases with  $\mu$ . **D**, Predictions of the model for the attentional modulation of inter-areal gamma-band synchrony between sender populations and receiver population (synchronization ratio  $S$ ) in dependence on the parameter  $\mu$ .

mission of the simulated stimuli, we investigated the gating strength  $G$  (the ratio between the SCs of the simulated V4 LFP and the attended and non-attended signals) and the synchronization ratio  $S$  (the ratio between the synchrony of the V4 LFP with the input populations representing the attended and non-attended stimuli) in dependence of  $\mu$  (see Materials and Methods). As expected, a perfect anti-phase relation between an input population and V4 caused the strongest attenuation of the non-attended signal, leading to the highest value of  $G$  for  $\mu = 0$ . Also, the synchronization ratio  $S$  was lower for perfect anti-phase synchronization (at  $\mu = 0$ ) than for a random phase relation between V1 and V4 gamma activity. Surprisingly, the synchronization ratio  $S(\mu)$  was found to be not monotonous, but rather to exhibit a maximum at intermediate values of  $\mu$ . This feature is caused by two counteracting factors contributing to the denominator of  $S$ , which is given by the SC between the V4 LFP and the input population representing the non-attended stimulus: in particular, one factor is the contribution of episodes for which the signals are in an in-phase relationship, and the second factor is the contribution of episodes for which the signals are in an anti-phase relationship. When averaged in complex space, these two factors contribute to the SC with opposite signs, and their relative contribution depends on  $\mu$ . In addition, during in-phase episodes the simulated flicker signals of the non-attended stimuli add effectively on top of the V4 activity, hence causing the in-phase contribution to be stronger than the contribution of anti-phase coherence. In consequence, the denominator undergoes a transition from a state where anti-phase contributions dominate ( $\mu = 0$ ) to a state where in-phase contributions dominate ( $\mu = 1$ ), thus revealing a minimum SC for values of  $\mu$  in between those extremes.

For our experimentally observed mean of  $\sim G = 2.4$  (Fig. 4A), the model predicts that the input population representing the non-attended stimulus will assume a gamma phase-relationship with V4 activity that is a mixture of being in perfect anti-phase and in a random phase ( $\mu = 1/3$ ). For this regime, we would predict a synchronization ratio  $S \sim 3.3$ , which is well in line with

the experimental observation of gamma phase synchronization between V1 and V4 being approximately four times stronger in the attended condition (Grothe et al., 2012).

Complementing the experimental results, our model also allowed to address the question whether signal gating is an enhancement of the attended signal, a suppression of the non-attended signal, or a combination of these effects. The answer depends on how well the input populations are synchronized with V4 in a “neutral” condition, i.e., in a situation where there is no attention on any of the two stimuli: if there will be a purely random phase relationship, then suppression of the non-attended flicker signal and enhancement of the attended signal will both increase with decreasing  $\mu$ . If the neutral condition is identical to the non-attended condition, there will be only an enhancement of the attended signal but no change in transmission of the non-attended signal. For the particular value of  $\mu$  suggested by our experimental results ( $\mu = 1/3$ ), distractor suppression might

range from 0 to  $-35\%$ , whereas target enhancement might range from  $+60\%$  to  $+140\%$ , respectively.

## Discussion

In this study, we investigated at what processing level in V4 attention intervenes into the signal routing of two stimuli, both positioned within the same RF (Fig. 1). We showed that it is possible to simultaneously quantify the contribution of signals from both, highly similar, stimuli to neuronal processing in V4, by tagging the two stimuli with statistically independent broadband luminance modulations and estimating their spectral coherence with the V4 LFP. We found that the V4 LFP was predominantly coherent with the attended stimulus’ luminance fluctuations, which suggests that attention gates signals at the interface between afferent fibers and the local neurons.

Using the spectral coherence measure also allows determining the response function of the LFP and hence to explore signal routing and its underlying mechanisms in more detail. The V4 LFP could follow mainly flicker components  $<24$  Hz and showed strongest coherence and attentional signal gating with the lowest signal frequencies. A minimal model which implemented signal routing by selective gamma-band synchronization between neuronal populations (Fries, 2005; Kreiter, 2006) replicated both the observed characteristics of signal transfer to V4 and the attentional signal gating effect.

A substantial number of studies have shown that allocating attention to a stimulus accompanied by one or more distracters modifies the strength of neuronal responses to the stimuli (Moran and Desimone, 1985; Treue and Maunsell, 1996; Luck et al., 1997; Chelazzi et al., 1998; Kastner et al., 1998; Reynolds et al., 1999; Lee and Maunsell, 2010; Ni et al., 2012; Keitel et al., 2013). However, different mechanisms have been proposed for this attention-dependent intervention in signal processing.

One possibility, often referred to as “input gain” modulation (Fig. 1A*i*), requires opposite modulation of the output signals from the presynaptic sites (Ghose and Maunsell, 2008; Ghose,

2009). Quantitative modeling applying non-linearities found to explain center-surround interactions in V1, could reproduce experimentally observed attentional modulations of firing rates in V4. This required very strong activity gain modulation of the simulated sender neurons. Previous experimental work showed in contrast, that attention-dependent modulations of V1 and V2 neurons responding to individual stimuli within their RFs are typically very weak or absent in monkeys (Motter, 1993; Luck et al., 1997; McAdams and Maunsell, 1999; Mehta et al., 2000) and humans (Yoshor et al., 2007). Stronger effects have been observed for stimuli which imply strong contextual effects (Ito and Gilbert, 1999) including contour integration tasks (Roelfsema et al., 1998). However, none of these results indicate for the sender areas the strong rate modulations of 50% and higher used to model attention-dependent changes in V4 (Ghose, 2009). For the present task, previous work suggests no such modulation (Grothe et al., 2012) and the control experiment shows rather a small decrease of SC in V1 if the stimulus is attended and no indication for signal enhancement. Thus, the particularly strong attention-dependent selectivity of signal processing observed here in V4 cannot be explained by differential activity modulation of the presynaptic sites. A possible reason for the small attention-dependent decrease in SC in half of the V1 sites might not be a change in the flicker stimulus related LFP component, but a small attention-dependent increase of the overall power of other, internally generated signal components. Because the normalized SC measures the relative proportion of the stimulus related components to the entire signal (see Materials and Methods) this would result in the observed small but consistent decrease of the measure.

These considerations suggests that attention-dependent selective stimulus processing depended predominantly on mechanisms acting in V4. Some models of selective attention explain response modulations by changing the neuron's output gain (Fig. 1Aiii). An attention-dependent mechanism enhancing the output gain of neurons which prefer the features that characterize the attended stimulus and suppress neurons with different feature preferences would cause neurons to respond strongly if the preferred stimulus is attended and only weakly if the nonpreferred stimulus is in the RF (Treue and Martínez Trujillo, 1999; Maunsell and Treue, 2006). In the present study, the two stimuli resided within the same V4 neurons' RF, provided these neurons with similar amounts of afferent signals, and likely did not drive distinct subsets of feature specific neurons, because very similar shapes were presented as a continuously morphing stream and taken out of the same set. Aside from different feature selectivity, spatially separated RFs matching the position of the two stimuli may define subsets of V4 neurons within the population that are separately accessible for attention-dependent output gain modulation. The assumption that the recorded LFP does not contain contributions from such different subsets is substantiated by the finding that the RF sizes that we measured were highly similar to RF sizes found for single units at the corresponding eccentricities in earlier studies (Desimone and Schein, 1987; Motter, 2009). This implies that the populations of neurons contributing to the LFP did not sample much more visual space than the single neurons that they were composed of. In other words, we expect both stimuli would fit into the RFs of the corresponding single neurons as well. Thus, selection of anatomically or functionally defined subsets of V4 neurons for attention-dependent modulation is implausible in our experimental design.

The output gain would be predicted to remain unchanged if attention switches between two equivalent stimuli within a RF.

Their temporal broadband modulations should therefore equally influence the temporal pattern of the spiking output of the V4 neurons. The results show, on the contrary, a predominance of signals caused by the attended stimulus, suggesting that V4 neurons' output gain modulation does not provide a straightforward explanation for the observed attention-dependent processing in a majority of the sites.

We suspect that in the sites not showing a significant attentional change in signal strength, the effect may be masked due to the high noise level and limited amount of data we had available. This is substantiated by the observation that at least part of those sites seem to show no reliable transfer of the flicker tag over all frequencies from 5 up to 24 Hz: their spectral coherence values only reach significance in less than three consecutive frequency bins.

An alternative mechanism is input gating by differential modulation of synaptic transmission of input signals arriving at the V4 neurons, i.e., "synaptic input gain". This mechanism biases the effectiveness by which specific subsets of afferent synaptic input signals are transmitted to V4 neurons in favor of those signals caused by the attended stimulus and does not depend on modulations of the activity in the upstream populations of neurons delivering the afferent signals. Such selective input gating has been proposed by several models implementing biased competition (Reynolds et al., 1999; Hamker, 2004; Mishra et al., 2006) and may implement multiplication of input signals assumed by divisive normalization models of attention (Lee and Maunsell, 2009; Reynolds and Heeger, 2009; Ni et al., 2012). The receiver neurons in such input gating models are much stronger influenced by the signals of the attended stimulus. Correspondingly, neurons in this study would be expected to preferentially express the temporal tagging patterns imposed on attended stimuli, despite of the absence of feature differences between target and distracter stimuli, which is well in line with the observations in this study. Our results therefore indicate that attention can already gate afferent neuronal input signals very effectively.

This raises the question how such differential effective gain modulations for multiple inputs are achieved. Recent observations of selective modulations of gamma-band synchrony between V4 neurons and subgroups of their V1 input (Bosman et al., 2012; Grothe et al., 2012) suggest that signal routing by synchrony may serve as a corresponding mechanism. The minimal model based on this concept does not only replicate the observed attentional gating effects, but also explains the spectral transfer characteristics of the signal. Attentional allocation was solely implemented by selective gamma-band synchrony and did not require modulation in the activity of the upstream population nor a specific stimulus selectivity of the V4 population for one of the stimuli. Hence, also the model supports the hypothesis that attention can differentially gate signals already at the level of afferent synaptic input from upstream cortical areas and does not need to rely on feature differences between the stimuli, nor on output gain changes. Moreover, for the observed averaged gating ratio of  $\sim G = 2.4$ , the model delivers the prediction that V1 neurons representing the non-attended stimulus will assume a gamma phase-relationship with V4 activity that is a mixture of being in anti-phase and in a random phase. The synchronization ratio corresponding to this regime is well in line with the experimental observation of gamma phase synchronization between V1 and V4 being approximately four times stronger in the attended condition (Grothe et al., 2012). The latter and our current findings are in line with previous proposals (Fries, 2005; Kreiter, 2006) and previous theoretical work (Masuda, 2009; Tiesinga and

Sejnowski, 2010; Battaglia et al., 2012) proposing that such switches of gamma-band synchrony may route the information selectively throughout the neuronal processing system.

The tagging technique used in the present work for tracking multiple visual stimuli is related to frequency tagging used in electroencephalographic studies (Morgan et al., 1996; Müller et al., 2003) and often for BCI purposes (starting with Middendorff et al., 2000). Usually, two or more stimuli exhibit regular flicker with different constant frequencies (but see VanRullen and McDonald, 2012). Despite the similar principle, tagging with a broad spectrum of frequencies provides an elegant way of overcoming difficulties of tagging with a periodic signal. In particular, we avoid the problem that single frequencies should be different enough to not interfere with each other or have similar harmonics, and that stimuli might be not equally difficult to attend or to perceive.

In summary, with the combination of stimulus design and analysis method we provide a novel way to investigate neuronal information processing: tagging stimuli by broadband temporal noise enables to track the information flow through cortex and to map transmission for all frequencies simultaneously. This method allows to investigate the linear components of signal gating, and to characterize the system response at different processing stages simultaneously, thus providing an elegant probing technique for sensory, but importantly also optogenetic stimulation. Taking the experimental and simulation data together, our findings support the concept that attention can gate information flow already at the level of afferent synaptic input, thereby possibly explaining previous studies' attention-associated firing rates changes.

## References

- Battaglia D, Witt A, Wolf F, Geisel T (2012) Dynamic effective connectivity of inter-areal brain circuits. *PLoS Comput Biol* 8:e1002438. [CrossRef Medline](#)
- Bichot NP, Rossi AF, Desimone R (2005) Parallel and serial neural mechanisms for visual search in macaque area V4. *Science* 308:529–534. [CrossRef Medline](#)
- Bosman CA, Schoffelen JM, Brunet N, Oostenveld R, Bastos AM, Womelsdorf T, Rubehn B, Stieglitz T, De Weerd P, Fries P (2012) Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron* 75:875–888. [CrossRef Medline](#)
- Buia CI, Tiesinga PH (2008) Role of interneuron diversity in the cortical microcircuit for attention. *J Neurophysiol* 99:2158–2182. [CrossRef Medline](#)
- Chelazzi L, Duncan J, Miller EK, Desimone R (1998) Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol* 80:2918–2940. [CrossRef Medline](#)
- Deco G, Rolls ET (2005) Neurodynamics of biased competition and cooperation for attention: a model with spiking neurons. *J Neurophysiol* 94:295–313. [CrossRef Medline](#)
- Desimone R, Schein SJ (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurophysiol* 57:835–868. [CrossRef Medline](#)
- Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci* 9:474–480. [CrossRef Medline](#)
- Fries P (2009) Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu Rev Neurosci* 32:209–224. [CrossRef Medline](#)
- Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291:1560–1563. [CrossRef Medline](#)
- Ghose GM (2009) Attentional modulation of visual responses by flexible input gain. *J Neurophysiol* 101:2089–2106. [CrossRef Medline](#)
- Ghose GM, Maunsell JH (2008) Spatial summation can explain the attentional modulation of neuronal responses to multiple stimuli in area V4. *J Neurosci* 28:5115–5126. [CrossRef Medline](#)
- Grothe I, Neitzel SD, Mandon S, Kreiter AK (2012) Switching neuronal inputs by differential modulations of gamma-band phase-coherence. *J Neurosci* 32:16172–16180. [CrossRef Medline](#)
- Hamker FH (2004) Predictions of a model of spatial attention using sum- and max-pooling functions. *Neurocomputing* 56:329–343. [CrossRef](#)
- Harnack D, Ernst UA, Pawelzik KR (2015) A model for attentional information routing through coherence predicts biased competition and multistable perception. *J Neurophysiol* 114:1593–1605. [CrossRef Medline](#)
- Ito M, Gilbert CD (1999) Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron* 22:593–604. [CrossRef Medline](#)
- Jensen O, Kaiser J, Lachaux JP (2007) Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci* 30:317–324. [CrossRef Medline](#)
- Kastner S, De Weerd P, Desimone R, Ungerleider LG (1998) Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 282:108–111. [CrossRef Medline](#)
- Keitel C, Andersen SK, Quigley C, Müller MM (2013) Independent effects of attentional gain control and competitive interactions on visual stimulus processing. *Cereb Cortex* 23:940–946. [CrossRef Medline](#)
- Kreiter AK (2006) How do we model attention-dependent signal routing? *Neural Netw* 19:1443–1444. [CrossRef Medline](#)
- Kronlandt-Martinet R, Morlet J, Grossmann A (1987) Analysis of sound patterns through wavelet transforms. *Int J Pattern Recognit Artif Intell* 1:273–302. [CrossRef](#)
- Lee J, Maunsell JH (2009) A normalization model of attentional modulation of single unit responses. *PLoS One* 4:e4651. [CrossRef Medline](#)
- Lee J, Maunsell JH (2010) Attentional modulation of MT neurons with single or multiple stimuli in their receptive fields. *J Neurosci* 30:3058–3066. [CrossRef Medline](#)
- Lee J, Williford T, Maunsell JH (2007) Spatial attention and the latency of neuronal responses in macaque area V4. *J Neurosci* 27:9632–9637. [CrossRef Medline](#)
- Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol* 77:24–42. [CrossRef Medline](#)
- Masuda N (2009) Selective population rate coding: a possible computational role of gamma oscillations in selective attention. *Neural Comput* 21:3335–3362. [CrossRef Medline](#)
- Maunsell JH, Treue S (2006) Feature-based attention in visual cortex. *Trends Neurosci* 29:317–322. [CrossRef Medline](#)
- McAdams CJ, Maunsell JH (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19:431–441. [Medline](#)
- Mehta AD, Ulbert I, Schroeder CE (2000) Intermodal selective attention in monkeys. I: distribution and timing of effects across visual areas. *Cereb Cortex* 10:343–358. [CrossRef Medline](#)
- Middendorff M, McMillan G, Calhoun G, Jones KS (2000) Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Trans Rehabil Eng* 8:211–214. [CrossRef Medline](#)
- Mishra J, Fellous JM, Sejnowski TJ (2006) Selective attention through phase relationship of excitatory and inhibitory input synchrony in a model cortical neuron. *Neural Netw* 19:1329–1346. [CrossRef Medline](#)
- Moran J, Desimone R (1985) Selective attention gates visual processing in extrastriate cortex. *Science* 229:782–784. [CrossRef Medline](#)
- Morgan ST, Hansen JC, Hillyard SA (1996) Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proc Natl Acad Sci U S A* 93:4770–4774. [CrossRef Medline](#)
- Motter BC (1993) Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol* 70:909–919. [CrossRef Medline](#)
- Motter BC (2009) Central V4 receptive fields are scaled by the V1 cortical magnification and correspond to a constant-sized sampling of the V1 surface. *J Neurosci* 29:5749–5757. [CrossRef Medline](#)
- Müller MM, Malinowski P, Gruber T, Hillyard SA (2003) Sustained division of the attentional spotlight. *Nature* 424:309–312. [CrossRef Medline](#)
- Ni AM, Ray S, Maunsell JH (2012) Tuned normalization explains the size of attention modulations. *Neuron* 73:803–813. [CrossRef Medline](#)
- Palmigiano A, Geisel T, Wolf F, Battaglia D (2017) Flexible information routing by transient synchrony. *Nat Neurosci* 20:1014–1022. [CrossRef Medline](#)
- Reynolds JH, Heeger DJ (2009) The normalization model of attention. *Neuron* 61:168–185. [CrossRef Medline](#)

- Reynolds JH, Chelazzi L, Desimone R (1999) Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci* 19:1736–1753. [CrossRef Medline](#)
- Rock I, Gutman D (1981) The effect of inattention on form perception. *J Exp Psychol Hum Percept Perform* 7:275–285. [CrossRef Medline](#)
- Rock I, Linnett CM, Grant P, Mack A (1992) Perception without attention: results of a new method. *Cognit Psychol* 24:502–534. [CrossRef Medline](#)
- Roelfsema PR, Lamme VA, Spekreijse H (1998) Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395:376–381. [CrossRef Medline](#)
- Rotermund D, Taylor K, Ernst UA, Kreiter AK, Pawelzik KR (2009) Attention improves object representation in visual cortical field potentials. *J Neurosci* 29:10120–10130. [CrossRef Medline](#)
- Rotermund D, Ernst UA, Mandon S, Taylor K, Smiyukha Y, Kreiter AK, Pawelzik KR (2013) Toward high performance, weakly invasive brain computer interfaces using selective visual attention. *J Neurosci* 33:6001–6011. [CrossRef Medline](#)
- Sokolov A, Lutzenberger W, Pavlova M, Preissl H, Braun C, Birbaumer N (1999) Gamma-band MEG activity to coherent motion depends on task-driven attention. *Neuroreport* 10:1997–2000. [CrossRef Medline](#)
- Spratling MW, Johnson MH (2004) A feedback model of visual attention. *J Cogn Neurosci* 16:219–237. [CrossRef Medline](#)
- Sripati AP, Johnson KO (2006) Dynamic gain changes during attentional modulation. *Neural Comput* 18:1847–1867. [CrossRef Medline](#)
- Taylor K, Mandon S, Freiwald WA, Kreiter AK (2005) Coherent oscillatory activity in monkey area V4 predicts successful allocation of attention. *Cereb Cortex* 15:1424–1437. [CrossRef Medline](#)
- Tiesinga PH, Sejnowski TJ (2010) Mechanisms for phase shifting in cortical networks and their role in communication through coherence. *Front Hum Neurosci* 4:196. [CrossRef Medline](#)
- Treue S, Maunsell JH (1996) Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* 382:539–541. [CrossRef Medline](#)
- Treue S, Martínez Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399:575–579. [CrossRef Medline](#)
- VanRullen R, Macdonald JS (2012) Perceptual echoes at 10 Hz in the human brain. *Curr Biol* 22:995–999. [CrossRef Medline](#)
- Wolfe JM, Bennett SC (1997) Preattentive object files: shapeless bundles of basic features. *Vision Res* 37:25–43. [CrossRef Medline](#)
- Womelsdorf T, Schoffelen JM, Oostenveld R, Singer W, Desimone R, Engel AK, Fries P (2007) Modulation of neuronal interactions through neuronal synchronization. *Science* 316:1609–1612. [CrossRef Medline](#)
- Yoshor D, Ghose GM, Bosking WH, Sun P, Maunsell JH (2007) Spatial attention does not strongly modulate neuronal responses in early human visual cortex. *J Neurosci* 27:13205–13209. [CrossRef Medline](#)