

Spatially-Specific Attentional Modulation Revealed by fMRI

David C. Somers and Stephanie A. McMains

ABSTRACT

Functional MRI and visual psychophysics were employed to investigate space-based attentional selection mechanisms in human occipital cortex. Our 1999 findings, along with nearly simultaneous findings from other laboratories, demonstrated that spatially specific attention could operate robustly even in primary visual cortex. Spatial deployment of attention operates in a “push-pull” fashion, both increasing responses at attended locations and decreasing responses at nonattended locations. Parametric studies suggest that spatial attention acts primarily as an “additive bias” signal, whose amplitude is largely independent of stimulus strength. Finally, the spatial window of attention, as reflected in retinotopic cortical activation, is highly flexible and it may be split into multiple “spotlights,” if the task so demands.

I. INTRODUCTION

Complex visual environments bombard our eyes with more information than our cognitive systems can act on at one time. Although the retina performs massively parallel processing, visual cognition operates on no more than a few items at once. Given these limitations, what we do or do not perceive is largely determined by attentional mechanisms that select information for enhanced cognitive processing. William James proposed that visual attention acts like a spotlight, selecting a single region of visual space, while largely neglecting the rest of the visual field. The “spotlight” model, further refined by Eriksen (Eriksen and St. James, 1986) and others as a “zoom lens” mechanism that selects one contiguous, convex spatial

window, has successfully addressed a wealth of psychophysical data. This article summarizes functional MRI-based research explorations of the mechanisms of spatial attention in human occipital cortex, including primary visual cortex.

II. ATTENTIONAL MODULATION IN VI AND OTHER EARLY VISUAL CORTICAL AREAS

Our initial fMRI experiments began with the question of how visual conical processing of a moving stimulus is influenced by attention (Somers et al., 1999a). These experiments depended critically on the ability to distinguish the visual cortical representations of different parts of the visual field (e.g., (Serenio et al., 1995) see Figure 62.1D). The cortical representation of the center of the visual field (or fovea) occupies a large swath of cortex in the center of the flattened occipital cortex patches (see Figure 62.1C). More peripheral eccentricities are represented as horizontal bands above and below the central region. The stimulus display was configured to exploit the eccentricity bias of the cortical map. A two-part stimulus display was utilized (see Figure 62.1A). It consisted of an annular region containing a grating pattern, which rotated either clockwise or counterclockwise on a given trial, and a central disk in which letters were displayed in a rapid serial visual presentation (RSVP) format. In this annulus-disk configuration, the RSVP letters would drive the central eccentricity band, and the motion annulus would drive the horizontal bands above and below the central region.

Subjects held central fixation on this two-part display and directed their attention to one portion or the other. Eye position measurements performed in the

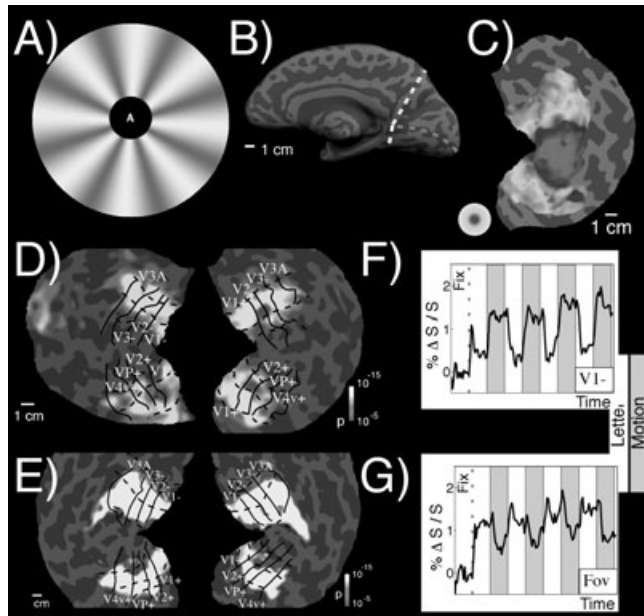


FIGURE 62.1 Attentional modulation of striate and extrastriate visual cortex. **A.** Visual stimuli were composed of an annulus with rotating radial wedge patterns and a central target that was either a fixation point or single letters in an RSVP stream. **B.** The medial view of a mathematically inflated cortical surface revealing the buried sulci. The posterior portion is cut off and cut along the calcarine sulcus. The resulting patch is flattened for data visualization. **C.** Functional mapping of visual eccentricity, with the foveal representation in the center of the flattened patch. **D** and **E.** Patterns of statistically significant increased activation for attend extrafoveal motion vs. attend foveal letters for both hemispheres for two subjects, extending across all labeled retinotopic areas. **F** and **G.** Time course data for the fovea and peripheral V1, averaged across subjects for attend foveal letters vs. attend extrafoveal motion.

scanner confirmed that subjects could hold central fixation during these experiments. Subjects performed alternating blocks of trials in which they either judged the rotation direction of the motion annulus or identified five consecutive letters appearing in the central RSVP stream. The central task was designed to be highly demanding in order to strongly draw attention away from the motion annulus. Comparison of fMRI activation between the two conditions revealed spatially specific attentional modulations across all early visual cortical areas (V1, V2, V3, VP, V3A, V4v). When attention was directed to the motion annulus, activation increased in the iso-eccentricity bands corresponding to the cortical representation of the annulus (see Figure 62.1D, E, F). When attention was directed to the fovea, fMRI activation increased in the cortical representation of the fovea (see Figure 62.1G). These findings were also confirmed with a comparison of equivalent tasks, in which the RSVP stream was

replaced by a motion disk stimulus that rotated independently of the annulus.

Two features of these results were remarkable. First, the attentional modulations closely followed the iso-eccentricity bands and thus attentional effects were strongly spatially selective. Perhaps the more surprising observation was that attention robustly modulated primary visual cortex (see Figure 62.1F). These attentional modulations in area V1 were as large as 1% signal change in some subjects. Overall, the attentional modulations alone comprised 60% of the combined stimulus plus attention activation (compared to passive viewing of a blank fixation display). This implies that the attentional modulation amplitude was of the same order as the pure stimulus drive activation and was perhaps even greater than the stimulus drive. The conventional wisdom had held that primary visual cortex was a “preprocessing” stage that was “cognitively impenetrable.” The seminal primate electrophysiology studies of Moran and Desimone in 1985 had observed attentional modulations in area V4, but not V1. As late as 1998, a leading research group predicted that attention could not operate in area V1. Our findings flew in the face of this conventional wisdom and were supported by nearly simultaneous reports of V1 attention from other fMRI laboratories using widely different stimuli and task paradigms (see Chapter 61) (Tootell et al., 1998; Watanabe et al., 1998; Gandhi et al., 1999; Martinez et al., 1999). This wave of studies resulted in the quick acceptance of the V1 attention finding. Thus, the V1 attention result is an example of a result that was widely accepted on the basis of human fMRI data before it was embraced on the basis of primate physiology. Motter (Motter, 1993) had reported V1 attentional effects in monkeys, but this finding was questioned due to concerns that the effects might be due to small eye movement artifacts. The large areas of cortex modulated in the fMRI studies could not be explained away as an eye movement effect. Ironically, the relatively coarse spatial resolution of fMRI was an important asset in establishing this finding.

III. MECHANISMS OF SPATIAL ATTENTION SELECTION

Further experiments were aimed at determining whether the observed attentional modulations reflected an enhancement of activity at the attended locations, suppression of activity at the unattended locations, or both (Somers et al., 1999a). In the initial experiments this issue could not be resolved since

there was a direct comparison of attention directed to two spatially complementary regions. This issue was addressed by performing experiments in which each attentional state (attend center, attend periphery) was paired with a condition in which the stimulus was passively viewed. The passive viewing condition represents the attentional baseline condition. Each experiment makes complementary predictions about the modulations to be observed in the two spatial regions. In the comparison of Attend Periphery vs. Passive Viewing, modulations observed in the periphery would be interpreted as attentional enhancements, whereas modulations observed in the fovea region would be interpreted as attentional suppressions. Conversely, in the “Attend Center” vs. Passive Viewing comparison, peripheral modulations would reflect suppression and central modulations would reflect enhancement. In both sets of experiments, attentional modulations were observed in both spatial regions (in antiphase relationship), implying that spatial attention acts in a “push-pull” manner, increasing responses at the cortical representations of the attended locations and diminishing responses at the nonattended cortical representations (see Figure 62.2A) (see Chapters 71 and 50).

Many of the stimulus response properties of visual cortical neurons have been extremely well characterized in electrophysiology studies. Our next question was as follows: By what computational means does attention modulate these stimulus responses? Does attention have a multiplicative effect, an additive effect, or some other form of influence on stimulus responses? We performed parametric fMRI studies in order to reveal these computational mechanisms. Contrast was the stimulus property varied. Contrast Response Functions (CRFs) have been well-established in psychophysics and single-unit physiology. As stimulus contrast increases, responses increase monotonically over a broad range of contrasts. Boynton et al. (1996) extended this CRF result to fMRI. In our experiments (Somers et al., 1999b), we constructed two fMRI CRFs using seven different contrast values, each with attention covertly directed to the stimulus and with attention directed away from the stimulus in alternate blocks. Comparison of these two CRFs will reveal how attention interacts with stimulus contrast responses (see Figure 62.2C).

Three hypotheses were proposed. Attention *might* multiplicatively “gate” stimulus responses. The multiplicative model predicts that the effects of attention should increase as the stimulus drive (and contrast) increase. Attention *might* provide an additive bias signal to the attended locations. The additive model predicts that attention effects should be constant across

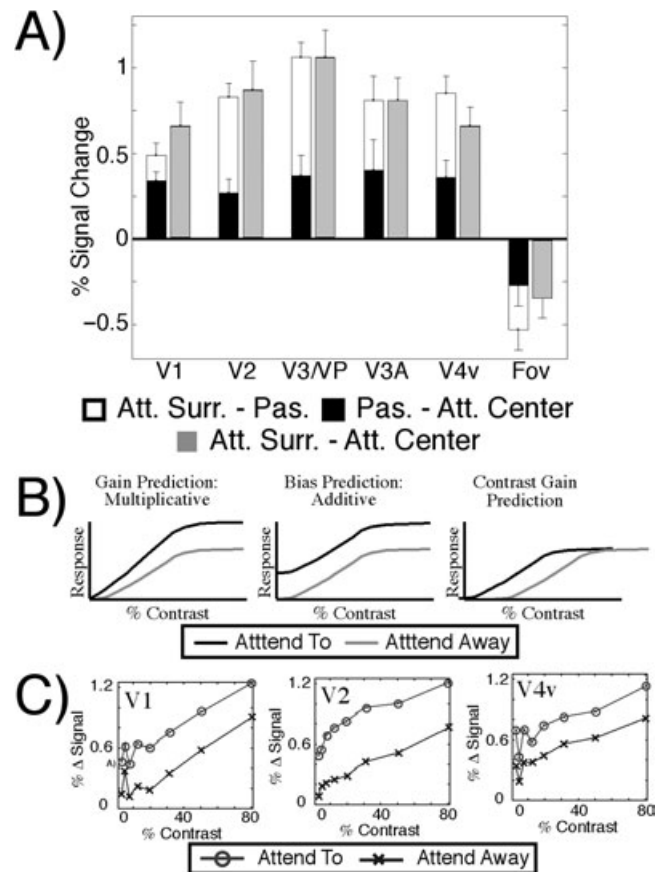


FIGURE 62.2 Mechanisms of attention. A. Average attentional modulation amplitudes for attend conditions vs. passive viewing, demonstrating the push-pull manner of spatial attention. B. Three hypothesis for how attention might interact with stimulus contrast responses. C. Contrast response functions for three visual areas, with and without attention, supporting the additive bias model.

all stimulus drive (and contrast) levels. Attention *might* increase the contrast gain, making a stimulus effectively appear as if it were at a higher contrast. This predicts that attention should shift the CRF to the left and that attentional effects should be strongest for the middle range of stimulus drives/contrasts (see Figure 62.2B).

In these experiments, the “with attention” and “without attention” CRFs were observed to differ primarily by a constant factor across all stimulus contrasts (see Figure 62.2C). This result supports the additive bias model prediction and is inconsistent with the multiplicative gain model. In addition to the dominant additive effect, the attentional effects exhibit a modest peak at the middle contrast ranges, consistent with some weak contribution from a contrast gain mechanism. However, since fMRI CRFs do not saturate at high contrasts (while single unit CRFs do saturate), the fMRI results are not well described by a

leftward shift of the CRF or by the contrast gain modulation hypothesis.

The additive bias finding is somewhat surprising. Adding a constant factor to a response actually *decreases* the signal-to-noise ratio at that location. If perceptual decisions were computed locally this would be a destructive influence of attention. However, if perceptual decisions were computed upstream in a higher cortical area and different locations had to compete for access to this circuitry then an additive signal would be a constructive influence as it would bias the competition toward that location. This is highly consistent with the predictions of the Biased Competition model of attention (Desimone, 1998) (see Chapter 50). The finding of an “attentional pedestal” of activity that is independent of stimulus drive (see Chapter 51) appears consistent with Kastner et al.’s fMRI finding that the expectation of the appearance of a stimulus can yield increased activation even in the absence of a stimulus (see Chapter 50). However, these findings differ substantially from a primate study of attentional effects on CRFs in area V4 (Reynolds et al., 2000), which reported that the primary effect was a leftward shift of the CRF1 (contrast gain hypothesis) (see Chapter 70). A small additive effect was also reported. It is difficult to reconcile these two sets of findings in different species other than to suggest that single-unit recordings reflect activation from a tiny minority of the neurons reflected in the fMRI recordings. Perhaps, a small set of highly responsive neurons exhibit robust contrast gain modulation effects under directed attention, whereas a much larger population of more weakly responsive neurons exhibits an additive pedestal influence of spatial attention. Without a more complete understanding of the workings of the neural code, it is difficult to know which neural population plays a greater role in constructing perception. Hopefully, future work will better reconcile the differences between human fMRI and primate electrophysiology studies.

IV. FLEXIBLE SPATIAL WINDOWS OF ATTENTION

William James suggested that spatial attention acts like a “spotlight,” selecting what is in its beam and filtering out what falls outside the spotlight. This “spotlight” metaphor was further developed into the “zoom lens” model (Eriksen and St. James, 1986). The zoom lens model suggests that spatial attention is flexible in that it may be narrowed or expanded (and distorted) to suit task demands. However, the zoom lens model also constrains attention to act over a single contigu-

ous region of space. This model appears too limiting in that some tasks may require attention to separated regions of space. Other tasks may benefit from excluding strong distracting stimuli from the window of selection. Some psychophysical studies have suggested that the window of spatial attention may indeed be more flexible than the zoom lens model suggests (e.g., Awh and Pashler, 2000).

We performed a set of experiments that overcame these limitations and demonstrate that spatial attention can be split into multiple distinct “spotlights” of activation in area V1 and other human occipital cortical regions (McMains and Somers, 2004). Since fMRI studies are limited by their temporal resolution and since spatial attention is capable of moving as fast as every 200–500 ms (e.g., Weichselgartner and Sperling, 1987), it is extremely difficult to rule out the movement of spatial attention as a factor in the pattern of activation observed if one relies purely on the fMRI data. A central feature of these experiments was to employ a psychophysical task that excluded the possibility that spatial attention was rapidly switching between locations of interest.

Subjects were required to compare the identity of targets simultaneously displayed in two separated locations. The targets appeared in RSVP streams and thus were visible only briefly (173 ms) and were masked immediately by another stimulus. The targets were digits appearing in RSVP streams of letters (use of letter targets did not change the psychophysical findings) and there were no other overt cues to the appearance of a target. Subjects were required to report whether the target digits were the same or different. This task can be performed only at above chance levels if both targets are identified. These two RSVP streams were imbedded in a display consisting of five RSVP streams, a central one surrounded by four equidistant peripheral streams, one per visual field quadrant (see Figure 62.3A). The central RSVP stream lay directly inbetween the two attended RSVP streams, which were placed in opposing visual field quadrants. In order to make the central stream highly distracting, only digits appeared in this stream. If information appearing in the central stream were selected along with the information in the two peripheral streams of interest, it likely would interfere with performance of the digit comparison task. Thus for this task, it was advantageous to avoid selecting the central region while selecting the two peripheral streams. Psychophysical performance in this task was investigated by parametrically varying the letter presentation duration of the RSVP streams. Threshold level performance ($d' = 1$) was observed at letter durations of 67 ms. This rate is much faster than the minimum estimates for

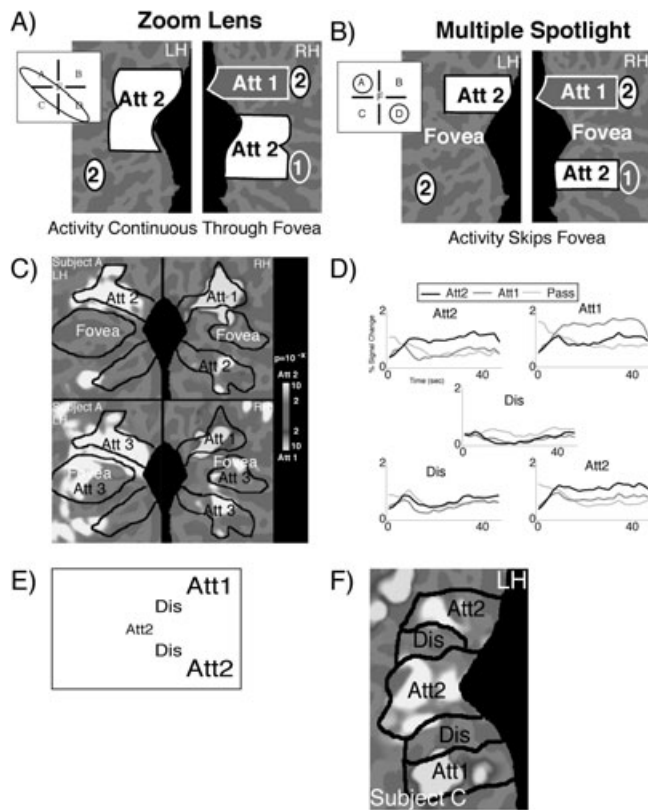


FIGURE 62.3 Multiple spotlights of spatial attention. **A** and **B**. Predicted activation patterns for Attend2 vs. Attend1 for the two competing hypotheses. **C**. Activation patterns for Attend2 vs. Attend1, for two subjects, supporting the multiple spotlight hypothesis. **D**. Time course data averaged across all subjects. **E**. Spatial layout for hemifield experiment. **F**. Activation maps for one subject in the right hemifield configuration comparing Attend2 vs. Attend1.

how quickly spatial attention can select a target, deploy to a new location, and select a second target (200–500 ms). In addition to the Attend2 task, an Attend1 condition was included in which attention was covertly directed only to a (third) peripheral RSVP stream, while viewing the same display as that used in the Attend2 condition. In the Attend1 task subjects had to identify digits appearing in the stream and report whether they matched a predefined target digit. Although the Attend1 task was somewhat easier than the Attend2 task, performance differences were far less than would be predicted if the Attend2 task were done by purely serial processing. The threshold in the Attend1 task was 59 ms, only 8 ms shorter than for the Attend2 task. These modest task performance differences were observed across all RSVP speeds of interest. Thus the psychophysical results demonstrate that spatial attention was simultaneously deployed to the two RSVP streams of interest.

Having thus ruled out the “rapid switching” hypothesis for this task, two hypotheses then remained that could explain selection of the two peripheral streams. The zoom lens model predicts that attention stretches to select the two RSVP streams and necessarily also selects the distracting central RSVP stream (Figure 62.3A). The alternate hypothesis suggests that the spatial window of attention can be split into two distinct spotlights that select the two RSVP streams of interest while filtering out the central stream (Figure 62.3B). This question was addressed using fMRI of occipital lobe activation during task performance. Comparison of Attend2 vs. Attend1 activation (see Figure 62.3C, D) revealed two hotspots of activation at the visual cortical representations of the two Attend2 streams. Critically, the cortical representation of the central RSVP stream was not activated (nor was it activated in comparisons with the passive viewing condition). This result demonstrates that the window of spatial attention may be split to select multiple distinct regions. This pattern was observed for all subjects. Group analysis revealed that this split attentional spotlight effect occurred in all “early” visual cortical areas, including primary visual cortex. As expected, the Attend1 condition (vs. Attend2) revealed a single hotspot of activation at the cortical representation of the attended RSVP stream (see Figure 62.3C, D).

We repeated these experiments using a different spatial configuration of RSVP streams (see Figure 62.3E). All streams were placed in one hemifield and the two RSVP streams of interest in the Attend2 task were placed in one visual field quadrant—one at the fovea and one in the periphery. The critical distracter was placed inbetween at a mid-eccentricity. The Attend1 RSVP stream was placed in the other visual quadrant. The fMRI results again revealed two hotspots of activation corresponding to the locations of the two targets, with a sparing of the intermediate region (see Figure 62.3F).

Our results demonstrate that attention may be split between targets located in opposite hemifields and thus split across different cortical hemispheres. We also observed splitting of the spotlight within a visual quadrant and within a cortical hemisphere. The finding that the foveal representation may either be included or excluded as one focus of attention demonstrates that attention may be divided between overt and covert targets or divided between two covertly monitored targets. In these regards, the ability to divide spatial attention exhibits remarkable flexibility.

In hindsight, the requirement that spatial attention be a serial process seems unnecessary (see Chapter 97). The capacity for parallel processing is an implicit

feature of spatial representations. Once spatial attention was demonstrated to operate in early visual cortical areas (see Chapter 61) (Tootell et al., 1998; Watanabe et al., 1998; Gandhi et al., 1999; Martinez et al., 1999; Somers et al., 1999a), the general requirement for a single spotlight of spatial attention was ready to fall. These cortical areas contain primarily neurons with spatially limited receptive fields (RFs), so multiple targets need not compete for resources at these stages, provided that targets are far enough apart. Even in higher visual cortical areas, where attention shrinks down the normally large RFs to isolate single targets (Desimone, 1998), attentional selection may divide the pool of relevant neurons between multiple targets rather than assigning them in a “winner-take-all” fashion to one target location. This strategy might result in some cost for attending to multiple targets but would not necessarily result in a complete failure to select some locations. A modestly parallel spatial attention system of this form appears much simpler to implement physiologically than does a rapidly switching serial spatial attention system.

V. SUMMARY

fMRI is a powerful technique for revealing how spatial attention is deployed across the cortical representations of the visual field. fMRI has helped to reveal that spatial attention may have robust effects in primary visual cortex and that attention may be divided between multiple, distinct, spatial windows or “spotlights.” Parametric fMRI studies hold the promise of a greater understanding of the computational mechanisms of attention. Our studies revealed an additive push-pull bias signal that is consistent with the Biased Competition hypothesis (Desimone, 1998).

References

- Awh, E., and Pashler, H. (2000). Evidence for split attentional foci. *J. Exp. Psychol. Hum. Percept. Perform.* **26**, 834–846.
- Boynton, G. M., Engel, S. A., Glover, G. H., and Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *J. Neurosci.* **16**, 4207–4221.
- Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **353**, 1245–1255.
- Eriksen, C. W., and St. James, J. D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. *Percept. Psychophys.* **40**, 225–240.
- Gandhi, S. P., Heeger, D. J., and Boynton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proc. Natl. Acad. Sci. USA* **96**, 3314–3319.
- Martinez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., Wong, E. C., Hinrichs, H., Heinze, R. I., and Hillyard, S. A. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nat. Neurosci.* **2**, 364–369.
- McMains, S. A., and Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron.* **42**, 677–686.
- Motter, B. C. (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.
- Reynolds, J. H., Pasternak, T., Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron.* **26**, 703–714.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R., and Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science.* **268**, 889–893.
- Somers, D. C., Dale, A. M., Seiffert, A. E., and Tootell, R. B. (1999a). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proc. Natl. Acad. Sci. USA* **96**, 1663–1668.
- Somers, D. C., Seiffert, A. E., Dale, A. M., and Tootell, R. B. (1999b). Effects of task difficulty and stimulus contrast on attentional modulation in human striate and extrastriate cortex revealed by fMRI. In: *Society for Neuroscience*, 6.10.
- Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., and Dale, A. M. (1998). The retinotopy of visual spatial attention. *Neuron.* **21**, 1409–1422.
- Watanabe, T., Harner, A. M., Miyauchi, S., Sasaki, Y., Nielsen, M., Palomo, D., and Mukai, I. (1998). Task-dependent influences of attention on the activation of human primary visual cortex. *Proc. Natl. Acad. Sci. USA* **95**, 11489–11492.
- Weichselgartner, E., and Sperling, U. (1987). Dynamics of automatic and controlled visual attention. *Science.* **238**, 778–780.